IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

The Bell phenomenon in classical frameworks

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1996 J. Phys. A: Math. Gen. 29 247

(http://iopscience.iop.org/0305-4470/29/2/005)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 171.66.16.68
The article was downloaded on 02/06/2010 at 02:39

Please note that terms and conditions apply.

J. Phys. A: Math. Gen. **29** (1996) 247–261. Printed in the UK

The Bell phenomenon in classical frameworks

E G Beltrametti† and S Bugajski‡

† Department of Physics, University of Genoa and Istituto Nazionale di Fisica Nucleare, Sezione di Genova, Via Dodecaneso 33, I-16146 Genova, Italy

‡ Institute of Physics, University of Silesia, Uniwersytecka 4, PL 40007 Katowice, Poland

Received 10 July 1995

Abstract. The experimentally tested occurrence of probabilities that do not find a representation within the usual Kolmogorov probability theory is mathematically formalized through the notion of the Bell phenomenon for probability measures and for observables. Reference is made to a physically natural definition of observables that includes both the classical and the quantum usual versions. It is shown that even a classical framework can host the Bell phenomenon, provided fuzzy observables are called into play.

1. Introduction

Since the 1964 paper of Bell [1] it has been clear that quantum phenomena can give rise to probabilistic behaviours that cannot be viewed as merely expressing some degree of ignorance inside a standard deterministic frame, such as a 'non-contextual' local hidden-variable theory. The growth of interest on the issue of Bell inequalities made it also clear that, quite independently from the specific physical content of the (Bohm version of the) Einstein–Podolski–Rosen correlation, one can tackle the more general mathematical problem of the conditions that allow the representation of a given set of probabilities inside the standard Kolmogorov probability theory. The preferred form of these conditions is still in terms of requiring some suitable linear combination of the given probabilities to lie inside a numerical interval, which are what we may call Bell-type inequalities [2–4]. Quantum mechanics provides a variety of situations in which probabilities arise which do not satisfy these conditions, thus escaping a representation within the standard Kolmogorov probability theory. It is this feature that, loosely speaking, will be captured by what we shall define in the following as the 'Bell phenomenon'.

The description of a physical system can be based on the structure of the convex set S of its states, an observable being defined as an affine map of S into the convex set of the probability measures on the measurable space which collects the possible values of that observable. Indeed the physical notion of an observable corresponds to specifying its possible outcomes and the probabilities of their occurrence, for each state of the physical system. The classical or quantum nature of a model is coded in the convex structure of S, which is a simplex in the classical case while it cannot be such in the quantum one, in view of the well known phenomenon of the non-unique decomposition of quantum mixtures into pure states. In sections 3 and 5 the notion of the Bell phenomenon will be made precise with reference to probability measures on measurable spaces and with reference to the mentioned definition of observables. Section 4 will deal with the particular case of probability measures

0305-4470/96/020247+15$19.50 © 1996 IOP Publishing Ltd

on spaces consisting of just two points, a framework that allows a comparison of the Bell phenomenon with similar notions occurring in the literature.

The Bell phenomenon appears historically as a distinguishing feature of the quantum context; no analogue is generally acknowledged in classical frameworks. Could it be hosted in a classical context? This question takes a particular motivation in view of the fact that quantum mechanics admits an extension of classical nature, as discussed in [5, 6], and briefly reviewed in the next section. The extended classical model has a wider set of observables, but each quantum observable finds its representative in the extended model and the relevant quantum features are preserved in the classical extension. As shown in section 6 the Bell phenomenon is carried into the classical context, but fuzzy observables are called into play.

2. State-observable structure

In this section we summarize, without proofs, a number of results given in [5, 6].

Let S be the convex set formed by the states of the physical system under discussion. It is natural to define an observable as an affine map of S into the convex set $M_1^+(\Xi)$ of all the probability measures on some measurable space Ξ in which the observable takes values (typically the real line). Indeed, the notion of observable carries the specification of the possible measurement results and of their probability distribution for each state of the physical system. With some abuse of notation we write Ξ for a measurable space in place of $(\Xi, \mathcal{B}(\Xi))$ where $\mathcal{B}(\Xi)$ stands for the σ-algebra of measurable subsets of Ξ, and we shall assume that the one-point subsets of Ξ are measurable. This definition of observable is not new (see, e.g., [7–12]) and encompasses the standard definitions used in quantum and in classical mechanics. For a fixed $X \in \mathcal{B}(\Xi)$, the pair formed by X and by an observable B determines an affine function $E_{B,X} : S \to [0, 1]$ defined by $E_{B,X}(\alpha) := (B\alpha)(X)$, $\alpha \in S$: it is called an effect.

In the quantum case S is the set S_{Q} of all density operators on a separable Hilbert space $\mathcal{H}$. Typical of S_{Q} is the non-unique decomposition of mixed states into pure ones, which mirrors the fact that S_{Q} is not a simplex. The effects on S_{Q} are known [12] to be in one-to-one correspondence with the positive operators of $\mathcal{H}$ which have mean value not bigger than 1: explicitly, if $\mathcal{P}$ is such an operator and $D \in S_{\mathrm{Q}}$ then the function $\mathrm{Tr}(D\mathcal{P}) : S_{\mathrm{Q}} \to [0, 1]$ is the effect associated with $\mathcal{P}$. Thus, our notion of observable, when referred to S_{Q}, gives back the so-called positive operator valued (POV) measures, which are typical ingredients of the 'operational' approach to quantum mechanics: we also quote the recent volume [13] for reference to the relevant literature. Notice that when positive operators are restricted to projectors and the measurable space Ξ is specified as the real line $\mathbb{R}$ then the POV measures become the familiar PV measures on $\mathbb{R}$, hence they correspond to the self-adjoint operators of $\mathcal{H}$ (via the spectral theorem). Therefore, in the case of Hilbert-space quantum mechanics, the definition of observable we are adopting here is fully equivalent to the operational one, and it includes as a special case the traditional observables represented by self-adjoint operators.

In the classical case the set S of states takes the typical structure of the set $M_1^+(\Omega)$ of all the probability measures on the 'phase space' Ω of the physical system. Of course Ω is understood as a measurable space, and we additionally assume that the one-point subsets of Ω are measurable. The measures concentrated at one point of Ω (the Dirac measures) correspond to the pure states and the unique decomposability of mixed states into pure states becomes the distinguishing feature that mirrors the simplex nature of the convex set $M_1^+(\Omega)$. According to our definition, an observable taking values in the measurable space Ξ is now an affine map of $M_1^+(\Omega)$ into $M_1^+(\Xi)$. The observable $B : M_1^+(\Omega) \to M_1^+(\Xi)$

will be called *regular* if

$$B\mu = \int_\Omega (B\delta_\omega)\,\mathrm{d}\mu(\omega)$$

for any $\mu \in M_1^+(\Omega)$, where δ_ω denotes the Dirac measure concentrated at $\omega \in \Omega$, and the integral is understood in the sense that

$$(B\mu)(X) = \int_\Omega (B\delta_\omega)(X)\,\mathrm{d}\mu(\omega)$$

for any measurable subset X of Ξ. This equality can be read by saying that the value of the effect $E_{B,X}$ at the state μ is the integral, with respect to the measure μ, of the (measurable) function $(B\delta_\omega)(X) : M_1^+(\Omega) \to [0,1]$. This way of reading the above equality shows the equivalence between the notion of regularity given here and the one used in [6].

In the following we shall be mainly concerned with regular observables. Among them, one has to outline the particular family of the observables that capture a notion of sharpness. Following a traditional terminology we shall say (as in [6] and [13]) that an observable is sharp when its effects are extreme elements of the convex set of all effects. In this respect it is worth remarking that the usual observables employed in standard classical mechanics and in classical probability theory are regular observables generated by functions from Ω into Ξ in the sense that for any such observable B there is a unique function $f_B : \Omega \to \Xi$ such that $B\delta_\omega = \delta_{f_B(\omega)}$ for every $\omega \in \Omega$. In other words, the value of the effect $E_{B,X}$ at the state μ takes the form

$$(B\mu)(X) = \int_\Omega \delta_{f_B(\omega)}\,\mathrm{d}\mu(\omega) = \int_\Omega \chi_{f_B^{-1}(X)}(\omega)\,\mathrm{d}\mu(\omega) = \mu(f_B^{-1}(X)) \qquad (1)$$

where χ_Y is the characteristic function of $Y \in \mathcal{B}(\Xi)$. An observable like that is sharp (see theorem 1 of [6]) and one may wonder whether any regular sharp observable must have this form. The answer is yes [14, 15], provided one makes the additional assumption that the extreme points of $M_1^+(\Xi)$ coincide with the Dirac measures on Ξ. Actually this assumption is met by any measurable space Ξ contained in $\mathbb{R}^n$ (n an arbitrary positive integer) with the σ-algebra of Borel sets, hence in all cases of physical interest. This is why in the following the sharpness of a regular observable B will be equivalently understood as the property that for each $\omega \in \Omega$ there exists a unique $\xi \in \Xi$ such that $B\delta_\omega = \delta_\xi$.

The regular observables which are not sharp will be called *fuzzy*. This name is motivated by the fact that when a regular observable is not sharp then at least some of its effects will define fuzzy subsets of Ω, in the sense of Zadeh's fuzzy set theory [16]. Indeed, the effects of a regular sharp observable are characterized by the above expression (1), whose restriction to the Dirac measures on Ω reads

$$(B\delta_\omega)(X) = \delta_\omega(f_B^{-1}(X))$$

so that $E_{B,X}$ defines a $0,1$-function on Ω, hence an ordinary subset of Ω. If a regular observable is not sharp then, at least for some $X \in \mathcal{B}(\Xi)$ and $\omega \in \Omega$, the quantity $(B\delta_\omega)(X)$ will not be restricted to the values 0, 1, so that $E_{B,X}$ will define a function on Ω taking values in the interval [0, 1], hence a fuzzy subset of Ω. (The reader might ask why we have not used the word 'unsharp' to qualify an observable which is not sharp: the reason is that 'unsharp' was often used, as in [13], with a slightly different meaning.)

In general, a physical system admits different levels of description. For instance one might include or ignore some degrees of freedom; as another example, the quantum description of a system which is part of a compound system might be made on the basis of the Hilbert space of the subsystem or on the basis of the Hilbert space of the compound system. In other words, different levels of coarse graining can be adopted. We say that the descriptive model based on the convex set of states S admits an extension based on a set $\tilde{S}$ of states when there exists an affine surjective map $R : \tilde{S} \to S$. Such a map will be called the reduction map, for it reduces the $\tilde{S}$-based model to the S-based one. So, the extended model makes use of a richer set of states, and the reduction map R determines, in general, a many-to-one correspondence between $\tilde{S}$ and S: the elements of $\tilde{S}$ mapped into the same element of S form a 'coarse grain'. Since $\tilde{S}$ is richer than S, the set of observables on $\tilde{S}$ will be, in turn, richer than the set of observables on S. Indeed, if $B : S \to M_1^+(\Xi)$ is an observable of the S-based model then the map composition $\tilde{B} := B \circ R$ is obviously an affine mapping of $\tilde{S}$ into $M_1^+(\Xi)$, hence an observable of the $\tilde{S}$-based model. So, every observable on S finds a representative among the observables on $\tilde{S}$ (not the converse, of course).

The fact that we can have S quantum, that is of the form S_{Q}, and $\tilde{S}$ classical, that is of the form $M_1^+(\Omega)$ for some measurable space Ω is remarkable. In this sense we speak of a classical extension of quantum mechanics. Explicitly, if one considers the measurable space ∂S_{Q} formed by the extremal elements of S_{Q} (that is the pure states, or the one-dimensional projectors on the Hilbert space $\mathcal{H}$) then there exists [6] an affine surjective map

$$R_{\mathrm{M}} : M_1^+(\partial S_{\mathrm{Q}}) \to S_{\mathrm{Q}} \tag{2}$$

which carries the canonical classical extension of quantum mechanics. The reduction map has been denoted R_{M} to remind one that it was studied by Misra [17] some 20 years ago.

It turns out that the statistical distribution of results of a quantum observable and of its classical representative are the same, and the typical quantum features are preserved in the classical extension: so, for example, whenever two quantum observables obey an uncertainty relation so do their classical representatives (with the same uncertainty limit). This is not paradoxical because the classical representatives of the quantum observables are regular but not sharp: they are fuzzy. As it will be seen in section 6, this is also why the violation of Bell-type inequalities, typical of the quantum context, can survive in a classical framework.

3. The Bell phenomenon for measures

Consider a finite collection $\{\Xi_1, \Xi_2, \dots, \Xi_n\}$ of measurable spaces. The Cartesian product $\Xi^{(1,2,\dots,n)} := \Xi_1 \times \Xi_2 \times \cdots \times \Xi_n$ has the natural structure of measurable space with $\mathcal{B}(\Xi^{(1,2,\dots,n)})$ defined as the smallest σ-algebra of subsets of $\Xi^{(1,2,\dots,n)}$ containing all 'rectangles' $X_1 \times X_2 \times \cdots \times X_n$ with $X_i \in \mathcal{B}(\Xi_i)$, $i = 1, 2, \dots, n$. If $\{i_1, i_2, \dots, i_r\} \subset \{1, 2, \dots, n\}$, $i_1 < i_2 < \cdots < i_r$, then the product space $\Xi_{i_1} \times \Xi_{i_2} \times \cdots \times \Xi_{i_r}$, with the σ-algebra of measurable subsets defined as above, will be denoted $\Xi^{(i_1,i_2,\dots,i_r)}$.

In the following the notion of probability measures on measurable spaces will be repeatedly used: we will simply say 'measures' when meaning probability measures.

Definition 1. A consistent family of measures on the product of measurable spaces $\Xi^{(1,2,\dots,n)} := \Xi_1 \times \Xi_2 \times \cdots \times \Xi_n$ is a collection $\mathcal{M}$ of measures on some of the product spaces $\Xi^{(i_1,i_2,\dots,i_r)} := \Xi_{i_1} \times \Xi_{i_2} \times \cdots \times \Xi_{i_r}$ with $i_1 < i_2 < \cdots < i_r$ and $\{i_1, i_2, \dots, i_r\} \subseteq \{1, 2, \dots, n\}$, such that

(i) to any product space $\Xi^{(i_1,i_2,\dots,i_r)}$ there corresponds at most one (i.e. one or none) probability measure $\mu^{(i_1,i_2,\dots,i_r)}$ on the space $\Xi^{(i_1,i_2,\dots,i_r)}$;

(ii) if $\mu^{(i_1,i_2,\dots,i_r)}, \mu^{(j_1,j_2,\dots,j_s)} \in \mathcal{M}$ and $\{i_1, i_2, \dots, i_r\} \cap \{j_1, j_2, \dots, j_s\} =$

$\{k_1, k_2, \ldots, k_t\} \neq \emptyset$ then

$$\pi^{(k_1,k_2,\ldots,k_t)}_{(i_1,i_2,\ldots,i_r)} \mu^{(i_1,i_2,\ldots,i_r)} = \pi^{(k_1,k_2,\ldots,k_t)}_{(j_1,j_2,\ldots,j_s)} \mu^{(j_1,j_2,\ldots,j_s)}$$

where $\pi^{(k_1,k_2,\ldots,k_t)}_{(i_1,i_2,\ldots,i_r)} : M_1^+(\Xi^{(i_1,i_2,\ldots,i_r)}) \to M_1^+(\Xi^{(k_1,k_2,\ldots,k_t)}$ is the marginal projection;

(iii) for every $i \in \{1, 2, \ldots, n\}$ there is in $\mathcal{M}$ a measure $\mu^{(i)}$ on Ξ_i: these 'one-dimensional' measures $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(n)}$ will be called the basic elements of $\mathcal{M}$.

Let us make a few remarks.

Remark 3.1. Condition (ii) sounds as if it is a consistency requirement. The marginal measure $\pi^{(k_1,k_2,\ldots,k_t)}_{(i_1,i_2,\ldots,i_r)} \mu^{(i_1,i_2,\ldots,i_r)}$ is defined by taking the 'rectangle' $X_{k_1} \times X_{k_2} \times \cdots \times X_{k_t} \in \mathcal{B}(\Xi^{(k_1,k_2,\ldots,k_t)})$ and setting

$$(\pi^{(k_1,k_2,\ldots,k_t)}_{(i_1,i_2,\ldots,i_r)} \mu^{(i_1,i_2,\ldots,i_r)})(X_{i_1} \times X_{i_2} \times \cdots \times X_{i_r}) = \mu^{(i_1,i_2,\ldots,i_r)}(\mathrm{Cyl}(X_{k_1} \times X_{k_2} \times \cdots \times X_{k_t}))$$

where the cylinder set $\mathrm{Cyl}(X_{k_1} \times X_{k_2} \times \cdots \times X_{k_t})$ with base $X_{k_1} \times X_{k_2} \times \cdots \times X_{k_t}$ is the 'rectangle' $Y_{i_1} \times Y_{i_2} \times \cdots \times Y_{i_r}$ in $\mathcal{B}(\Xi^{(i_1,i_2,\ldots,i_r)})$ where $Y_l = X_l$ if $l \in \{k_1, k_2, \ldots, k_t\}$ and $Y_l = \Xi_l$ if $l \notin \{k_1, k_2, \ldots, k_t\}$. The marginal measure $\pi^{(k_1,k_2,\ldots,k_t)}_{(i_1,i_2,\ldots,i_r)} \mu^{(i_1,i_2,\ldots,i_r)}$ is then defined at any element of $\mathcal{B}(\Xi^{(k_1,k_2,\ldots,k_t)})$ by the standard methods of measure theory (see, e.g., [18], theorem 11.3).

Remark 3.2. Requirement (iii) is a simplifying feature and looks natural in view of the physical interpretation discussed in the following; in particular it says that the sequence of value spaces $\Xi_1, \Xi_2, \ldots, \Xi_n$ is not redundant since all of them are called into play by the family $\mathcal{M}$.

Remark 3.3. The definition of a consistent family of measures does not imply that such a family has to contain a measure on every marginal product space $\Xi^{(i_1,i_2,\ldots,i_r)}$.

Remark 3.4. The above notion of a consistent family of measures generalizes the one used in the standard theory of stochastic processes, where one is faced with all the 'places' in the hierarchy of product spaces occupied by measures, so that a consistent family of measures contains exactly one measure for every product space constructed from the given sequence of measurable spaces.

It is clear that in the case when the family $\mathcal{M}$ contains a measure $\mu^{(1,2,\ldots,n)}$ on $\Xi^{(1,2,\ldots,n)}$ then this measure generates all the members of the family by the mechanism of marginal projections. In such a case $\mathcal{M}$ is said to be complete, and $\mu^{(1,2,\ldots,n)}$ is called the generating measure of $\mathcal{M}$. If the consistent family $\mathcal{M}$ is not complete, namely if it does not contain a measure on $\Xi^{(1,2,\ldots,n)}$, then the question arises whether it can be made complete by the addition of a suitable measure on $\Xi^{(1,2,\ldots,n)}$, namely whether there exists $\mu^{(1,2,\ldots,n)} \in M_1^+(\Xi^{(1,2,\ldots,n)})$ such that $\mathcal{M} \cup \{\mu^{(1,2,\ldots,n)}\}$ is a complete consistent family of measures. In the case when such a $\mu^{(1,2,\ldots,n)}$ exists then it generates all the members of $\mathcal{M}$ by the mechanism of marginal projections, and we say that $\mathcal{M}$ admits a generating measure.

The relevant and not so obvious fact is that there are consistent families of measures which do not admit a generating measure: in other words they cannot be thought of as sub-families of complete families. When a consistent family of measures $\mathcal{M}$ does not admit a generating measure we say that it exhibits the *Bell phenomenon*. An explicit example will be provided in the next section.

Here we add some further remarks.

Remark 3.5. When referred to the basic elements $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(n)}$ of $\mathcal{M}$ (or to some of them) the definition of generating measure reduces to the familiar definition of joint measure: indeed a measure $\mu^{(1,2,\ldots,n)}$ on $\Xi^{(1,2,\ldots,n)}$ is a joint measure for the $\mu^{(i)}$s if $\mu^{(i)} = \pi^{(i)}_{(1,2,\ldots,n)}\mu^{(1,2,\ldots,n)}$. It is known that the set of 'one-dimensional' measures $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(n)}$ always admits at least one joint measure: hence a consistent family $\mathcal{M}$ formed only by its basic elements $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(n)}$ certainly admits a generating measure.

Remark 3.6. Though the basic measures $\{\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(n)}\}$ of a consistent family $\mathcal{M}$ always determine at least one joint measure on $\Xi^{(1,2,\ldots,n)}$, this joint measure need not be a generating measure for the whole family $\mathcal{M}$ when the latter contains other measures beyond the basic elements.

Remark 3.7. If, besides the basic measures $\{\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(n)}\}$, $\mathcal{M}$ contains only measures having the form of products of the $\mu^{(i)}$s, then $\mathcal{M}$ admits $\mu^{(1)} \times \mu^{(2)} \times \cdots \times \mu^{(n)}$ as a generating measure. This follows from the fact that a marginal projection of a product measure is still a product measure.

Remark 3.8. A generating measure of a consistent family of measures $\mathcal{M}$ need not be unique. Actually, when there are more than one then there are infinitely many because any convex combination of two generating measures is still a generating measure. However we have the following theorem.

Theorem 1. If for any $\{i_1, i_2, \ldots, i_r\} \subseteq \{1, 2, \ldots, n\}$ the set $\{\mu^{(i_1)}, \mu^{(i_2)}, \ldots, \mu^{(i_r)}\}$ of basic measures possesses only one joint measure then $\mathcal{M}$ admits a unique generating measure.

Proof. Noticing that a generating measure is also a joint measure of the basic elements, we have that the uniqueness of the latter entails the uniqueness of the former, if it exists. However, its existence comes from the fact that any $\mu^{(i_1,i_2,\ldots,i_r)} \in \mathcal{M}$ is in turn the unique joint measure of $\mu^{(i_1)}, \mu^{(i_2)}, \ldots, \mu^{(i_r)}$ so that it must coincide with the projected measure $\pi^{(i_1,i_2,\ldots,i_r)}_{(1,2,\ldots,n)}\,\mu^{(1,2,\ldots,n)}$. □

4. The case of events

In this section we restrict ourselves to the special case of measurable spaces consisting of just two points of the real line, say $\Xi_i = \{\xi'_i, \xi''_i\}$, $i = 1, 2, \ldots, n$. This two-valuedness fits the notion of 'event' which either occurs or does not occur: if $\mu^{(i)}$ is a measure on Ξ_i we can agree that $\mu^{(i)}(\xi'_i)$ (respectively $\mu^{(i)}(\xi''_i)$) gives the probability of occurrence (respectively non-occurrence) of the ith event. Similarly, if $\mu^{(i_1,i_2,\ldots,i_r)}$ is a measure on $\Xi_{i_1} \times \Xi_{i_2} \times \cdots \times \Xi_{i_r}$, then the number $\mu^{(i_1,i_2,\ldots,i_r)}(\xi'_{i_1}, \xi'_{i_2}, \ldots, \xi'_{i_r})$ will be naturally read as the probability of the joint occurrence of the events labelled by $i_1, i_2, \ldots, i_r$, while the number $\mu^{(i_1,i_2,\ldots,i_r)}(\xi'_{i_1}, \xi'_{i_2}, \ldots, \xi''_{i_r})$ will represent the probability of joint occurrence of the events $i_1, i_2, \ldots, i_{r-1}$ and of non-occurrence of the event i_r, and so on.

We come now to an explicit example of a consistent family of measures that exhibits the Bell phenomenon (another example will be given in section 6). Consider four measurable spaces $\Xi_1, \Xi_2, \Xi_3, \Xi_4$ and take the consistent family of measures $\mathcal{M} := \{\mu^{(1)}, \mu^{(2)}, \mu^{(3)}, \mu^{(4)}, \mu^{(1,3)}, \mu^{(1,4)}, \mu^{(2,3)}, \mu^{(2,4)}\}$. This family has the typical structure encountered in the so-called Einstein–Podolski–Rosen (EPR) spin correlation: we are going to check that, for a particular choice of its elements, $\mathcal{M}$ exhibits the Bell phenomenon, namely it does not admit a generating measure.

Suppose $\mathcal{M}$ admits a generating measure $\mu^{(1,2,3,4)}$ and notice that this measure would be defined on the 16-point space $\Xi^{(1,2,3,4)} = \Xi_1 \times \Xi_2 \times \Xi_3 \times \Xi_4$. A point of this space will be denoted $\{\xi_1, \xi_2, \xi_3, \xi_4\}$ where ξ_i takes values in $\{\xi'_i, \xi''_i\}$, $i = 1, 2, 3, 4$. The elements of $\mathcal{M}$ would now be obtained as marginal measures of $\mu^{(1,2,3,4)}$; so, for example, $\mu^{(1)}(\xi_1)$ would be the sum of $\mu^{(1,2,3,4)}(\xi_1, \xi_2, \xi_3, \xi_4)$ over ξ_2, ξ_3, ξ_4 while $\mu^{(1,3)}(\xi_1, \xi_3)$ would be the sum of $\mu^{(1,2,3,4)}(\xi_1, \xi_2, \xi_3, \xi_4)$ over ξ_2, ξ_4. Then, by direct inspection, one sees that the quantity

$$\mu^{(1)}(\xi_1') + \mu^{(3)}(\xi_3') - \mu^{(1,3)}(\xi_1', \xi_3') - \mu^{(1,4)}(\xi_1', \xi_4') - \mu^{(2,3)}(\xi_2', \xi_3') + \mu^{(2,4)}(\xi_2', \xi_4') \tag{3}$$

is the sum of the values taken by $\mu^{(1,2,3,4)}$ at the points $(\xi_1', \xi_2', \xi_3'', \xi_4')$, $(\xi_1', \xi_2', \xi_3'', \xi_4'')$, $(\xi_1', \xi_2'', \xi_3', \xi_4'')$, $(\xi_1', \xi_2'', \xi_3'', \xi_4'')$, $(\xi_1'', \xi_2', \xi_3', \xi_4')$, $(\xi_1'', \xi_2', \xi_3'', \xi_4')$, $(\xi_1'', \xi_2'', \xi_3', \xi_4')$, $(\xi_1'', \xi_2'', \xi_3', \xi_4'')$. So quantity (3) is the value taken by $\mu^{(1,2,3,4)}$ at a (eight point) subset of the 16-point space $\Xi^{(1,2,3,4)}$ space, what implies that this quantity must lie between 0 and 1. But this condition can be violated by making, for instance, the explicit choice

$$\begin{aligned}
&\mu^{(i)}(\xi_i') = \tfrac{1}{2} \qquad i = 1, 2, 3, 4 \\
&\mu^{(1,3)}(\xi_1', \xi_3') = \mu^{(1,4)}(\xi_1', \xi_4') = \frac{1}{4}\left(1 - \frac{\sqrt{2}}{2}\right) \\
&\mu^{(2,3)}(\xi_2', \xi_3') = 0 \qquad \mu^{(2,4)}(\xi_2', \xi_4') = \tfrac{1}{4}
\end{aligned} \tag{4}$$

(notice that the remaining values follow from $\mu^{(i)}(\xi_i') + \mu^{(i)}(\xi_i'') = 1$, $\mu^{(1,3)}(\xi_1', \xi_3') + \mu^{(1,3)}(\xi_1', \xi_3'') = \mu^{(1)}(\xi_1') = \frac{1}{2}$, and so on).

Thus we see that, with the above choice, the consistent family $\mathcal{M}$ does not admit a generating measure, hence it exhibits the Bell phenomenon.

The condition that the quantity (2) lies between 0 and 1 (and the analogous conditions obtained by interchanging index 1 with index 2, or 3 with 4, or by making both interchangings) is just the familiar Bell inequality. More specifically it has the form which has already been studied by Clauser and Horne in a well known paper [19] and which has then often been referred to as the Clauser–Horne inequality.

In view of the widespread opinion that the issue of locality is an important ingredient to produce Bell inequalities, one might ask why it did not play any role in our approach. A similar question applies also to other deductions of Bell inequalities, as in [2] where a discussion can be found. In short, when we deal with the EPR spin correlation for a two-particle system and with the four observables corresponding to measuring the spin of one particle along the direction 1 or 2 and the spin of the other particle along the direction 3 or 4, the last four terms in expression (3) represent conditionals on different, alternative measurements. As such, there is no logical need to represent them into a single classical Kolmogorovian probability space. With the words of our approach, there is no need to have them coming from a single generating measure. It is when one embeds this particular experimental situation into the framework of hidden variable theories (as in Bell's original paper) that the issue of locality emerges as a physically natural argument in favour of collapsing these conditional probabilities into absolute probabilities associated with a common state of the physical system. And this collapse corresponds to the fulfilment of Bell inequalities.

The particular framework considered in this section, with dichotomic measurable spaces $\Xi_i = \{\xi_i', \xi_i''\}$, allows a comparison of the notion of the Bell phenomenon with parallel notions occurring in the literature. Let $\mathcal{M}$ be a consistent family of measures on the product of measurable spaces $\{\xi_1', \xi_1''\} \times \{\xi_2', \xi_2''\} \times \cdots \times \{\xi_n', \xi_n''\}$ and consider the collection of numbers

$$\mathcal{P} := \{\mu^{(i_1,i_2,\ldots,i_r)}(\xi_{i_1}', \xi_{i_2}', \ldots, \xi_{i_r}') | \mu^{(i_1,i_2,\ldots,i_r)} \in \mathcal{M}\}.$$

As said above, the quantity $\mu^{(i_1,i_2,\ldots,i_r)}(\xi_{i_1}', \xi_{i_2}', \ldots, \xi_{i_r}')$ is naturally interpreted as the probability of the joint occurrence of the events labelled by $i_1, i_2, \ldots, i_r$. In view of the definition of a consistent family of measures, $\mathcal{P}$ will include the probabilities $\mu^{(i)}(\xi_i')$, $i = 1, 2, \ldots, n$ of occurrence of the events associated with the n basic elements of $\mathcal{M}$.

A collection such as $\mathcal{P}$ is called a 'correlation sequence of probabilities' in [2–4], and it is said to admit a Kolmogorovian (or classical) representation if there exist a Boolean algebra L, a measure ν on L, and elements $a_1, a_2, \ldots, a_n \in L$ such that

$$\mu^{(i)}(\xi_i') = \nu(a_i) \qquad i = 1, 2, \ldots, n$$
$$\mu^{(i_1,i_2,\ldots,i_r)}(\xi_{i_1}', \xi_{i_2}', \ldots, \xi_{i_r}') = \nu(a_{i_1} \cap a_{i_2} \cap \cdots \cap a_{i_r})$$

(we use the set intersection symbol $\cap$ having in mind a realization of L in terms of subsets of a set).

Now we have the following theorem.

Theorem 2. $\mathcal{M}$ admits a generating measure if and only if $\mathcal{P}$ has a Kolmogorovian representation.

Proof. First assume that $\mathcal{M}$ has a generating measure $\mu^{(1,2,\ldots,n)}$. Then

$$\mu^{(i_1,i_2,\ldots,i_r)}(\xi_{i_1}', \xi_{i_2}', \ldots, \xi_{i_r}') = (\pi^{(i_1,i_2,\ldots,i_r)}_{(1,2,\ldots,n)} \mu^{(1,2,\ldots,n)})(\xi_{i_1}', \xi_{i_2}', \ldots, \xi_{i_r}')$$
$$= \mu^{(1,2,\ldots,n)}(\mathrm{Cyl}(\{\xi_{i_1}'\} \times \{\xi_{i_2}'\} \times \cdots \times \{\xi_{i_r}'\}).$$

This means that the correlation sequence $\mathcal{P}$ has a Kolmogorovian representation over the Boolean algebra of measurable subsets of $\{\xi_1', \xi_1''\} \times \{\xi_2', \xi_2''\} \times \cdots \times \{\xi_n', \xi_n''\}$ by means of the measure $\mu^{(1,2,\ldots,n)}$. Conversely, assume that $\mathcal{P}$ has a Kolmogorovian representation in terms of the Boolean algebra L, of the measure ν, and of the elements $a_1, a_2, \ldots, a_n$. Then define a measure $\mu^{(1,2,\ldots,n)}$ on $\{\xi_1', \xi_1''\} \times \{\xi_2', \xi_2''\} \times \cdots \times \{\xi_n', \xi_n''\}$ as follows:

$$\mu^{(1,2,\ldots,n)}(\xi_1', \xi_2', \ldots, \xi_n') = \nu(a_1 \cap a_2 \cap \cdots \cap a_n)$$
$$\mu^{(1,2,\ldots,n)}(\xi_1', \xi_2', \ldots, \xi_{n-1}', \xi_n'') = \nu(a_1 \cap a_2 \cap \cdots \cap a_{n-1} \cap a_n^{\perp}) \qquad \text{etc}$$

where $a_n^{\perp}$ is the complement of a_n in L. The so-defined measure $\mu^{(1,2,\ldots,n)}$ clearly generates $\mathcal{M}$. □

In [2–4] it is shown that the notion of Kolmogorovian representability of correlation sequences of probabilities provides the natural framework and generalization of the familiar Bell inequalities (the latter being particular conditions of Kolmogorovian representability): the discussion above shows that the notion of the Bell phenomenon for a consistent family of measures encompasses the idea of Bell-type inequalities.

5. The Bell phenomenon for observables

Let S be a convex set of states as in section 2. Consider again a finite collection $\{\Xi_1, \Xi_2, \ldots, \Xi_n\}$ of measurable spaces, with the Cartesian product $\Xi^{(1,2,\ldots,n)} := \Xi_1 \times \Xi_2 \times \cdots \times \Xi_n$ as in section 3.

Definition 2. A consistent family of observables on S, with value space $\Xi^{(1,2,\ldots,n)}$, is a collection $\mathcal{O}$ of affine mappings of S into some of the convex sets $M_1^+(\Xi^{(i_1,i_2,\ldots,i_r)})$, where $i_1 < i_2 < \cdots < i_r$ and $\{i_1, i_2, \ldots, i_r\} \subseteq \{1, 2, \ldots, n\}$, such that:

(i) to any sequence $\{i_1, i_2, \ldots, i_r\}$ there corresponds at most one (i.e. one or none) affine mapping of S into $M_1^+(\Xi^{(i_1,i_2,\ldots,i_r)})$ that we denote $B^{(i_1,i_2,\ldots,i_r)}$;

(ii) if $B^{(i_1,i_2,\ldots,i_r)}$, $B^{(j_1,j_2,\ldots,j_s)} \in \mathcal{O}$ and $\{i_1, i_2, \ldots, i_r\} \cap \{j_1, j_2, \ldots, j_s\} = \{k_1, k_2, \ldots, k_t\} \neq \emptyset$ then

$$\pi^{(k_1,k_2,\ldots,k_t)}_{(i_1,i_2,\ldots,i_r)} \circ B^{(i_1,i_2,\ldots,i_r)} = \pi^{(k_1,k_2,\ldots,k_t)}_{(j_1,j_2,\ldots,j_s)} \circ B^{(j_1,j_2,\ldots,j_s)}$$

where $\pi^{(k_1,k_2,\ldots,k_t)}_{(i_1,i_2,\ldots,i_r)} : M_1^+(\Xi^{(i_1,i_2,\ldots,i_r)}) \to M_1^+(\Xi^{(k_1,k_2,\ldots,k_t)})$ is the marginal projection;

(iii) for every $i \in \{1, 2, \ldots, n\}$ there is in $\mathcal{O}$ an element $B^{(i)} : S \to M_1^+(\Xi_i)$: the observables $B^{(1)}, B^{(2)}, \ldots, B^{(n)}$ will be called the basic elements of $\mathcal{O}$.

If the consistent family of observables $\mathcal{O}$ contains an element $B^{(1,2,\ldots,n)} : S \to M_1^+(\Xi^{(1,2,\ldots,n)})$, then this observable generates all the members of the family by the mechanism of marginal projections: indeed one would have $B^{(i_1,i_2,\ldots,i_r)} = \pi^{(i_1,i_2,\ldots,i_r)}_{(1,2,\ldots,n)} \circ B^{(1,2,\ldots,n)}$. In such a case $\mathcal{O}$ is said to be complete and $B^{(1,2,\ldots,n)}$ is called the generating element of $\mathcal{O}$. If the consistent family $\mathcal{O}$ is not complete, then the question arises whether there exists an element $B^{(1,2,\ldots,n)} : S \to M_1^+(\Xi^{(1,2,\ldots,n)})$ such that $\mathcal{O} \cup \{B^{(1,2,\ldots,n)}\}$ is a complete consistent family of observables. When such a $B^{(1,2,\ldots,n)}$ does exist then it generates all the members of $\mathcal{O}$ by the mechanism of marginal projections, and we say that $\mathcal{O}$ admits a generating observable.

The issue of completeness, and of a generating element, of a consistent family of observables $\mathcal{O}$ reminds us of the parallel issue for measures discussed in section 3. There are analogies, but also differences, as shown by the following remarks.

Remark 5.1. When referred to the basic elements $B^{(1)}, B^{(2)}, \ldots, B^{(n)}$ of $\mathcal{O}$ (or to some of them) the definition of generating observable reduces to the familiar definition of joint observable: given the finite collection of observables $B^{(i)} : S \to M_1^+(\Xi_i)$, $i = 1, 2, \ldots, n$, we say indeed that $B^{(1,2,\ldots,n)} : S \to M_1^+(\Xi^{(1,2,\ldots,n)}$ is a joint observable of $B^{(1)}, B^{(2)}, \ldots, B^{(n)}$ if

$$B^{(i)} = \pi^{(i)}_{(1,2,\ldots,n)} \circ B^{(1,2,\ldots,n)} \qquad i = 1, 2, \ldots, n.$$

When a joint observable exists it is also customary [5] to say that the $B^{(i)}$s are mutually comeasurable (or coexistent [13]). The notion of a generating observable generalizes the one of a joint observable: indeed, if $B^{(1,2,\ldots,n)}$ is a joint observable of $B^{(1)}, B^{(2)}, \ldots, B^{(n)}$ then it is a generating observable of the consistent family of observables $\{B^{(1)}, B^{(2)}, \ldots, B^{(n)}\}$; and conversely, whenever a consistent family of observables $\mathcal{O}$ admits a generating observable the latter is a joint observable of the basic elements of $\mathcal{O}$.

Remark 5.2. Within the level of generality considered here (in particular about S) the existence of a joint observable is by no means guaranteed. Think, for example, of the quantum context where $S = S_{\mathrm{Q}}$, and consider two quantum observables associated with non-commuting self-adjoint operators: they do not have, in this context, any joint observable.

Remark 5.3. Even in one case of the basic observables $B^{(1)}, B^{(2)}, \ldots, B^{(n)}$ of a consistent family $\mathcal{O}$ admitting a joint observable, this need not be a generating observable for the whole family $\mathcal{O}$ when the latter contains other observables beyond the basic ones.

Remark 5.4. A generating observable of a consistent family $\mathcal{O}$ need not be unique. Actually, when there are more than one then there are infinitely many because any convex combination of two generating observables is still a generating observable.

However, we have the following

Theorem 3. If for any $\{i_1, i_2, \ldots, i_r\} \subseteq \{1, 2, \ldots, n\}$ the set $\{B^{(i_1)}, B^{(i_2)}, \ldots, B^{(i_r)}\}$ of basic observables of the consistent family of observables $\mathcal{O}$ possesses only one joint observable then $\mathcal{O}$ admits a unique generating observable.

Proof. Let $B^{(1,2,\ldots,n)}$ be the unique joint observable of the collection $\{B_1, B_2, \ldots, B_n\}$ of all the basic elements of $\mathcal{O}$. If $B^{(i_1,i_2,\ldots,i_r)} \in \mathcal{O}$ then it has to be the unique joint observable of the collection $\{B^{(i_1)}, B^{(i_2)}, \ldots, B^{(i_r)}\}$ of basic observables. On the other hand, $\pi^{(i_1,i_2,\ldots,i_r)}_{(1,2,\ldots,n)} \circ B^{(1,2,\ldots,n)}$ is a joint observable of the same set $\{B^{(i_1)}, B^{(i_2)}, \ldots, B^{(i_r)}\}$ of basic observables. Hence the uniqueness assumption implies $\pi^{(i_1,i_2,\ldots,i_r)}_{(1,2,\ldots,n)} \circ B^{(1,2,\ldots,n)} = B^{(i_1,i_2,\ldots,i_r)}$, so that $B^{(1,2,\ldots,n)}$ is the unique generating observable of $\mathcal{O}$. □

Come now to a connection between observables and measures. The elements of a consistent family of observables $\mathcal{O}$ are defined as affine mappings of S into the set of measures on some measurable space: so, if one evaluates all the members of $\mathcal{O}$ at some fixed $\alpha \in S$ one gets a family of measures, to be denoted $\mathcal{M}(\mathcal{O}, \alpha)$, which is clearly a consistent family of measures in the sense of section 3. It is now clear that if $\mathcal{O}$ admits a generating observable $B^{(1,2,\ldots,n)}$ then, for every $\alpha \in S$, $B^{(1,2,\ldots,n)}\alpha$ is a generating measure for $\mathcal{M}(\mathcal{O}, \alpha)$. Therefore, if $\mathcal{O}$ admits a generating observable then there is no $\alpha \in S$ such that $\mathcal{M}(\mathcal{O}, \alpha)$ exhibits the Bell phenomenon; conversely, if for some $\alpha \in S$ the family of measures $\mathcal{M}(\mathcal{O}, \alpha)$ shows the Bell phenomenon then $\mathcal{O}$ cannot admit a generating observable.

In view of these facts it is natural to say that the consistent family of observables $\mathcal{O}$ shows the *Bell phenomenon for observables* if at some $\alpha \in S$ the family $\mathcal{M}(\mathcal{O}, \alpha)$ exhibits the Bell phenomenon for measures, i.e. $\mathcal{M}(\mathcal{O}, \alpha)$ has no generating measure. When $\mathcal{O}$ has no generating observable we say that $\mathcal{O}$ exhibits the *weak Bell phenomenon*. Clearly, for a consistent family of observables the Bell phenomenon implies the weak Bell phenomenon, but not the converse.

In the standard quantum context, where the states are density operators of some Hilbert space $\mathcal{H}$ and the observables correspond to self-adjoint operators, the notion of weak Bell phenomenon would simply reflect the existence in $\mathcal{O}$ of non-commuting observables.

To provide an example of a consistent family $\mathcal{O}$ showing the Bell phenomenon one could consider the familiar EPR situation: take $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ and the family of observables

$$\mathcal{O} = \{B^{(1)}, B^{(2)}, B^{(3)}, B^{(4)}, B^{(1,3)}, B^{(1,4)}, B^{(2,3)}, B^{(2,4)}\} \tag{5}$$

where $B^{(1)} = \boldsymbol{\sigma}\cdot\boldsymbol{a}\otimes I$, $B^{(2)} = \boldsymbol{\sigma}\cdot\boldsymbol{b}\otimes I$, $B^{(3)} = I\otimes\boldsymbol{\sigma}\cdot\boldsymbol{b}$, $B^{(4)} = I\otimes\boldsymbol{\sigma}\cdot\boldsymbol{c}$, $B^{(1,3)}$ is the unique joint observable of the pair of commuting observables $B^{(1)}$, $B^{(3)}$, and similarly for $B^{(1,4)}$, $B^{(2,3)}$, $B^{(2,4)}$. We have denoted $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ three unit vectors in $\mathbb{R}^3$ and $\boldsymbol{\sigma}$ the vector whose components are the Pauli matrices. The value space of $B^{(i)}$, $i = 1, 2, 3, 4$ is the two-point space $\Xi_i = \{\xi_i', \xi_i''\}$ with $\xi_i' = 1$, $\xi_i'' = -1$ (we use the same notation as in section 4), while the value space of $B^{(i,j)}$, $i = 1, 2$, $j = 3, 4$, is the four-point space $\Xi^{(i,j)} = \Xi_i \times \Xi_j$. The explicit form of $B^{(1,3)}$, evaluated at a pure state α of $\mathbb{C}^2 \otimes \mathbb{C}^2$, reads

$$(B^{(1,3)}\alpha)(\xi_1', \xi_3') = (\alpha, P_{\boldsymbol{a}} \otimes P_{\boldsymbol{c}}\alpha) \qquad (B^{(1,3)}\alpha)(\xi_1', \xi_3'') = (\alpha, P_{\boldsymbol{a}} \otimes P_{-\boldsymbol{c}}\alpha)$$
$$(B^{(1,3)}\alpha)(\xi_1'', \xi_3') = (\alpha, P_{-\boldsymbol{a}} \otimes P_{\boldsymbol{c}}\alpha) \qquad (B^{(1,3)}\alpha)(\xi_1'', \xi_3'') = (\alpha, P_{-\boldsymbol{a}} \otimes P_{-\boldsymbol{c}}\alpha)$$

where $P_{\pm\boldsymbol{a}} = \frac{1}{2}(I \pm \boldsymbol{\sigma}\cdot\boldsymbol{a})$ is the projector in $\mathbb{C}^2$ on the eigenvector of $\boldsymbol{\sigma}\cdot\boldsymbol{a}$ with eigenvalues ± 1 (similarly for $P_{\pm\boldsymbol{c}}$), and $(\alpha, P_{\pm\boldsymbol{a}} \otimes P_{\pm\boldsymbol{c}}\alpha)$ stands for the scalar product in $\mathbb{C}^2 \otimes \mathbb{C}^2$. The explicit values of $B^{(1,4)}$, $B^{(2,3)}$, $B^{(2,4)}$ at a pure state α have analogous expressions. Taking for α the singlet (i.e. the zero spin) state in $\mathbb{C}^2 \otimes \mathbb{C}^2$ one gets

$$(\alpha, P_{\boldsymbol{a}} \otimes P_{\boldsymbol{c}}\alpha) = (\alpha, P_{-\boldsymbol{a}} \otimes P_{-\boldsymbol{c}}\alpha) = \tfrac{1}{4}(1 - \boldsymbol{a}\cdot\boldsymbol{c})$$
$$(\alpha, P_{\boldsymbol{a}} \otimes P_{-\boldsymbol{c}}\alpha) = (\alpha, P_{-\boldsymbol{a}} \otimes P_{\boldsymbol{c}}\alpha) = \tfrac{1}{4}(1 + \boldsymbol{a}\cdot\boldsymbol{c})$$

and making the further choice of $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ planar with $\boldsymbol{a}\cdot\boldsymbol{b} = \boldsymbol{a}\cdot\boldsymbol{c} = \frac{\sqrt{2}}{2}$, $\boldsymbol{b}\cdot\boldsymbol{c} = 0$ one would associate with the family of observables given in equation (5) a family $\mathcal{M}(\mathcal{O}, \alpha)$ of measures that would exactly reproduce the example given in equation (4), hence an example of Bell phenomenon for measures.

If one takes $\boldsymbol{b} = \boldsymbol{c}$, so that $B^{(3)} = B^{(4)}$, $B^{(1,3)} = B^{(1,4)}$, $B^{(2,3)} = B^{(2,4)}$, it is easily seen that all the Bell-type inequalities (obtained by requiring that quantities as in equation (3) lie between 0 and 1) are satisfied so that there is no Bell phenomenon, but the family of observables $\mathcal{O}$ shows the weak Bell phenomenon because $B^{(1)}$ and $B^{(2)}$ do not admit a joint observable, hence $\mathcal{O}$ does not admit a generating observable.

6. The Bell phenomenon in classical frames

In the last section the notion of the Bell phenomenon for observables has been introduced within a general frame, without any special assumption about the underlying convex set of states S. Here we come to the relevant case in which S is classical, that is S takes the form of the simplex $M_1^+(\Omega)$ for some measurable space Ω. An observable with value space Ξ now becomes an affine mapping of $M_1^+(\Omega)$ into $M_1^+(\Xi)$. As already stated in section 2, the classical frame contains the class of regular observables: these are the observables we shall be concerned with. They can be divided into sharp and fuzzy ones, according to the definition given in section 2.

Definition 3. Let $B^{(i)} : M_1^+(\Omega) \to M_1^+(\Xi_i)$, $i = 1, 2, \ldots, n$, be regular observables. Let P be the map from the set $\{\delta_\omega | \omega \in \Omega\}$ of the Dirac measures on Ω into $M_1^+(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n)$ defined by $P\delta_\omega := \times_i B^{(i)}\delta_\omega$. The map P is integrable and the integral $\int_\Omega (P\delta_\omega)\,\mathrm{d}\mu(\omega)$ uniquely defines (see [6], theorem 2) a regular observable which maps $M_1^+(\Omega)$ into $M_1^+(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n)$: it will be called the product observable and denoted $B^{(1)} \times B^{(2)} \times \cdots \times B^{(n)}$.

Now we have the following facts.

Remark 6.1. A joint observable for the collection of regular observables $\{B^{(i)} : M_1^+(\Omega) \to M_1^+(\Xi_i), | i = 1, 2, \ldots, n\}$ is provided by the product $B^{(1)} \times B^{(2)} \times \cdots \times B^{(n)}$. This is an immediate consequence of the definition of product observable.

Remark 6.2. If the regular observables $B^{(1)}, B^{(2)}, \ldots, B^{(n)}$ are sharp then there is no other joint observable besides their product. To prove this fact consider a measurable rectangle $R := X_1 \times X_2 \times \cdots \times X_n$ in the product space $\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n$ and notice that the involved effects have to meet the known property $E_{B,R} = \bigwedge_i E_{B^{(i)},X_i}$ where B is an abbreviation for $B^{(1)} \times B^{(2)} \times \cdots \times B^{(n)}$. Now suppose A is another joint observable: we would have $E_{A,R} \leqslant E_{A,X_1 \times \Xi_2 \times \cdots \times \Xi_n} = E_{B^{(1)},X_1}$, and similarly $E_{A,R} \leqslant E_{B^{(i)},X_i}$ for $i = 1, 2, \ldots, n$, hence $E_{A,R} \leqslant E_{B,R}$. Since the last relation has to hold for every rectangle, it has to be an equality, otherwise, viewing R as a member of a (finite) sequence $R, R', R'', \ldots$ of disjoint measurable rectangles which sum up to $\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n$, we would get $E_{A,R} + E_{A,R'} + E_{A,R''} + \cdots < E_{B,R} + E_{B,R'} + E_{B,R''} + \cdots$: a contradiction since both sides represent the unit function on Ω. Thus $A = B = B^{(1)} \times B^{(2)} \times \cdots \times B^{(n)}$.

Remark 6.3. If a consistent family of regular observables on $M_1^+(\Omega)$, with basic elements $B^{(1)}, B^{(2)}, \ldots, B^{(n)}$, contains only product observables (besides its basic elements) then the product $B^{(1)} \times B^{(2)} \times \cdots \times B^{(n)}$ is a generating observable for the whole family. Indeed, every marginal projection of a product observable is a product observable (see the analogous property for measures stated in remark 3.7) and a regular observable is uniquely determined by the values it takes at the set $\{\delta_\omega | \omega \in \Omega\}$ of Dirac measures.

A consistent family of regular observables on $M_1^+(\Omega)$ need not admit a generating observable; it is different when we restrict ourselves to regular sharp observables, as specified by the following theorem.

Theorem 4. A consistent family $\mathcal{O}$ of regular sharp observables on $M_1^+(\Omega)$ admits a unique generating observable which coincides with the product $B^{(1)} \times B^{(2)} \times \cdots \times B^{(n)}$ of all the basic elements of $\mathcal{O}$.

Proof. In view of the property stated in remark 6.2, for every $\{i_1, i_2, \ldots, i_r\} \subseteq \{1, 2, \ldots, n\}$ the set $\{B^{(i_1)}, B^{(i_2)}, \ldots, B^{(i_r)}\}$ of basic observables of the consistent family $\mathcal{O}$ admits a unique joint observable given by the product $B^{(i_1)} \times B^{(i_2)} \times \cdots \times B^{(i_r)}$. Theorem 3 above then ensures that $\mathcal{O}$ admits a unique generating observable. Since a generating observable is also a joint observable for the basic elements of $\mathcal{O}$ we have, using again the property of remark 6.2, that such a generating observable is the product $B^{(1)} \times B^{(2)} \times \cdots \times B^{(n)}$ of the basic elements of $\mathcal{O}$. □

As discussed in [6], the standard classical statistical models make use only of sharp observables: in view of the above theorem these standard models cannot display the Bell phenomenon, nor the weak Bell phenomenon. But classical models can host also fuzzy observables: for these models the Bell effect is not excluded.

A typical situation occurs with the classical extension of quantum mechanics that has been shortly recalled in section 2. The quantum model based on the convex set of states S_Q (the set of density operators on a Hilbert space $\mathcal{H}$) admits the canonical classical extension based on the set $M_1^+(\partial S_Q)$ of the probability measures on the pure states of S_Q. The reduction map R_M of equation (2) provides a one-to-one correspondence between the pure states of S_Q and the Dirac measures of $M_1^+(\partial S_Q)$ but it provides a many-to-one correspondence when referred to non-pure states [6]. If B is an observable on S_Q then $\tilde{B} := B \circ R_M$ is an observable on $M_1^+(\partial S_Q)$, but of course not every observable on $M_1^+(\partial S_Q)$ is the representative of an observable on S_Q. If $\mathcal{O}$ is a consistent family of observables on S_Q then the family $\tilde{\mathcal{O}}$ obtained from $\mathcal{O}$ by composing all its members with R_M is a consistent family of observables on $M_1^+(\partial S_Q)$. Moreover, all probability distributions are preserved by the classical extension, in the sense that given any $\alpha \in S_Q$ we have $B\alpha = \tilde{B}\beta$ for every β in the counterimage of α under R_M. From these facts it follows that whenever a consistent family $\mathcal{O}$ of observables on S_Q exhibits the Bell phenomenon so does the family $\tilde{\mathcal{O}}$ of the classical representatives of the elements of $\mathcal{O}$. In other words, all occurrences of the Bell phenomenon in quantum mechanics are carefully reproduced in the classical extension.

Comparing the above conclusion with theorem 4 the consequence follows that the classical representatives of the quantum observables cannot be regular and sharp: indeed it is known [6] that they are regular and fuzzy.

According to remark 6.1 the basic elements of $\tilde{\mathcal{O}}$ do admit a joint observable, no matter whether the basic elements of $\mathcal{O}$ do or do not. However, the joint observable of the basic elements of $\tilde{\mathcal{O}}$ need not be a generating observable for the whole $\tilde{\mathcal{O}}$ as outlined in remark 5.3. The fact that the classical representatives of the observables on S_Q do not exhaust the observables on $M_1^+(\partial S_Q)$ supports the question of whether $\tilde{\mathcal{O}}$ can admit a generating observable when $\mathcal{O}$ does not: this is an open possibility provided the family of measures $\mathcal{M}(\mathcal{O}, \alpha)$ does not exhibit the Bell phenomenon for any $\alpha \in S_Q$, otherways a contradiction would arise with the fact that the occurrence of the Bell phenomenon in quantum mechanics is reproduced by the classical extension. So, there can be cases in which the weak Bell phenomenon is not reproduced by the classical extension of quantum mechanics. As a trivial example, consider the case of a family $\mathcal{O}$ consisting of two non-commuting quantum observables: it exhibits the weak Bell phenomenon but not the Bell phenomenon (since two measures always have a joint measure), while $\tilde{\mathcal{O}}$ does not exhibit the Bell phenomenon nor the weak one.

An example of occurrence of the Bell phenomenon for classical observables can be obtained by considering the consistent family $\mathcal{O}$ given in equation (5) and looking at its classical counterpart

$$\tilde{\mathcal{O}} = \{\tilde{B}^{(1)}, \tilde{B}^{(2)}, \tilde{B}^{(3)}, \tilde{B}^{(4)}, \tilde{B}^{(1,3)}, \tilde{B}^{(1,4)}, \tilde{B}^{(2,3)}, \tilde{B}^{(2,4)}\}.$$

Taking the singlet state α as in section 5 one obtains a consistent family of measures $\mathcal{M}(\tilde{\mathcal{O}}, \alpha)$ that reproduces precisely the example of equation (4), hence an example of Bell phenomenon. This example, though embedded into a classical framework, still preserves a quantum root if one has in mind the usual EPR spin correlation, but it acquires a more genuine classical flavour if one considers the macroscopical situation conjectured by Aerts (see [20] also for further references) that imitates the quantum mechanical violation of Bell inequalities (the randomness of some parameters of the observables considered there corresponds to their fuzzyness).

The well known experiments on the violation of Bell inequalities, which show that there are empirical probabilities not admitting a Kolmogorovian representation, are usually interpreted as a proof of the impossibility of embedding the quantum description into a classical framework. In view of the analysis made in this paper we can say that a classical embedding is possible but fuzzy observables have to be called into play.

An example of the Bell phenomenon for a physical system strictly described within standard classical mechanics cannot exist: indeed, the deterministic nature of classical mechanics leaves no room for fuzzy observables, hence no room for the Bell phenomenon, as stated by theorem 4 above. To call into play fuzzy observables one needs some randomness in the measurement process, which could realistically simulate a complexity of interactions or external influences not fully under control. In this sense one can look for the occurrence of Bell phenomenon for a classical system (of course randomly behaved systems are also commonly encountered in human sciences such as sociology, economics, psychology, etc).

To visualize a simple example think of a small ball, whose initial kinematical (pure) state is uniquely specified, that undergoes an interaction having random features, and then is recorded by one of an array of eight detectors, say a, b, c, d, e, f, g, h (one could have in mind a ball rolling down an incline where a number of pins cause a random walk, and eight boxes on the lower edge where the ball can fall). Different observables will correspond to different random interactions (different pin patterns) suffered by the ball. We are going to define a consistent family of observables and we first specify operationally their value spaces. We shall deal with three two-valued observables $B^{(i)}$, $i = 1, 2, 3$, and three four-valued observables $B^{(i,j)}$, $i < j = 1, 2, 3$: as in section 4 we denote $\Xi_i = \{\xi_i', \xi_i''\}$ the value space of $B^{(i)}$, while we take the product space $\Xi_i \times \Xi_j$ as the value space of $B^{(i,j)}$. Let us assume that $B^{(1)}$ has outcome ξ_1' if the ball is detected by any one of a, b, c, d and outcome ξ_1'' if it is detected by any one of e, f, g, h: we summarize this by writing

$$\xi_1' \leftrightarrow \{a, b, c, d\} \qquad \xi_1'' \leftrightarrow \{e, f, g, h\}.$$

Referring to $B^{(2)}$, $B^{(3)}$, and $B^{(i,j)}$ let similarly

$$\begin{array}{llll}
\xi_2' \leftrightarrow \{c, d, e, f\} & \xi_2'' \leftrightarrow \{a, b, g, h\} & \xi_3' \leftrightarrow \{b, c, f, g\} & \xi_3'' \leftrightarrow \{a, d, e, h\} \\
(\xi_1', \xi_2') \leftrightarrow \{c, d\} & (\xi_1', \xi_2'') \leftrightarrow \{a, b\} & (\xi_1'', \xi_2') \leftrightarrow \{e, f\} & (\xi_1'', \xi_2'') \leftrightarrow \{g, h\} \\
(\xi_1', \xi_3') \leftrightarrow \{b, c\} & (\xi_1', \xi_3'') \leftrightarrow \{a, d\} & (\xi_1'', \xi_3') \leftrightarrow \{f, g\} & (\xi_1'', \xi_3'') \leftrightarrow \{e, h\} \\
(\xi_2', \xi_3') \leftrightarrow \{c, f\} & (\xi_2', \xi_3'') \leftrightarrow \{d, e\} & (\xi_2'', \xi_3') \leftrightarrow \{b, g\} & (\xi_3'', \xi_3'') \leftrightarrow \{a, h\}.
\end{array}$$

We complete the definition of these observable by specifying that acting on the (initial) state of the ball they will produce the measures

$$\begin{array}{ll}
\mu^{(i)}(\xi_i') = \mu^{(i)}(\xi_i'') = \tfrac{1}{2} \qquad i = 1, 2, 3 & \\
\mu^{(1,i)}(\xi_1', \xi_i') = \mu^{(1,i)}(\xi_1'', \xi_i'') = \tfrac{1}{2} & \mu^{(1,i)}(\xi_1', \xi_i'') = \mu^{(1,i)}(\xi_1'', \xi_i') = 0 \qquad i = 2, 3 \\
\mu^{(2,3)}(\xi_2', \xi_3') = \mu^{(2,3)}(\xi_2'', \xi_3'') = 0 & \mu^{(2,3)}(\xi_2', \xi_3'') = \mu^{(2,3)}(\xi_2'', \xi_3') = \tfrac{1}{2}.
\end{array} \tag{6}$$

It can be easily checked that $B^{(i,j)}$ behaves as a joint observable of $B^{(i)}$ and $B^{(j)}$,

$i < j = 1, 2, 3$, as anticipated by the notation. Now, the consistent family

$$\{B^{(1)}, B^{(2)}, B^{(3)}, B^{(1,2)}, B^{(1,3)}, B^{(2,3)}\} \tag{7}$$

exhibits the Bell phenomenon: indeed the corresponding measures do not admit a generating measure (on an eight-point space) since the pairs $\mu^{(1)}, \mu^{(2)}$ and $\mu^{(1)}, \mu^{(3)}$ are positively correlated whereas the pair $\mu^{(2)}, \mu^{(3)}$ is negatively correlated. As expected from theorem 2 of section 4, we should contextually have the violation of some Bell-like inequality: indeed the inequality

$$0 \leqslant \mu^{(1)}(\xi_1') - \mu^{(1,2)}(\xi_1', \xi_2') - \mu^{(1,3)}(\xi_1', \xi_3') + \mu^{(2,3)}(\xi_2', \xi_3') \leqslant 1$$

called the Bell–Wigner inequality in [2], is violated.

We make use of this example to add a couple of remarks.

One might ask what changes if the elements of randomness are shifted from the observables to the state of the system, in other words if instead of considering fuzzy observables acting on a pure state one takes sharp observables acting on a mixed state. Of course we expect from theorem 4 that the above example will no longer carry any Bell phenomenon. We know that a mixed state in a classical context has an unambiguous convex decomposition into pure states, and pure states have no dispersion on sharp observables (the strict connection between the simplex nature of the set of states and the absence of dispersion for pure states has been discussed in [11]). So, let the state of the ball be a mixture $\frac{1}{2}\alpha + \frac{1}{2}\beta$ of two pure states α, β and let us preserve the operational structure of the value spaces of the observables specified above. This mixed state could still be mapped into the measures of equation (6), but as soon as we come to a pure state, say α, we would unavoidably end in some family of measures such as

$$B^{(i)}\alpha = \delta_{\xi_i'},\ i = 1, 2, 3 \qquad B^{(i,j)}\alpha = \delta_{\xi_i', \xi_j'},\ i = 1, 2 \qquad B^{(2,3)}\alpha = \delta_{\xi_2', \xi_3''}$$

which is not a consistent family of measures in the sense of definition 1 since $B^{(2,3)}$ is no longer a joint observable of $B^{(2)}, B^{(3)}$. Thus we see that when we force the observables of the above example into sharp ones the set of equation (7) looses the essential nature of a consistent family of observables.

As a final remark let us outline that the Bell phenomenon emerging from the above example has nothing to do with any notion of locality. This stresses the fact, partially anticipated in section 4, that locality is not an ingredient of the Bell phenomenon: historically the issue of locality became a concern only occasionally in the particular context of the EPR correlation and of its interpretation inside the hidden variable philosophy.

Acknowledgment

One of the authors (S Bugajski) acknowledges the support provided by the Science Research Committee (KBN, Warsaw) under the grant 2/2420/92/03.

References

[1] Bell J S 1964 *Physics* **1** 195

[2] Pitowsky I 1989 *Quantum Probability—Quantum Logic (Lecture Notes in Physics 321)* (Berlin: Springer)

[3] Beltrametti E G and Maczynski M J 1991 *J. Math. Phys.* **32** 1280–6; 1993 *J. Math. Phys.* **34** 4919–29

[4] Beltrametti E G and Maczynski M J 1993 On the intrinsic characterization of classical and quantum probabilities *Symposium on the Foundations of Quantum Physics* ed P Busch, P Lahti and P Mittelstaedt (Singapore: World Scientific)

[5] Beltrametti E G and Bugajski S 1995 *Int. J. Theor. Phys.* **34** 1221–9

[6] Beltrametti E G and Bugajski S 1995 *J. Phys. A: Math. Gen.* **28** 3329–43

[7] Ali S T and Prugovecki E 1977 *J. Math. Phys.* **18** 219–28

[8] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)

[9] Singer M and Stulpe W 1992 *J. Math. Phys.* **33** 131–42

[10] Bugajski S 1993 *Int. J. Theor. Phys.* **32** 389–98

[11] Beltrametti E G and Bugajski S 1993 *Int. J. Theor. Phys.* **32** 2235–44

[12] Cassinelli G and Lahti P 1993 *J. Math. Phys.* **34** 5468–75

[13] Busch P, Grabowski M and Lahti P 1995 *Operational Quantum Physics* (Berlin: Springer)

[14] Stulpe W 1986 Bedingte Erwartungen und Stochastische Prozesse in der Generalisierten Wahrscheinlichkeitstheorie *Doctoral Dissertation* Berlin

[15] Bugajski S, Hellwig K E and Stulpe W 1995 On fuzzy random variables and statistical maps, to be published

[16] Zadeh L A 1965 *Information and Control* **8** 338–53

[17] Misra B 1974 *Physical Reality and Mathematical Description* ed C P Enz and J Mehra (Dordrecht: Reidel) pp 455–76

[18] Billingsley P 1979 *Probability and Measure* (New York: Wiley)

[19] Clauser J F and Horne M A 1974 *Phys. Rev.* D **10** 526–35

[20] Aerts D 1991 *Helv. Phys. Acta* **64** 1–23