An attempt to close the Einstein–Podolsky–Rosen debate

Thomas Krüger

Abstract: Based on a new *rigorous* ensemble approach to quantum mechanics, and without stressing any idea or concept of reality, the entire Einstein–Podolsky–Rosen (EPR) problem can be boiled down to the question of whether the separability principle of the natural sciences is universally valid. To give a precise answer first of all Bell's inequality is deduced from said ensemble point of view and with minimal requirements only. (In the final discussion of the results it turns out that Bell's inequality defines the upper bound for those basic correlations that are due to a mere conservation law.) Then, by use of Wheeler's gedanken experiment with coin halves, I show that the statistical operator representing an ensemble under investigation may be either separable (*in a simplified sense*) or not. The conceptual consequences of nonseparability are explained, and a general EPR-type experiment is re-examined. Thereby, it is proven that, if and only if, the statistical operator is *non*separable, Bell's inequality may be violated. Experimental evidence demands nonseparable operators. So, if quantum mechanics is assumed to make statistical statements on the results of measurements on ensembles only, there is no way to avoid acceptance of its (operationally) *holistic* character, and the question posed at the outset must be negated.

PACS Nos.: 03.65.Ta, 03.65.Ud

Résumé : Sur la base d'une nouvelle approche d'ensemble de la mécanique quantique et sans insister sur l'idée ou le concept de réalité, l'ensemble du problème Einstein–Podolsky–Rosen (EPR) peut se ramener à la question de savoir si le principe de *séparabilité* en sciences naturelles est universellement valide. Afin de donner une réponse précise, l'inégalité de Bell est d'abord déduite de ce point de vue d'ensemble avec un minimum de conditions. (Dans la discussion finale des résultats, il appert que l'inégalité de Bell définit la limite supérieure pour ces corrélations de base qui sont dues à une simple loi de conservation.) En utilisant alors l'expérience par la pensée de Wheeler sur les pièces de monnaie, nous montrons que l'opérateur statistique représentant un ensemble étudié peut être soit séparable (au sens simple) ou non. Nous expliquons les conséquences conceptuelles de la non-séparabilité et réexaminons une expérience générale de type EPR. Il est alors démontré que si et seulement si l'opérateur statistique est non-séparable, alors l'inégalité de Bell peut être violée. Une démonstration expérimentale exige donc des opérateurs non-séparables. Ainsi, si la mécanique quantique est présumée porter des jugements statistiques sur les résultats de mesures sur des ensembles seulement, il n'y a aucune façon d'éviter l'acceptation de son caractère (opérationnellement) holistique et nous devons répondre non à la question posée au début.

[Traduit par la Rédaction]

Received 11 February 2003. Accepted 30 November 2003. Published on the NRC Research Press Web site at http://cjp.nrc.ca/ on 16 February 2004.

T. Krüger. Theoretical Physics, Faculty of Science, University of Paderborn, Warburger Str. 100, 33098 Paderborn, Germany (e-mail: t.krueger@phys.uni-paderborn.de).

1. Introduction

1.1. Sketch of the situation and aim of the article

The present situation of quantum mechanics (QM) has been reviewed extensively by Laloë [1]. Laloë points out that "the discussions are still lively", which is mainly due to the unsolved question about the "true nature" of the so-called state vector $|\Psi\rangle$. Does it represent single entities or ensembles? Does it describe the (physical) reality or does it contain the (partial) knowledge of this reality that we may have or obtain? And what is "reality" supposed to mean?

Further problems are connected with $|\Psi\rangle$. If we assume that the state vector acts as an information carrier only — does it contain *all* the information about the entity in question or does it contain the information that *we* have about the entity? In other words: Is said information absolute, that is, a feature of the entity, or is it relative, that is, does it reflect our individual knowledge about the entity? In the latter case, two observers would have to use different state vectors for one and the same entity.

In spite of the fact that the convincing solution of the latter problem can be found in a (quite often ignored) paper by Bunge [2], additional questions arise. What is it about the measurement process? Assuming that the realm of QM may be extended to cover macroscopic measurement devices too, a simple calculation shows that a superposition of state vectors pertaining to our entity produces a superposition of state vectors pertaining to the device. However, these macroscopic superpositions are never observed. Instead every experiment yields one single result and not a linear combination of pointer positions. To explain this fact the reduction postulate had been introduced into quantum mechanics. Its task is to take care of the proper just-in-time collapse of the superposition, but, strictly speaking, it is not satisfactory to solve a problem by just postulating it away.

Things become really difficult if Schrödinger's cat paradox is taken into account. The measurement of the cat's "state" takes place if the box is opened. In this moment, the postulate operates and in consequence we obtain an unambiguous result. At any time before, however, the cat's state is a superposition of $|\Psi_{\text{dead}}\rangle$ and $|\Psi_{\text{alive}}\rangle$, which is the most extreme variety of a macroscopic superposition, and it is hard to believe that the cat is a zombie as long as the box is still closed.

Things become even worse if the 1935 argument of Einstein, Podolsky, and Rosen (EPR) [3] is taken into account. EPR have added two seemingly reasonable assumptions to the framework of QM as it was generally accepted in the mid-thirties [4]:

1. *Principle of reality*: "If, without in any way disturbing a system, we can predict with certainty . . . the value of a physical quantity, then there exists an element of physical reality corresponding to that quantity."

2. *Principle of separability*: "If a physical system remains, for a certain time, . . . isolated from other systems, then the evolution of its properties during this whole time interval cannot be influenced by operations carried out on other systems."

Then, by precise logical reasoning, the following theorem can be obtained [1]: "If the predictions of QM are correct . . . and if physical reality can be described in a local (or separable) way, then QM is necessarily incomplete." EPR were convinced that

- $|\Psi\rangle$ corresponds to an entity and not to the mere knowledge about it, and that

- the entities belonging to the microworld are real in the same sense as the macroscopic "objects" of our daily experience.

This belief, combined with special relativity, automatically yields the conclusion that the principle of separability is correct. Applied to the above-mentioned theorem, we then obtain the following statement: QM is incomplete.

Recall, however, EPR's line of reasoning. These two assumptions are the essential prerequisites of the canonization of the separability principle. But can these assumptions still be taken for granted? On the one hand, the use of the state vector in QM causes a lot of conceptual problems that are far from being solved. On the other hand, "reality" is by sure one of the most leached, dwelled on, exhausted, and confused notions in contemporary language. Before talking honestly about reality, this notion has to be defined and analyzed in detail — and with respect to the whole of Occidental philosophy and thinking. The implications and complications connected with this notion are nearly innumerable, and it is not at all surprising that no elaboration on microphysical reality is at hand that could not be attacked quite easily [5].

Moreover, it must be emphasized that reality cannot be a question of the formalism. The formalism *may allow for* a realistic interpretation or not, but no formalism at all can *contain* a realistic feature or quality or something else. The formalism is number and figure, not meaning, and *no* formalism can carry the interpretation in itself [6]. So the general validity of the separability principle is not guaranteed.

The aim of the present paper is to unveil what I assume to be the actual core of EPR's discovery. It is not intended to present a review-like survey over the tremendous number of attempts to solve the EPR problem. (Some of them have been discussed by Home and Selleri [7] in an older review that still is worth reading.) Therefore, only those papers are quoted that are related directly to the Ansatz followed up in this work.

1.2. The rigorous ensemble approach

To elucidate things as far as possible, we first have to get rid off all problematic aspects of the theory. Therefore, the EPR analysis will be based on the following thesis:

QM makes

- statistical statements on
- the results of
- measurements on
- ensembles.

Why this ensemble interpretation? First of all, it virtually demands the use of statistical operators whereby we may ignore the notion of a state vector and all the questions regarding its meaning that have been posed above. In consequence, all problems connected with the reduction postulate are removed as well. Moreover, we are not concerned about the reality of individual microentities. Even the whole discussion about reality itself becomes irrelevant. But this approach to EPR still offers a couple of further advantages: It reflects the practical situation of the EPR experiments performed insofar as it just deals with ensembles of entities and not with single constituents. No experiment has ever been published where less than some thousand single runs have been done. And last but not least, this approach avoids the problematic notion of the primary probability of a single event (Popper's propensity). Here, we speak about probabilities in the quite trivial sense of relative frequencies only.

The ensemble interpretation employed here does *not* rest on the idea that each individual member of the ensemble always *has* (in the sense of EPR's principle of reality) precise values (properties) for all its property types, that is, it does *not* belong to the class of so-called pre-assigned initial value (PIV) interpretations [8]. Recall that Gillespie has proven that, in the case of certain quantum systems, these interpretations are not possible at all [9].

Instead, the proposed interpretation is *rigorous* in the following sense: There are, of course, outcomes of single runs, but we do not ascribe them to individual entities! These pointer positions are considered as single constituents of the global outcome only. This global outcome, however, is the only relevant number, and it is ascribed to the ensemble. Only the *whole*, that is, the ensemble and, on another level, the global outcome, is treated as an element of QM. Only the ensemble is visible. The individual entity remains veiled, and therefore, we do *not* make any statements regarding individuals.

The only thing we can say is: if there "is" an ensemble, then there must "be" constituents. They, however, do not have an "existence" on their own. They, so to say, "live" as constituents of a whole only. There "is" no individual ant despite the fact that we can perform experiments with it. The individual ant "exists" solely as a part of a well-structured collective. An individual ant, separated from the collective, is nonviable.

Comparing this rigorous ensemble approach with the other interpretations that have been used to explain EPR from an ensemble point of view (see, for example, Chap. 6.1 of ref. 8), it is seen immediately that the present approach is a *new* Ansatz to solve the old problem.

One could object that the EPR problem has already been discussed exhaustively on a comparably radical level by Ludwig [10], but closer inspection shows that this is not the case. Ludwig's work is based on a purely operationalistic interpretation of QM *not* allowing for any *ontic* thinking about the microworld. "... the axiomatic basis of quantum mechanics is formulated without the need for the assumption of the existence of microsystems" [11]. So Ludwig analyzes the EPR problem *technically*, in terms of preparation and registration, using "the well-defined language of craftsmen," whereas the present author attempts to rescue the ontic point of view. In my opinion this aim can be achieved with certainty if and only if the problem is re-formulated in terms of statistical operators. This Ansatz, however, has not been addressed by Ludwig.

2. Bell's inequality from the rigorous ensemble point of view

Bell's inequality [12,13] has opened the way to the experimental investigation of the EPR problem. So the first step of the analysis must be the derivation of Bell's inequality, which is at the heart of all EPR-type experiments, by considering said ensembles and the trivial notion of probability only.

At a time t_0 an entity $(\mathrm{UV})_\mathrm{i}$ shall decay into a pair of fragments $(\mathrm{U_i}, \mathrm{V_i})$, which are separated spatially so that they do not interact any more. At a time $t_1 \gg t_0$, let $\mathrm{U_i}$ $(\mathrm{V_i})$ impinge on an apparatus A (B) measuring a quantity represented by a self-adjoint operator $\hat{A}$ $(\hat{B})$ and yielding some value ± 1. Then the outcome of the coincidence measurement is given by the simple product

$$O_i(\boldsymbol{a}, \boldsymbol{b}) = A_i(\boldsymbol{a}) \times B_i(\boldsymbol{b}) \tag{1}$$

where the vector $\boldsymbol{a}$ $(\boldsymbol{b})$ is an experimental parameter determining the actual internal status of apparatus A (B). Now, the experiment is repeated N_1 times $(N_1 \gg 1)$, and we obtain the final result of this measurement series No. 1 according to

$$O(\boldsymbol{a}, \boldsymbol{b}) = \frac{1}{N_1} \sum_i O_i(\boldsymbol{a}, \boldsymbol{b}) \tag{2}$$

Assume that a second ensemble of fragments is used to perform a second series of measurements that differs from the first one insofar as the setting of apparatus B has been changed from $\boldsymbol{b}$ to $\boldsymbol{b}'$. In consequence, we obtain

$$O(\boldsymbol{a}, \boldsymbol{b}) - O\left(\boldsymbol{a}, \boldsymbol{b}'\right) = \frac{1}{N_1} \sum_i A_i^{(1)} B_i - \frac{1}{N_2} \sum_i A_i^{(2)} B_i' \tag{3}$$

The superscripts have been introduced to make it explicit that the ensemble used in series No. 1 is different from the one used in series No. 2, because no single entity is measured twice. Furthermore, note that the ith pointer position of apparatus A may be different in the two series. Even if $N_1 = N_2$, the subensemble $\{\mathrm{U_i}\}^{(1)}$ used in the first measurement series could cause a sequence of single outcomes $A_1^{(1)}, A_2^{(1)}, \ldots$ as

$+ - - - + - - \ldots$

where in the second case an order as

$- - + + - - - \ldots$

might be detected (although the final results are identical). Therefore, it is not possible, in general, to combine the two sums on the right-hand side of (3). Let us, however, assume for the moment that $N_1 = N_2 = N$ and $\{\mathrm{U_i}\}^{(1)} \cong \{\mathrm{U_i}\}^{(2)}$, that is, the outcome $A_i^{(1)}$ shall be equal to $A_i^{(2)}\ \forall\, i$. In this case, we say that the two $\mathrm{U_i}$ subensembles are *equivalent*. Then

$$O(\boldsymbol{a}, \boldsymbol{b}) - O\left(\boldsymbol{a}, \boldsymbol{b}'\right) = \frac{1}{N} \sum_i A_i \left(B_i - B_i'\right) \tag{4}$$

Since the absolute value of a sum is always less than or at best equal to the sum of the absolute values of the single parts, we obtain

$$|O(\boldsymbol{a}, \boldsymbol{b}) - O\left(\boldsymbol{a}, \boldsymbol{b}'\right)| \le \frac{1}{N} \sum_i |A_i \left(B_i - B_i'\right)| = \frac{1}{N} \sum_i |A_i| \times |B_i - B_i'| \tag{5}$$

In the same way as above, two additional measurement series yield

$$|O\left(\boldsymbol{a}', \boldsymbol{b}\right) + O\left(\boldsymbol{a}', \boldsymbol{b}'\right)| \le \frac{1}{N} \sum_i |A_i'| \times |B_i + B_i'| \tag{6}$$

The physical quantity in question allows for single run outcomes ± 1 only. Therefore, $|A_i| = |A_i'| = 1 \forall i$. Addition of (5) and (6) then leads to

$$\Delta \overset{\text{def}}{=} |O(\boldsymbol{a}, \boldsymbol{b}) - O\left(\boldsymbol{a}, \boldsymbol{b}'\right)| + |O\left(\boldsymbol{a}', \boldsymbol{b}\right) + O\left(\boldsymbol{a}'\boldsymbol{b}'\right)| \le \frac{1}{N} \sum_i \left(|B_i - B_i'| + |B_i + B_i'|\right) \tag{7}$$

and it is easy to see that the term in brackets is equal to $2\ \forall\, i$. So, we finally arrive at Bell's inequality

$$\Delta \le 2 \tag{8}$$

Note that this result has been achieved using two prerequisites:

1. The determination of A_i does *not* influence B_i and vice versa, that is, there is no action-at-a-distance between the two *detectors*.
2. There is an equivalence of subensembles:

   $\{\mathrm{U_i}\}^{(2)} \cong \{\mathrm{U_i}\}^{(1)}$ and $\{\mathrm{U_i}\}^{(4)} \cong \{\mathrm{U_i}\}^{(3)}$

The second requirement, however, is not really necessary to establish the inequality. If, in contrast to the deduction of (8), the four subensembles are mutually *in*equivalent, we nevertheless may write $\mathrm{A_i}^{(2)} = f_i \mathrm{A_i}^{(1)}$ and $\mathrm{A_i}^{(4)} = f_i' \mathrm{A_i}^{(3)}$ where the only possible values of f and f' are ± 1. Instead of (7) we then obtain

$$\Delta \le \frac{1}{N} \sum_i \left(|B_i - f_i B_i'| + |B_i + f_i' B_i'|\right) \tag{9}$$

Now, assume first that, for a particular i, $f_i' = f_i$. In this case, the term in brackets becomes equal to 2 for all possible values of B_i and B_i'. If, however, for this index i we have $f_i' = -f_i$, then with four of the eight possible B_i-B_i' combinations the term in brackets becomes equal to 0 but with the other four combinations a value of 4 results. In the regime of the large-N limit, it is to be expected that in 50% of the cases f_i' will be equal to f_i, whereas, in the other cases $f_i' = -f_i$ will hold. So half of the sum terms will be equal to 2 while the other terms will be equal to 0 or 4, but since the two latter values will occur with the same probability, summation over all i will always yield $2N$ in the large-N limit so that Bell's inequality is universally valid — even without the second prerequisite.

This would mean that a violation of (8) would automatically prove the "existence" of an action-at-a-distance at the macroscopic scale! There is, however, a *further* implicit requirement for the validity of this conclusion: the independence of A_i and $B_i \forall i$, see (1), that is, A_i may not depend on $\boldsymbol{b}$ and B_i may not depend on $\boldsymbol{a}$ for *all* pairs forming the total ensemble. This prerequisite is analyzed in detail in the following section.

3. The coin-halves experiment as an aid to understand nonseparability

3.1. The experiment

Let us perform a gedanken experiment, which dates back to the late J. A. Wheeler [14]: Suppose that a coin is sawn apart to that one half contains the head of the original coin and the other half the tail. Now each half is put into an envelope. Envelope 1 is sent to an observer named Alice whilst envelope 2 is sent to another observer named Bob. If Alice opens her envelope she will not only realize which half of the coin she has obtained, but she will also know with certainty what Bob will discover if he opens his envelope. Obviously these two processes are correlated. If Alice observes a "head" Bob will observe a "tail" and vice versa. Since Alice and Bob cannot be in possession of the same side of the original coin (unless magic is involved), a combination as (A_H, B_H) will never be found.

Of course this experiment can also be described using the language of QM. (At the moment we do not pay attention to the fact that coin halves are classical entities.) If we repeat it N times with $N \gg 1$ the ensemble of all coins cut into halves is represented by a statistical operator ρ_c, that is, by a self-adjoint operator with a non-negative spectrum and $\mathrm{Tr}\rho_c = 1$. ρ_c acts on the 2×2-dimensional Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Let $\{|\alpha_i\rangle|\beta_i\rangle\}$ be an orthonormal basis of $\mathcal{H}$. Then, in the most general case, ρ_c is given by

$$\rho_c = \sum_{i,j,k,l} c_{ij,kl} \hat{A}_{ij} \otimes \hat{B}_{kl} \tag{10}$$

where $\hat{A}_{ij} = |\alpha_i\rangle\langle\alpha_j|$ and $\hat{B}_{kl}$ defined analogously.

There are, *in principle*, four possible outcomes of a single run, namely, (A_H,B_H), (A_H,B_T), (A_T,B_H), and (A_T,B_T). The probability for, say, (A_H,B_H) to occur is

$$W_{HH} = \mathrm{Tr}\left(\rho_c(\hat{A}_{HH} \otimes \hat{B}_{HH})\right) = c_{HH,HH} \tag{11}$$

Now, since magic has been excluded, W_{HH} (and W_{TT} as well) must be equal to 0 so that two of the 16 coefficients $c_{ij,kl}$ are already fixed. We know, on the other hand, that (A_H,B_T) and (A_T,B_H) will be observed with equal probability so that $c_{HH,TT} = c_{TT,HH} = \frac{1}{2}$. (Note that any other value would lead to a violation of $\mathrm{Tr}\rho_c = 1$.)

The 12 nondiagonal coefficients are not defined yet. However, both the self-adjointness of ρ_c and the requirement of a non-negative spectrum impose several conditions on these coefficients, and it can be shown that all of them must vanish. For example, consider the case where only the four antidiagonal coefficients are different from 0. Then, because of the self-adjointness of ρ_c, $c_{HT,HT} = c_{TH,HT} = a + ib$

and $c_{\mathrm{HT,TH}} = c_{\mathrm{TH,TH}} = a - ib$, and the eigenvalues of ρ_{c} are

$$\lambda_{1,2} = \pm\left(a^2 + b^2\right)^{1/2}$$

and

$$\lambda_{3,4} = \frac{1}{2}\left(1 \pm \left(a^2 + b^2\right)^{1/2}\right)$$

respectively. The requirement $\lambda_i \geq 0\,\forall\, i$ is fulfilled only if $a^2 + b^2 = 0$, that is, if ρ_{c} is diagonal. So we finally obtain

$$\rho_{\mathrm{c}} = \frac{1}{2}\left(\hat{A}_{\mathrm{HH}} \otimes \hat{B}_{\mathrm{TT}} + \hat{A}_{\mathrm{TT}} \otimes \hat{B}_{\mathrm{HH}}\right) \tag{12}$$

Now reconsider the experimental situation. Alice can do with her half of the coin what she wants. None of her actions, not even the dissolution of the half in nitrohydrochloric acid, exerts any influence on Bob's half. Therefore, the two subensembles $\{\mathrm{A_i}\}$ and $\{\mathrm{B_i}\}$ may be considered totally *independent*, and this system should be a realization of EPR's 1935 principle of separability (see above). In consequence, and instead of (10), one would make the following Ansatz for the statistical operator representing the total ensemble $\{\{\mathrm{A_i}\}, \{\mathrm{B_i}\}\}$,

$$\rho_{\mathrm{c}}' = \rho_A \otimes \rho_B \tag{13}$$

with

$$\rho_A = \sum_{i,j} a_{ij}\hat{A}_{ij} \tag{14}$$

and ρ_B defined analogously.

3.2. When is a statistical operator separable?

I call a statistical operator *separable* if and only if it can be decomposed according to (13). Note that this definition differs from the usual one where an operator is called separable if it can be decomposed into a *convex sum* of direct products as on the right side of (13). However, the definition used in this article offers a couple of advantages with respect to the usual one, which is discussed in detail in ref. 15.

It is of fundamental importance to realize that a definition in itself can be neither right nor wrong. One can be more appropriate or suitable than the other, but this is obviously a matter of debate. An operationalist, for example, will always prefer (13) over the usual definition, because (13) reflects directly the independence that might be observed in the laboratory. Recall that *both* definitions yield the *same* upper bound of the correlation function in a general EPR-type experiment, which amounts to $\sqrt{2}$! This has been proven in ref. 5. So it is easy to see that both the derivation and the validity of Bell's inequality ($\Delta \leq 2$) are *independent* of the separability definition applied.

Furthermore, it is worth noting that there is some ambiguity in these definitions anyway. Abouraddy et al. [16] say that $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is *factorizable* if and only if $|\Psi\rangle = |\Psi_A\rangle \otimes |\Psi_B\rangle$ and *entangled* if not. Tsallis et al. [17], however, divide the set of statistical operators as follows: ρ is *uncorrelated* if $\rho = \rho_A \otimes \rho_B$. It is *separable* if $\rho = \sum_i p_i \rho_{A,i} \otimes \rho_{B,i}$, and it is *entangled* (which is considered equivalent to *non*separable) if not. But Lomonaco Jr. [18] states explicitly that a *pure* ensemble is separable if it satisfies a condition equivalent to the one of Abouraddy et al. while in the case of a *mixed* ensemble "one possible definition" is the one used by Tsallis et al. Obviously there is no definition that is accepted by all colleagues in common. This situation really demands a simplified approach as presented in refs. 5 and 15.

However, to avoid a misunderstanding and to enhance the clarity of the approach, the statistical operators in the form of (13) will be called K-separable (≡ Krüger-separable) from now on.

It is easy to prove that ρ_c given by (12) can never be written in the form of (13), that is, ρ_c is *non*-K-separable. What does this mean? It simply means that the two subensembles of coin halves, $\{A_i\}$ and $\{B_i\}$, are *not independent* of one another — despite the fact that there is no kind of interaction between them! The reason is that a half can only be defined with respect to the whole. A "head" makes sense only with respect to a complete coin. There is no "head" without a "tail". If the "tail" is dissolved in nitrohydrochloric acid, then, strictly speaking, the "head" disappears too. Of course a piece of metal remains, even without any obvious change in its appearance, but this remainder *has* been the "head" of a coin. Now it is something else. It has changed its character. Putting it into plain words: there is no "existence" of one half independent of the other. This is the ontic meaning of non-K-separability.

3.3. Is non-K-separability detectable?

Which one of the two possible statistical operators is the proper representation of the total ensemble? This question can be answered only if ρ_c and ρ_c' lead to different predictions that can be checked experimentally. The correlation function Δ, introduced in (7), is the cornerstone of all experimental tests of nonseparability performed so far. So it seems to be the essential tool for the present task too. Since

$$O(\boldsymbol{a}, \boldsymbol{b}) = \mathrm{Tr}\left(\rho(\hat{A} \otimes \hat{B})\right) \tag{15}$$

depends on the specific choice of the statistical operator, it is to be expected that Δ will depend on it as well. In contrast to, for example, a particle with spin, a half coin possesses *one* degree of freedom only. So the only chance to construct four different measurement series is to adjust the apparatuses A and B so that they can detect which kind of half coin the incoming half is. Let A be able to discriminate between the "head" and the "tail", the corresponding operator shall be given by

$$\hat{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{16}$$

which is equal to Pauli's matrix σ_3. So apparatus A will show $+1$ if the incoming half is a "head" and -1 if it is a "tail". $\hat{A}'$, however, may not be equal to $\hat{A}$. We, therefore, choose $\hat{A}' = \sigma_1$ which yields 0 in both cases. The rotation of $\hat{A}$ to $\hat{A}'$ corresponds to a complete detuning of the detector efficiency, that is, $\hat{A}'$ represents A if the apparatus is switched off. Let $\hat{B}$ be equivalent to $\hat{A}'$ and $\hat{B}'$ equivalent to $\hat{A}$. Recall that $\hat{A}$ and $\hat{A}'$ as well as $\hat{B}$ and $\hat{B}'$ do not commute.

Let us first investigate the K-separable case. From

$$O'(\boldsymbol{a}, \boldsymbol{b}) = Tr(\rho_A \hat{A})\,\mathrm{Tr}(\rho_B \hat{B}) \tag{17}$$

we obtain

$$\begin{aligned} O'(\boldsymbol{a}, \boldsymbol{b}) &= (a_{\mathrm{HH}} - a_{\mathrm{TT}})(b_{\mathrm{HT}} + b_{\mathrm{HT}}^*) \\ O'(\boldsymbol{a}, \boldsymbol{b}') &= (a_{\mathrm{HH}} - a_{\mathrm{TT}})(b_{\mathrm{HH}} - b_{\mathrm{TT}}) \\ O'(\boldsymbol{a}', \boldsymbol{b}) &= (a_{\mathrm{HT}} + a_{\mathrm{HT}}^*)(b_{\mathrm{HT}} + b_{\mathrm{HT}}^*) \end{aligned} \tag{18}$$

and

$$O'(\boldsymbol{a}', \boldsymbol{b}') = (a_{\mathrm{HT}} + a_{\mathrm{HT}}^*)(b_{\mathrm{HH}} - b_{\mathrm{TT}}) \tag{19}$$

Since all diagonal coefficients have the value $\frac{1}{2}$, only the third of the above equations survives and we arrive at

$$\Delta' = |\left(a_{\mathrm{HT}} + a_{\mathrm{HT}}^*\right)\left(b_{\mathrm{HT}} + b_{\mathrm{HT}}^*\right)| \tag{20}$$

With $a_{\text{HT}} = u + iv$ and $b_{\text{HT}} = x + iy$ (19) yields

$$\Delta' = 4|ux| \tag{21}$$

It is easy to see that the spectrum of ρ_A is non-negative only if $u^2 + v^2 \leq a_{\text{HH}}a_{\text{TT}} = 1/4$. An analogous relation holds for x and y, and we finally obtain $\Delta' \leq 1$. For *real* a_{HT} and b_{HT} $\Delta' = 1$.

Let us now turn to the non-K-separable operator ρ_{c} defined by (12). ρ_{c} corresponds to the symmetric matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

With

$$\hat{A} \otimes \hat{B} = \begin{pmatrix} a_{\text{HH}}\hat{B} & a_{\text{HT}}\hat{B} \\ a_{\text{TH}}\hat{B} & a_{\text{TT}}\hat{B} \end{pmatrix} \tag{22}$$

a straightforward calculation yields $\Delta = 1$. So there is no difference in the value of the experimentally accessible correlation function, and in consequence it will not be possible to discriminate between K-separability and non-K-separability in the case of coin halves.

Of course nobody will rack his brains over this result, because coin halves are classical entities, and one should never treat them quantum mechanically. But, nevertheless, this simple gedanken experiment allowed us

- to introduce (non-)K-separability as *the* essential feature of statistical operators and
- to explain its conceptual consequences.

Both items will be useful in the next section.

4. A general EPR-type gedanken experiment

Both in the Bohm–Aharonov gedanken experiment [19] and in the experiments actually performed with ensembles of photon pairs, produced by parametric down-conversion, a rotationally variant property type as spin or polarization is used to evaluate Δ. So we require that the single measurements still yield yes–no decisions, that is, the eigenvalues of $\hat{A}$ and $\hat{B}$ must be equal to ± 1, but that these decisions depend on rotation.

$\hat{A}$ shall be given by (16) again. Apparatus B differs from A insofar as the vector $\boldsymbol{b}$ determining its actual internal status is not equal to $\boldsymbol{a}$. Let $\boldsymbol{a}$ be a principal axis in the laboratory coordinate system. Then $\hat{B}$ emerges from $\hat{A}$ by a rotation around the angle χ between $\boldsymbol{a}$ and $\boldsymbol{b}$.

$$\Rightarrow \hat{B} = \begin{pmatrix} \cos\chi & \sin\chi \\ \sin\chi & -\cos\chi \end{pmatrix} \tag{23}$$

In complete analogy we define two further operators, $\hat{A}'$ and $\hat{B}'$, where $\hat{A}'$ represents apparatus A rotated with respect to its first position (determined by $\boldsymbol{a}$) by an angle φ. $\hat{B}'$ stands for B rotated with respect to $\boldsymbol{a}$ by an angle ψ. Note that $\hat{A}$ and $\hat{A}'$ as well as $\hat{B}$ and $\hat{B}'$ are in general noncommuting. The determinants of the commutators attain their maximum if $\varphi = \frac{\pi}{2}$ and $\psi = \chi + \frac{\pi}{2}$, respectively. In the following these angle settings will be employed throughout.

In the 2×2 case the most general statistical operator is the one defined by (10). It is seen immediately that this operator is non-K-separable, and we will, therefore, denote it by ρ_{ns}. As has been shown in ref. 5, the calculation of Δ_{ns} is simplified significantly if χ, which has not been fixed until now, is made equal to $\frac{\pi}{4}$. We then obtain

$$\Delta_{\mathrm{ns}} = \sqrt{2}\left[|1 - 2(c_{11,22} + c_{22,11})| + |c_{12,12} + c_{12,21} + c_{21,12} + c_{21,21}|\right] \tag{24}$$

Due to the self-adjointness of $\rho_{\mathrm{n\,s}}$, and by writing $c_{12,12} = u + iv$, (23) reduces to

$$\Delta_{\mathrm{ns}} = \sqrt{2}\left[|1 - 2(c_{11,22} + c_{22,11})| + 4|u|\right] \tag{25}$$

It is easy to see that the maximum of the first absolute value amounts to 1. The maximum value of $|u|$, however, is dependent on the conditions imposed on the nondiagonal coefficients to guarantee the non-negativity of the spectrum. So in general we have

$$\Delta_{\mathrm{ns}} \leq (1 + 4|u|)\sqrt{2} \tag{26}$$

which, in the singlet case usually discussed in the EPR context, yields the well-known upper bound of $2\sqrt{2}$.

The most general K-*separable* operator is given by

$$\rho_{\mathrm{s}} = \left(\sum_{i,j} a_{ij}\hat{A}_{ij}\right) \otimes \left(\sum_{k,l} b_{kl}\hat{B}_{kl}\right) \tag{27}$$

see (13), where $a_{11} + a_{22} = 1$ and $|a_{12}|^2 \leq a_{11}a_{22}$. Analogous relations hold for the b_{kl}. Proceeding in the same way as above we obtain [5]

$$\Delta_{\mathrm{s}} = \sqrt{2}\left[|(a_{11} - a_{22})(b_{11} - b_{22})| + |(a_{12} + a_{21})(b_{12} + b_{21})|\right] \tag{28}$$

With $a_{12} = u + iv$ and $b_{12} = x + iy$ the second absolute value reduces to $4|ux|$. Because of $u^2 + v^2 \leq a_{11}a_{22}$ and $x^2 + y^2 \leq b_{11}b_{22}$, $|ux|$ is maximized if $v = y = 0$.

$$\Rightarrow \Delta_{\mathrm{s}} \leq \sqrt{2}\left[|(2a_{11} - 1)(2b_{11} - 1)| + 4\left(a_{11}(1 - a_{11})b_{11}(1 - b_{11})\right)^{1/2}\right] \tag{29}$$

Numerical evaluation yields $[\ldots] \leq 1$ so that

$$\Delta_{\mathrm{s}} \leq \sqrt{2} \tag{30}$$

The results (25) and (29) have been obtained for maximally noncommuting operators and $\chi = \frac{\pi}{4}$. In ref. 5, it has been shown that by use of *commuting* operators both Δ_{s} and Δ_{ns} have an upper bound of 2.

5. Discussion

The *rigorous* ensemble approach has been applied to the EPR problem for the first time. By *complete* abdication of the possibility to make any statements regarding individual entities the following results have been achieved:

1. Bell's inequality holds from the rigorous ensemble point of view, as well:

    - *If* there is no action-at-a-distance between the two detectors A and B, and
    - *if* the outcome A_i (B_i) of one single run does not depend on the setting of B (A), and if this is true for all N runs (parameter independence),

    then the correlation function Δ is bound from above by the value 2.

2. In the deduction of Bell's inequality, reference has been made neither to the representation of the ensembles nor to the operators representing the apparatuses, that is, the inequality is operator *independent* and *non*contextual. Therefore, it is an upper bound for those basic correlations only, which are due to a *conservation law*. This is reflected in Bell's inequality, no more and no less.

3. An ensemble is represented by a statistical operator ρ. The set of all statistical operators on a given Hilbert space $\mathcal{H}$, being the direct product of two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, consists of two mutually exclusive subsets, because ρ may be either K-separable or not. ρ is called K-separable if and only if it can be decomposed into a direct product of one statistical operator on $\mathcal{H}_A$ and another one on $\mathcal{H}_B$.

4. The upper bound of Δ depends on *both* the (non)-K-separability of the statistical operator in question *and* on the property type to be measured, that is, on the respective operators $\hat{A}$, $\hat{A}'$, $\hat{B}$, and $\hat{B}'$ characterizing the two apparatuses with their actual setting. So $\max(\Delta)$ is *contextual*. (This may be considered an operationalistic- and ensemble-based extension of ref. 20.)

5. Maximally noncommuting operators (in the sense that $det\left([\hat{A}, \hat{A}']\right)$, $det\left([\hat{B}, \hat{B}']\right)$ = extremum $\neq 0$) yield $\Delta_s \leq \sqrt{2}$ and $\Delta_{\mathrm{ns}} \leq 2\sqrt{2}$ if $\chi = \frac{\pi}{4}$. In contrast, commuting operators yield $\max(\Delta_{\mathrm{s}}) = \max(\Delta_{\mathrm{ns}}) = 2$.

6. Δ is accessible by experiment if four series of measurements are performed. *Each* of the series must consist of such a lot of single runs that the large-N limit is reached sufficiently for all practical purposes. *If*, during the whole experiment, there is no change in the device producing the entity ensembles *then* the four ensembles may be considered equivalent, and *only then* the statistical operator refers to each of them individually.

7. *Then and only then* it is possible to compare Δ_{exp} with Δ_{s} and Δ_{ns} obtained using ρ_{s} and ρ_{ns}, respectively, and the actual setting of the apparatuses. *If* $\Delta_{\mathrm{s}} \neq \Delta_{\mathrm{ns}}$ *then* the comparison allows for an unambiguous statement whether ρ_{s} or ρ_{ns} is the correct representation of the total ensemble, and in this case we know whether the subensembles can be separated from one another or not.

What does it mean if two subensembles, $\{U_i\}$ and $\{V_i\}$, created by successive decay of the elements $(UV)_i$ of a precursor-ensemble $\{(UV)_i\}$, are non-K-separable? Recall that we cannot make any statement regarding the *single* constituents of each of the subensembles, that is, we may not claim that, for example, U_k has an "existence" of its own. The individual entity remains veiled.

QM yields *statistics* of single events (pointer positions). At an ontological level, statistics demands an ensemble of entities that gives rise to a multitude of single measurement outcomes that allow for a statistical evaluation. On the other hand there are two different sets of descriptors for the ensemble of interest, namely, the set of all non-K-separable statistical operators on a given Hilbert space and the set of all K-separable ones. There is an overwhelming amount of evidence that, in EPR-type experiments, the operator underlying the statistics must be non-K-separable (see, however, refs. 21–23). So the corresponding subensembles are related although noninteracting. Taking into account the contextuality of this result, we can describe QM as an *operationally* holistic theory.

There is nothing curious about this holism. Two subensembles of entities produced in an EPR experiment "remember" their common origin, and due to this memory effect both retain knowledge of the original complete whole. This is *more* than a mere conservation law. A complete coin differs significantly from a "head" put beside a "tail". With a complete coin we can go shopping, but no sales clerk would ever accept a "head" put beside a "tail". So the whole is more than the sum of its parts.

All the features of $\{V_i\}$ are stored in $\{U_i\}$ as well — and vice versa. Whether we can extract *all* knowledge by experiment depends on how tricky the experiment is to perform. In view of this contextuality we call two subensembles EPR correlated if and only if it is experimentally possible to realize the Δ bound of $2\sqrt{2}$.

6. Summary

"Physics, in contrast to different pursuits, is the study of *reproducible* phenomena. In the microscopic realm it is an *empirical fact*, learned without *any* help from theory, that only the behavior of ensembles is, in general, reproducible, whereas that of individual systems is not." [24] This statement by Gottfried is a very stimulating hint to analyze quantum behavior from an ensemble point of view. In a previous paper [5], a general EPR-type experiment was formulated using an ensemble approach, and it has been shown that Bell's inequality is violated only if the statistical operator representing the ensemble is *non*separable with respect to a significantly simplified definition of separability discussed in ref. 15 (K-separability).

In this paper the said approach has been tightened in a *rigorous* manner. After application of the deduction of Bell's inequality, the consequences of K-separability and non-K-separability have been described in detail using a gedanken experiment with coin halves. Finally, the general experiment mentioned above has been re-examined. Without any doubt, the resume of all this is that, if QM is assumed to make statistical statements on the results of measurements on ensembles only, then the theory is operationally holistic.

A rigorous approach has never been used before to elucidate the actual core of EPR's discovery. Home and Whitaker pointed out 10 years ago that *some* ensemble interpretations are not able to provide us with a "*physical understanding* of *how* the experimental results emerge" [8], but the present author is convinced that the rigorous approach does, in fact, dispose of the EPR problem. The ingredients are

- to ignore the single entity completely and in every respect (especially regarding its assumed "reality"),
- the mathematical structure of ρ, and
- the thesis that individual subensembles do not have an individual "existence".

It should, however, be mentioned that this approach will hardly resolve the so-called measurement problem that arises if, in contrast to Bohr's opinion, both the quantum system *and* the measurement apparatus are treated on the QM level. It is to be assumed that a statistical operator such as, for example, $\rho_{\text{ensemble}} \otimes \rho_{\text{apparatus}}$ will yield statistics of macroscopic superpositions, thereby only shifting the domain of the reduction postulate without getting rid of it.

Let me conclude: there are zones of the World (in the farthermost sense) which exhibit non-K-separability. There the whole is more than the sum of its parts. This is the solution of the riddle.

References

1. F. Laloë. Am. J. Phys. **69**, 655 (2001).
2. M. Bunge. Int. J. Quant. Chem. **12**, 1 (1977).
3. A. Einstein, B. Podolsky, and N. Rosen. Phys. Rev. **47**, 777 (1935).
4. B. d'Espagnat. Conceptual foundations of quantum mechanics. Benjamin, Reading, Mass. 1976. Chap. 8.
5. T. Krüger. Found. Phys. **30**, 1869 (2000).
6. K. Gödel. Monatshefte Math. Phys. **38**, 173 (1931).
7. D. Home and F. Selleri. Riv. Nuovo Cimento, **14**, No. 9 (1991).
8. D. Home and M.A.B. Whitaker. Phys. Rep. **210**, 223 (1992).
9. D.T. Gillespie. Am. J. Phys. **54**, 889 (1986).
10. G. Ludwig. Foundations of quantum mechanics II. Springer, New York. 1985. Chap. XVII Sect. 4.4.
11. G. Ludwig. Foundations of quantum mechanics I. Springer, New York. 1983. p. 9.
12. J.S. Bell. Physics, **1**, 195 (1964).

13. J.S. Bell. Rev. Mod. Phys. **38**, 447 (1966).
14. Quoted from W.H. Zurek. Phys. Today, October 1991, p. 36.
15. T. Krüger. Z. Naturforsch. **56a**, 849 (2001).
16. A.F. Abouraddy, B.E.A. Saleh, A.V. Sergienko, and M.C. Teich. Phys. Rev. A, **64**, 050101(R) (2001).
17. C. Tsallis, S. Lloyd, and M. Baranger. Phys. Rev. A, **63**, 042104 (2001).
18. S.J. Lomonaco, Jr. LANL e-print archive, quant-ph 0007045 (2000), Chap. 7.4.
19. D. Bohm and Y. Aharonov. Phys. Rev. **108**, 1070 (1957).
20. S. Basu, S. Bandyopadhyay, G. Kar, and D. Home. Phys. Lett. **279A**, 281 (2001).
21. S. Sulcs, G. Oppy, and B.C. Gilbert. Found. Phys. Lett. **13**, 521 (2000).
22. L. Vaidman. Phys. Lett. **286A**, 241 (2001).
23. M. Iannuzzi and V. Merlo. Found. Phys. Lett. **14**, 355 (2001).
24. K. Gottfried. Phys. World, October 1991, p. 34.