1. Introduction

Since the publication of the study of Einstein, Podolsky, and Rosen (≡ EPR) on the assumed incompleteness of quantum mechanics (= QM) in 1935 [1] the discussion of the EPR gedankenexperiment did not come to a conclusive result. Meanwhile the debate has been boiled down to the question whether the relation between micro-entities is of local or non-local nature (see, e. g., [2]) or, more precisely, whether ensembles of micro-entities have to be described by a separable or a non-separable statistical operator [3,4]. In the latter case, none of two correlated but non-interacting ensembles would have an ‘existence of its own’. Following Bell’s work [5], these two possibilities give rise to different upper bounds of the correlation function $\Delta$ which is defined by

$$\Delta := |O(a, b) - O(a', b')| + |O(a', b) + O(a', b')|, \quad (1)$$

where the right-hand terms are explained by (2). All experiments performed so far lead to an impressive violation of $\Delta_{\text{separable}}$ (see, e. g., [6] and [7]) so that non-separability must be regarded a typical feature of the quantum world. The question, however, is: Which statistical operators cause a violation of Bell’s inequality and therefore represent non-separable ensembles? Since the notion of non-separability is given with respect to the precise meaning of separability only, one first of all has to compare the definitions of separability to one another. This is done in Section 3.1. Then the effect of these different approaches on the von Neumann entropy $S$ will be examined. $S$ is used to define a non-separability index which allows for a quantitative separability comparison of two given operators. Finally the applicability of a current separability criterion will be checked.

2. The Model System

We refer to a well-known gedankenexperiment: A given quantum entity shall decay at a time $t_0$ into two parts called $U_i$ and $V_i$. An experimental setup shall separate $U_i$ and $V_i$ spatially so that they cannot interact any more. $U_i$ ($V_i$) shall impinge on an apparatus A (B). A and B are of the same kind. At a time $t_1 > t_0$ a dichotomic property type $E$ shall be measured simultaneously on both $U_i$ and $V_i$, $E(U_i)$ and $E(V_i)$, the obtained properties (= numerical values of $E$), shall depend on certain experimental parameters $a$ and $b$, resp., which are assumed to determine the actual internal structure of the two apparatuses A and B. Then the correlated outcome of one single run is given by the product

$$O_i(a, b) = E(a; U_i)E(b; V_i). \quad (2)$$
The final result $O(a, b)$ is the mean of all single outcomes.

We associate the ensemble of produced or emitted entity pairs $(U_i, V_i)$ with a statistical operator $\rho$, i.e., with a self-adjoint operator of trace 1 which possesses a positive semi-definite spectrum. $\rho$ acts on a four-dimensional Hilbert space $\mathcal{H}$ which is the direct product of two orthogonal spaces $\mathcal{H}_U$ and $\mathcal{H}_V$. $\mathcal{H}_U$ ($\mathcal{H}_V$) is connected to the physical space wherein $E$ is measured on $U_i$ ($V_i$) using A (B). Let $\{|\alpha_1\rangle, |\alpha_2\rangle\}$ be an orthonormal basis of $\mathcal{H}_U$ and $\{|\beta_1\rangle, |\beta_2\rangle\}$ an orthonormal basis of $\mathcal{H}_V$.

3. Separability and Non-Separability of Statistical Operators

3.1. Definitions

A statistical operator may be either separable or not, but what, in fact, shall a separable operator represent? With this question we leave the realm of statements of merely mathematical interest and devote our attention to an ontological consideration. Every $\rho$ stands for an ensemble of entities. In the present case we have to deal with ensembles $\{(U_n, V_n)\}$ of pairs of entities which have a pairwise common origin. Said pairs are separated experimentally so that any possibility of an interaction is excluded. In this way we create two sub-ensembles, $\{U_n\}$ and $\{V_n\}$. The former ‘exists’ in the space region of apparatus A only whereas the latter solely ‘exists’ in the space region of B. So they may be thought of as being independent (in the quite trivial sense of the word). Out of this very reason it makes sense to represent the entire ensemble $\{(U_n, V_n)\}$ by a statistical operator

$$\rho_s := \rho(U) \otimes \rho(V),$$

where the sub-ensemble $\{U_n\}$ ($\{V_n\}$) is represented by $\rho(U)$ ($\rho(V)$) on $\mathcal{H}_U$ ($\mathcal{H}_V$) [3, 4]. This $\rho_s$ may be called separable, because it reflects the independence of the two sub-ensembles which allows to separate them, i.e., the ontological approach to separability results in the simple direct product form of $\rho_s$. Note that in the literature ‘states’ described by a statistical operator satisfying the condition (3) are called product ‘states’.

Usually, however, $\rho$ is called separable if it can be expanded in a series of products according to

$$\rho_s^{(\text{usual})} := \sum_{i=1}^{N} w_i \rho_i(U) \otimes \rho_i(V),$$

where the $\rho_i(U)$ ($\rho_i(V)$) are statistical operators on $\mathcal{H}_U$ ($\mathcal{H}_V$) and $\sum_{i=1}^{N} w_i = 1$, $w_i \geq 0 \forall i$. This definition goes back to Schrödinger. It reappeared in the present form in a paper of Werner [6] who called “a composite quantum system ... classically correlated if it can be approximated by convex combinations of product states”. Since that time Werner’s ansatz, which starts from the question of how to prepare a corresponding ‘state’, has found wide application.

Obviously the usual definition of separability is more general than the one supported here, but we have to ask whether there is any difference of practical relevance. Quantum correlations are investigated mostly with respect to their influence on certain property types as, e.g., spin or polarization. In this way it has been shown [3] that essential results of EPR-type experiments, obtained by use of $\rho_s$ and $\rho_s^{(\text{usual})}$ resp., are identical. But, in view of this fact, why then do we need this other definition of separability? We need it, because, in addition to its conceptual clarity, it is much more simple as well, since in most cases it can be seen at once whether a given operator is separable in the sense of (3) or not. Separability according to the usual definition, however, demands for special criteria, and often it is not easy to check whether the operator in question actually meets the criterion.

If in (3) both $\rho(U)$ and $\rho(V)$ are projectors, then also $\rho_s$ is a projector. In case of the usual definition, however, $(\rho_s^{(\text{usual})})^2 \neq \rho_s^{(\text{usual})}$ if all $\rho_i(U)$ and $\rho_i(V)$ are projectors. Only in the trivial case ($w_i = 1$ and $w_j = 0 \forall j \neq i$) we obtain $(\rho_s^{(\text{usual})})^2 = \rho_s^{(\text{usual})}!$

In Sect. 3.4 it will be shown that none of the two definitions gives “better” or “more strictly” separable operators in general. This is done by introducing an entropy-determined index of non-separability which rests on an idea of Barnett and Phoenix [8, 9] who have introduced an index of correlation $I_c$ which is independent of any property type, i.e., this index is a system immanent quantity.

To complete these considerations it should be mentioned that an operator is called non-separable iff it cannot be separated. Note, however, that non-separability is not the same as entanglement. Common definitions of entanglement pay attention to the question whether the operator is pure or not [7], but this difference does not play any role at all if the question of separability is concerned.
3.2. Separability and Entropy

The von Neumann entropy in QM is defined, in units of Boltzmann’s constant $k_B$, as

$$S = -\text{Tr}(\rho \ln \rho),$$

where $\rho$ acts on the total Hilbert space $\mathcal{H} = \mathcal{H}_U \otimes \mathcal{H}_V$. Every $\rho$ on $\mathcal{H}$ may be expanded according to

$$\rho = \sum_{k,l,m,n} u_{k|l} v_{m|n} U_{k|l} \otimes V_{m|n},$$

with $U_{k|l} = |a_k\rangle \langle a_l|$ and $V_{m|n}$ defined analogously. The operator $\ln \rho$ exists for all statistical operators and is given by the power series

$$\ln \rho = \sum_{k=1}^{\infty} (-1)^{k+1} (\rho - 1)^k / k$$

Each power of $(\rho - 1)$ can be transferred into a power of $\rho$ only so that the entropy of every $\rho$ is given by

$$S = -\text{Tr}(\rho \ln \rho) = -\sum_{k=1}^{\infty} \sum_{n=0}^{k} (-1)^{k+n+1} \binom{k}{n} k^{-1} \text{Tr}(\rho^{k-n+1}).$$

If $\rho$ is a projector, all $\text{Tr}(\rho^{k-n+1})$ are equal to 1, and we obtain $S = 0$ as expected.

Now consider the usual definition of separability. For the sake of simplicity the following analysis is confined to the $2 \times 2$ case. Then $\rho^{(\text{usual})}$ can be expressed by $\omega \rho_1 + (1 - \omega) \rho_2$ with $\omega \in [0, 1]$ and $\rho_1(2)$ being the direct product of one statistical operator $\rho_1(2)(U)$ on $\mathcal{H}_U$ and one, $\rho_1(2)(V)$, on $\mathcal{H}_V$. We assume that both $\rho_1$ and $\rho_2$ are given in their canonical representation. In this case we obtain

$$\rho_s^{(\text{usual})} = \omega a_{11} U_{11} + a_{12} U_{22} \otimes b_{11} V_{11} + b_{12} V_{22} + (1 - \omega) a_{21} U_{11} + a_{22} U_{22} \otimes b_{21} V_{11} + b_{22} V_{22},$$

where $a_{11} + a_{12} = a_{21} + a_{22} = b_{11} + b_{12} = b_{21} + b_{22} = 1$. Now the corresponding entropy can be computed as a function of $\omega$ and, if desired, some or all of the coefficients $a_{i,j}$ and $b_{i,j}$, using the power series expansion of $(\rho \ln \rho)$. (8) is then evaluated numerically in search of the respective maximum of $S$.

Assume first that both $\rho_1$ and $\rho_2$ are projectors. In this case $\max(S_s^{(\text{usual})})$, the maximal entropy for the given coefficients, amounts to $\ln 2$. Now let one of the operators deviate from the projector property by choosing, e.g., $a_{21} = b_{21} = 0.1$ and $a_{22} = b_{22} = 0.9$. This mixing results in an increase of $\max(S_s^{(\text{usual})})$, and it can be seen that the inequality $\max(S_s^{(\text{usual})}) \leq 2 \ln 2$ holds. The upper bound is attained if the mixture is homogeneous ($a_{21} = b_{21} = a_{22} = b_{22} = 0.5$). Note that only one of the two operators is a non-projector, but this inequality is valid even if both $\rho_1$ and $\rho_2$ are non-projectors, i.e., $2 \ln 2$ is the absolute maximum of the entropy of an operator which is separable according to the usual definition.

What can be said about the entropy of a $\rho$ which is separable according to the definition by (3)? If the canonical representation of $\rho(U)$ and $\rho(V)$ is presupposed, i.e., $\rho(U) = a_{11} U_{11} + (1 - a_{11}) U_{22}$ and $\rho(V)$ given analogously, then $S_s$ can be calculated quite easily in dependence of $a_{11}$ and $b_{11}$. If the coefficients are chosen so that both operators are projectors, we obtain $S_s = 0$. Allowing one of the operators to deviate from $\rho^2 = \rho$ results in $\max(S_s) = \ln 2$, and if both operators are no more projectors we obtain $\max(S_s) = 2 \ln 2$. In this respect the two separability definitions lead to a different dependence of the entropy on the degree of mixedness of the various operators involved, although the absolute maximum of $S = 2 \ln 2$ is the same. See Table 1 for a summary of these results.

For a certain class of statistical operators it can be proven that the upper bound of the von Neumann entropy is $2 \ln 2$ independent of i) how separability is defined, and ii) whether $\rho$ is separable at all: Consider the set $\mathcal{R}$ of all statistical operators which comply with the condition

$$\text{Tr}(\rho^k) = 1/r^{k-1} \text{Tr} \rho = r^{1-k}, \ r > 1 \forall k \geq 2.$$ (10)

This set contains all operators with a regular decrease in trace which is represented by the geometric progression $1/r, 1/r^2, 1/r^3, \ldots$ The parameter $r$ is a measure of the trace distance $\text{Tr} \rho - \text{Tr}(\rho^2)$.
Fig. 1. Entropy of an operator depending on the trace distance parameter $r$.

Insertion of (10) into (8) yields the entropy as a polynomial function of $r$. The result is shown in Figure 1. $S$ increases with increasing trace distance. If $r = 2$ we obtain $S = \ln 2$.

Is there any operator with $r$ larger than 2? For every $\rho \in R$, $\text{Tr}(\rho^3)$ must be equal to $1/r^2$ with $r = 1/\text{Tr}(\rho^2)$. Now, in the most general case $\rho$ is given by

$$\rho = \sum_{i,j} w_{ij} W_{ij},$$

so that

$$\text{Tr}(\rho^2) = \sum_i \left( \sum_j w_{ij} w_{ji} \right) := 1/r$$

and

$$\text{Tr}(\rho^3) = \sum_{i,l} \left( \sum_j w_{ij} w_{jl} \right) w_{li} := 1/r^2.$$  \(12\)

In the $2 \times 2$ case, by squaring (12) and equating it with (13), we obtain the condition

$$w_{11}^4 - 2w_{11}^3 + 5/4w_{11}^2 - 1/4w_{11}$$

$$+ |w_{12}|^2(2w_{11}^2 - 2w_{11} + |w_{12}|^2 + 1/4) = 0.$$  \(14\)

This equation has two allowed solutions, corresponding to $r = 1$ and $r = 2$, respectively. So there is no operator with $r > 2$, and due to the properties of the trace and the logarithm function it is evident that in a two-operator system as discussed above the maximum entropy is given by $2 \ln 2$.

3.3. An Example

In [4] an EPR-type experiment with molecules has been proposed which rests on the transfer of the correlation of molecular fragments to previously uncorrelated photon pairs. At least in principle the ensemble of photon pairs can be represented either by a separable or by a non-separable operator. For the former one we obtain

$$\rho_s = (\cos^2 \theta U_{11} + \sin^2 \theta U_{22}) \otimes (\cos^2 \theta V_{11} + \sin^2 \theta V_{22})$$  \(15\)

with $0 < \theta < \pi/2$. Regarding this experiment, the correct non-separable counterpart of $\rho_s$ is given by

$$\rho_{n-s} = (\cos^2 \theta U_{11} + \sin^2 \theta U_{22}) \otimes (\cos^2 \theta V_{11} + \sin^2 \theta V_{22})$$

$$- \cos^2 \theta \sin^2 \theta (U_{12} + U_{21}) \otimes (V_{12} + V_{21}).$$  \(16\)

The entropy of $\rho_s$ may be calculated quite easily, because there is a general formula for the trace of an arbitrary power of $\rho_s$:

$$\text{Tr}(\rho_s^n) = (\cos^{2n} \theta + \sin^{2n} \theta)^2.$$  \(17\)

In case of $\rho_{n-s}$, however, things are much more complicated due to the lack of a corresponding reduction formula. So the $(\rho \ln \rho)$ expansion had to be stopped after the fifth-order term which is sufficient for a principal comparison of $S_s$ and $S_{n-s}$. The result is shown in Figure 2. The entropy of the separable statistical operator is always larger than the entropy of the non-separable operator. While $\max(S_{n-s})$ is equal to $\ln 2$.
only, \( S_s \) reaches twice the value. The difference attains a maximum at \( \theta = \pi/4 \). It is worth to note that \( S_{n,s} \) does change only slightly between \( \theta = 30^\circ \) and \( \theta = 60^\circ \), irrespective of the fact that the prefactor of the interference terms has not yet achieved its maximum at the limit points of this interval.

High entropy corresponds to a low information content, and vice versa. So the non-separable operator must contain more information than its separable counterpart. This is obviously true, because in the non-separable case each of the two sub-ensembles represented has still a knowledge of the common origin. This knowledge is represented by the interference terms. If they are canceled to create a separable operator, then this knowledge gets lost automatically.

### 3.4. The Non-Separability Index

We define a measure of the extent of non-separability of a given operator \( \rho_a \) with respect to a second operator \( \rho_b \), independent of the observation of any property type, by

\[
I_{a/b} := -(S_a - S_b).
\]

(18)

Positive \( I_{a/b} \) indicates that \( \rho_a \) is non-separable relative to \( \rho_b \). If, in addition, \( \rho_b \) is separable, then \( \rho_a \) is non-separable in an absolute sense. Compare this definition to the proposal of Vedral et al. to use the quantum relative entropy as a measure of entanglement [10, 11]. See, however, [12].

It has already been mentioned that the functional dependence of \( S \) on the coefficients of the operators is different for an operator separable according to (3) and an operator of the Werner type. This behavior is reflected directly in the non-separability index \( I_{S/\text{usual}} \).

By numerical evaluation it is easy to see that \( I_{S/\text{usual}} \) can be positive as well as negative. If, e.g., \( \rho_{0.5/\text{usual}} = w \left( 0.2 \mathbf{U}_{11} + 0.8 \mathbf{U}_{22} \right) \otimes \left( 0.6 \mathbf{V}_{11} + 0.4 \mathbf{V}_{22} \right) + (1 - w) \left( 0.1 \mathbf{U}_{11} + 0.9 \mathbf{U}_{22} \right) \otimes \left( 0.1 \mathbf{V}_{11} + 0.9 \mathbf{V}_{22} \right) \), we obtain \( S_{\text{usual}} = 1.0543 \) in the case of equal shares \( (w = 0.5) \).

Now choose \( \rho_S \) to correspond to the first term of \( \rho_{\text{usual}} \). This results in \( S_S = 1.1650 \) and \( I_{S/\text{usual}} \) becomes negative. i.e., \( \rho_{\text{usual}} \) is non-separable with respect to \( \rho_S \), but if \( \rho_S \) corresponds to the second term, \( S_S \) becomes equal to 0.6285, so that the opposite is true. Obviously none of the two definitions gives "better" or "more strictly" separable operators in general!

In case of the operators involved in the example discussed above, \( I_{n,s/S} \) can be determined directly from Fig. 2 as a function of \( \theta \).

### 3.5. On a Common Separability Criterion

With respect to the usual definition it is not evident a priori whether a given operator is separable or not. For this reason several separability criteria have been proposed [13 - 17]. So it is an open question whether they are applicable to operators which are separable or non-separable according to the definition employed in the present contribution. In the case of \( 2 \times 2 \) and \( 2 \times 3 \) dimensions the criteria of Peres [13, 14] and of Horodecki, Horodecki, and Horodecki [15] (\( \equiv \text{PH}^3 \)) are concurrent, and we will concentrate on them.

Let \( \mathbf{R} \) be the density matrix corresponding to \( \rho \), i.e.,

\[
R_{ijpq} = \langle \alpha_i \beta_j | \rho | \alpha_p \beta_q \rangle = u_{ip} v_{jq} \tag{19}
\]

with \( \rho \) given by (6). Now define a new matrix \( \mathbf{S} \) according to

\[
S_{ijpq} := R_{pjiq} = u_{pi} v_{jq}. \tag{20}
\]

The indices belonging to the \( u \)-coefficients have been transposed, but not the other ones. This is not a unitary transformation, but nevertheless \( \mathbf{S} \) is Hermitian. It has been shown [13 - 15] that the non-existence of any negative eigenvalue of \( \mathbf{S} \) is a necessary and sufficient condition for the decomposability of any \( \rho \) according to the usual definition.

Now the question arises whether the \( \text{PH}^3 \) criterion can discriminate correctly also those operators which are separable or non-separable according to (3). Let us, e.g., consider the already mentioned \( \rho_{n,s} \).

\[
\Rightarrow \mathbf{R} = \begin{pmatrix}
\cos^4 \theta & 0 & 0 & -\cos^2 \theta \sin^2 \theta \\
0 & \cos^2 \theta \sin^2 \theta & -\cos^2 \theta \sin^2 \theta & 0 \\
0 & -\cos^2 \theta \sin^2 \theta & \cos^2 \theta \sin^2 \theta & 0 \\
-\cos^2 \theta \sin^2 \theta & 0 & 0 & \sin^4 \theta
\end{pmatrix} \tag{21}
\]

Forming the partial transpose \( \mathbf{S} \) results in a sign change of the elements \((2,2), (2,3), (3,2), \) and \((3,3)\).

This matrix has the eigenvalues \( \lambda_{1,2} = 0 \), \( \lambda_3 = (3 + \cos(4\theta))/4 \), and \( \lambda_4 = -2 \cos^2\theta\sin^2\theta \) so that for all \( \theta \neq 0, \pi/2, ... \) there is exactly one negative eigenvalue \( \lambda_A \) which implies that \( \rho_{n,s} \) is non-separable according to the \( \text{PH}^3 \) criterion as well. To corroborate this statement we may examine whether it is possible to write \( \rho_{n,s} \) in the form of (4). So we make the ansatz...
\[
\rho_{\text{trial}} = w \rho_1(U) \otimes \rho_1(V) + (1-w) \rho_2(U) \otimes \rho_2(V) \tag{22}
\]
with
\[
\rho_i(U) = u_{i11}U_{11} + u_{i12}U_{12} + u_{i21}U_{21} + u_{i22}U_{22} \tag{23}
\]
and the other operators given analogously. By inserting (23) into (22) we obtain expressions of the kind \( U_{12} \otimes V_{12} \) which have no counterpart in \( \rho_{\text{n-sep}} \). For arbitrary \( w \) these operator products can be removed only if \( u_{112} = u_{121} = v_{i12} = v_{i21} = 0 \forall i \). However, the consequence of this restriction is that also the prefactors of expressions as \( U_{12} \otimes V_{12} \) vanish so that there is finally no chance to reconstruct \( \rho_{\text{n-sep}} \) from \( \rho_{\text{trial}} \). This is what we wanted to show.

4. Summary

The recently proposed definition of a separable statistical operator as a onefold factorable one offers several advantages with respect to the usual definition: It is both conceptually and mathematically simple, and it allows for an ontological interpretation of separability in contrast to the touch of operationalism inherent in the usual approach. It is an open question whether common separability criteria are still applicable, but in view of the enormous simplicity of the present approach one will be able to do without these criteria anyway. It has already been shown that the upper bound of the \( \Delta \) function does not depend on the choice of the separability definition [4].

When a statistical operator \( \rho \) is separable according to Krüger, then it is also separable according to the usual definition. At least in one relevant case it has been shown that an operator which is non-separable according to Krüger is non-separable according to the usual definition as well.

A relative non-separability measure has been developed which allows for a simple characterization of statistical operators.

Results for various systems and applications have been presented, including operators previously used in the description of an EPR-type experiment with molecules. Also the entropic behavior of the set of all operators with a regular decrease in trace from \( \rho^n \) to \( \rho^{n+1} \) in analogy to the geometric progression has been examined.