Unexpected reemergence of the von Neumann theorem

Marek Żukowski
Institute of Theoretical Physics and Astrophysics, University of Gdańsk, ul. Wita Stwosza 57, PL-80-952 Gdańsk, Poland
(Received 31 August 2008; published 23 February 2009)

It is shown here that the "simple test of quantumness for a single system" of Alicki and Van Ryn has exactly the same relation to the discussion of the problem of describing the quantum system via a classical probabilistic scheme (that is in terms of hidden variables or within a realistic theory) as the von Neumann theorem [*Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932); *Mathematical Foundations of Quantum Mechanics* (Princeton University, Princeton, NJ, 1955)]. The latter one was shown by Bell [Rev. Mod. Phys. **38**, 447 (1966)] to stem from an assumption that the hidden variable values for a sum of two noncommuting observables have to be, for each individual system, equal to sums of eigenvalues of the two operators. One cannot find a justification for such an assumption to hold for noncommeasurable variables. On the positive side, the criterion may be useful in rejecting models which are based on stochastic classical fields. Nevertheless, the example used by the authors has a classical optical realization.

DOI: 10.1103/PhysRevA.79.024103 PACS number(s): 03.65.Ud

I. INTRODUCTION

The no-go theorem for classical probabilistic (i.e., hidden variable) description of single-quantum systems in Ref. [1], called by Alicki and Van Ryn *a simple test of quantumness for a single system*, reads

> If one can find two *non-negative* observables A and B, such that $B-A$ is a non-negative operator too, and a state for which the averages of these observables and their squares have the following property $0 \leq \langle A \rangle_{av} \leq \langle B \rangle_{av}$ and $\langle A^2 \rangle_{av} > \langle B^2 \rangle_{av}$, then these measurements on the system are not describable in terms of "minimal classical models."

This formulation of the test is in Ref. [2], in which the claim is scaled down a bit. Also, Alicki and Van Ryn admitted in [2] that the criterion does not test *general hidden variable models* but rather "eliminated a well-defined classical theory for some specific systems."

While one cannot challenge the logical value of this thesis, the following question can be put: which specific hidden variable models are excluded by the criterion? It will be shown here that the criterion in [1,2] excludes the same class of hidden variable models as the von Neumann theorem [3]. The latter one was shown [4] to be too restrictive in its assumptions to be useful in the discussion of whether one can find classical probabilistic models of quantum-mechanical processes. It will be shown that minimal classical models are equivalent to von Neumann's assumptions on properties of "dispersion free states." These assumptions are doubtful; see [4,5].

The usefulness of the criterion in the discussion of the foundations of quantum mechanics is highly limited. Nevertheless, it may be useful in pinpointing phenomena which have no description in terms of stochastic classical fields. However, the example given in [1,2] and realized in [6] does have a classical model like every second-order (in terms of the fields) photonic interference effect. The observed phenomena can be interpreted as nonclassical only due to the statistical properties of the parametric down-conversion (PDC) process, which are revealed in correlation experiments (at least fourth-order ones).

II. RELATION WITH THE VON NEUMANN THEOREM

"Any real linear combination of any two Hermitian operators [say A and B] represents an observable, and the same linear combination of its expectation values is the expectation value of the combination (quotation after Bell [4])." If one, following von Neumann, assumes that the same rule must hold also for all dispersion free states (i.e., deterministic classical models, the average over which should give the quantum averages), this immediately transfers this rule to the possible experimental results. That is, for the hidden dispersion free states von Neumann [3] tacitly assumed that $v(A+B)=v(A)+v(B)$, even if the quantum observables do not commute (are noncommeasurable). In more physical terms, if one has a system governed by hidden variables, then the allowed pure (classical) states are such that the above rule holds in each individual run of an experiment. But as the very conditions to measure $A+B$, A and B are different, see footnote [12], there is no reason whatsoever to assume this [4]. The von Neumann no-go theorem [3] is inconsequential [5].

Alicki and Van Ryn [1,2] based their reasoning on the following theorem: $0 \leq A \leq B \Rightarrow A^2 \leq B^2$ holds for all A and B belonging to an algebra $\mathcal{A}$ if and only if the algebra $\mathcal{A}$ is commutative (i.e., isomorphic to the algebra of continuous functions on a certain compact space). Therefore, the task now is to show that the assumptions of von Neumann [3] when applied to two observables A and B are sufficient and necessary for the following: if for all states ρ one has $0 \leq \mathrm{Tr}(\rho A) \leq \mathrm{Tr}(\rho B)$ then $\mathrm{Tr}(\rho A^2) \leq \mathrm{Tr}(\rho B^2)$.

To prove the sufficiency one can start with a variant of the assumption of von Neumann: $v(B-A)=v(B)-v(A)$. Since the eigenvalues of $B-A$ are non-negative one must have $v(B-A) \geq 0$. Thus $v(B) \geq v(A)$ for each individual hidden variable steered system (dispersion free state). The non-negativity of A and B implies that $v(A) \geq 0$ and $v(B) \geq 0$. This obviously implies that for an individual system $v(B^2)$

$=[v(B)]^2 > v(A^2)=[v(A)]^2$. The equations stem again from the quantum rules concerning eigenvalues of *commuting* observables, whereas the inequality has only an algebraic origin. This implies that after averaging, for all triples $A \geq 0$, $B \geq 0$, and $B-A \geq 0$, one always has

$$\langle A\rangle_{\text{av}} \leq \langle B\rangle_{\text{av}} \Rightarrow \langle A^2\rangle_{\text{av}} \leq \langle B^2\rangle_{\text{av}}. \quad (1)$$

The next task is to prove that the above rule [Eq. (1)] implies the von Neumann assumptions [3]. This can be done using the very theorem on the C^* algebras that Alicki and Van Ryn [1,2] used to get their result. Let $\mathcal{A}$ be a C^* algebra: if all pairs, for which one has $B \geq A \geq 0$, follow that $B^2 \geq A^2$, then $\mathcal{A}$ is a commutative algebra. With commutativity the von Neumann assumptions [3] are true for all observables A and B. There is no problem with commeasurability, and even pure *quantum* states which are dispersion free for these observables (i.e., their eigenstates) satisfy the von Neumann assumption [3] $v(B-A)=v(B)-v(A)$.

III. NONMINIMAL MODELS

Can one give a hidden variable model that reproduces the predictions of the example given in Refs. [1,2]? Of course, such a model will not be "minimal." However, this will not be put here as such a model can be found already in [6]. As a matter of fact already in Ref. [7] one can find an explicit hidden variable model for all (von Neumann) measurements on spin 1/2 (qubit).

One can ask the following question concerning the interpretation of the mathematical result in [1,2]. If one can find two observables $A \geq 0$ and $B \geq 0$, such that $B-A \geq 0$, and a state for which

$$\langle A\rangle_{\text{av}} \leq \langle B\rangle_{\text{av}}$$

and

$$\langle A^2\rangle_{\text{av}} > \langle B^2\rangle_{\text{av}}, \quad (2)$$

what is the implication of this property for the question of existence of classical probability models of such averages? Definitely this is not a general impossibility of a classical probabilistic model for these two observables, at least for one qubit (because of the existence of the aforementioned model in Ref. [6]). However, it is very easy to show that one can have a plethora of hidden variable models with such a property. Simply, as it will be shown below, the conjunction of these inequalities can be achieved if the observable A is governed by a hidden variable λ_1 whereas the observable B is governed by an independent hidden variable λ_2. Then inequality (1) in [1,2] does not apply, see [13], and it is definitely not impossible that

$$\int A(\lambda_1)\varrho_1(\lambda_1)d\lambda_1 \leq \int B(\lambda_2)\varrho_2(\lambda_2)d\lambda_2$$

and

$$\int [A(\lambda_1)]^2\varrho_1(\lambda_1)d\lambda_1 > \int [B(\lambda_2)]^2\varrho_2(\lambda_2)d\lambda_2,$$

while the joint hidden variable distribution is given by

$$\varrho(\lambda_1,\lambda_2) = \varrho_1(\lambda_1)\varrho_2(\lambda_2).$$

Of course such a model is not minimal anymore.

One can always build such a model for a given pair of quantum observables, for any quantum state ρ, by the following construction (which is given here for d-state systems):

(i) denote the eigenvalues of A and B by A_i and B_j, respectively, with $i,j=1,2,\ldots,d$, and calculate the quantum probabilities for the given state for getting these results, $P(X_k|\rho)$, with $X=A,B$ and $k=1,2,\ldots,d$;

(ii) put for the hidden (variables) probability: $P_{HV}(A_i,B_j)=P(A_i|\rho)P(B_j|\rho)$; see [14].

Such a construction is universal, and thus it applies also to pairs of non-negative quantum observables which have the property $\text{Tr}[\rho(B-A)] \geq 0$ for all ρ. The hidden probabilities P_{HV} reproduce correctly the quantum predictions for A^2 and B^2 for all states even if $\text{Tr}\,\rho(A^2-B^2) \geq 0$. For different state preparations ρ we have a different P_{HV}, but this is allowed even in classical physics. The model involves many hidden variables, but one can always have as many hidden variables as one wishes because they are hidden anyway.

Thus, the condition in [1,2] is not a condition of a genuine quantumness but rather one concludes that if condition (2) holds, then there is no chance to have a classical model in which for every individual system (or dispersion free state) one can put $v(A-B)=v(A)-v(B)$, that is for which the von Neumann assumption [3] is valid.

Minimality loophole

In Ref. [2] Alicki and Van Ryn scaled down their claim to the following: condition (2), if satisfied, prohibits a minimal classical model for the observables. Thus, what was showed above is that minimal classical models in [2] are equivalent to the von Neumann assumptions [3] and thus face the criticism in [4,5].

Let us now address the discussion of the "minimality loophole;" Ref. [2], page 3. One can read: "One could still argue that there may exist a nonminimal classical algebraic model (AM) describing the data (minimality loophole). In this case, the classical observable $B-A$ possesses negative outcomes (values of the function) which are not detectable by the differences of averages $\langle B\rangle-\langle A\rangle$." However, for noncommuting A and B, since $B-A$ is not commeasurable with neither A nor B, and even the latter ones are noncommeasurable too, there is no reason for the eigenvalues to follow the von Neumann rule [3]. Thus, for an individual system $v(B)-v(A)$ may be negative, while $v(B-A) \geq 0$. One can explicitly construct a nonminimal model without negative $v(B-A)$:

(i) denote the eigenvalues of A, B, and $B-A=C$ as A_i, B_j, and C_k (*all of them are non-negative*), respectively, with $i,j,k=1,2,\ldots,d$, and calculate the quantum probabilities for the given state for getting these results, $P(X_k|\rho)$ with $X=A,B,C$ and $k=1,2,\ldots,d$,

(ii) put for the hidden (variables) probability: $P_{HV}(A_i,B_j,C_k)=P(A_i|\rho)P(B_j|\rho)P(C_k|\rho)$.

Obviously, the marginals of such a distribution produce correct probability distributions for all the variables, e.g.,

$\Sigma_{A_i}\Sigma_{B_j}P_{HV}(A_i,B_j,C_k)=P(C_k|\rho)$. Note that the set of allowed values C_k gives $v(B-A)$'s. Such hidden variable models are highly contextual, and thus uninteresting, but since $C=B-A\geq 0$, there is no problem with $B-A$ possessing negative outcomes. Thus, one has a direct counterexample to the quoted claim.

IV. WELL DEFINED CLASSICAL MODELS

As far as a *direct* detection of nonexistence of *any* classical probabilistic models is concerned, we are left with the two theorems of Bell, see [5], which do not use the von Neumann assumptions [3] to limit the hidden variable theories. The two theorems are based, except for realism, on noncontextuality assumption [4] or the locality assumption [7]. The latter one is based on a very strong relativistically motivated criterion of direct causal independence of events which are spatially separated (they can have a common cause but cannot influence each other directly) and the additional "natural" assumption that one may have stochastic processes which are statistically independent (for details see [8]).

However, it would be interesting to find a useful realm of applicability of the criterion in [1,2]. Alicki and Van Ryn [1,2] wrote that the criterion rules out minimal classical description. Such a description was shown above to be equivalent to the von Neumann theorem [3] but does this make it useless, like the theorem? Certainly not. It will be shown here that the condition makes impossible a classical probabilistic description which is using the tools of classical field theory or aims at describing the system via phase-space methods (e.g., Wigner quasidistributions), as suggested by Alicki and Van Ryn [1,2]. Simply, under the condition [1,2], an attempt to use such methods would fail—because the quasiclassical description requires the possibility of having either singular and/or nonpositive distributions. Thus the form of nonclassicality detected by the criterion in, e.g., quantum optics is limited to the phenomena which do not have a model in terms of statistical distributions of random classical electromagnetic fields.

To illustrate this, let us use the example in Refs. [1,2] and embed it into quantum optics of a single mode field. This can easily be done, e.g., by assuming that the sole two eigenstates of the operator A are $|0\rangle$ and $|1\rangle$, i.e., the vacuum state and the one-photon state. Thus the eigenstates of the operator B are some linear combinations of the two. So is the pure state $|\varphi\rangle=0.391|0\rangle+0.920|1\rangle$, which gives the optimal realization of the criterion: $\langle B^2\rangle > \langle A^2\rangle$ despite $A\geq B\geq 0$.

A direct comparison with classical theory can be made if one uses the P representation. It is based on the overcomplete continuous basis of coherent states, denoted here as $|\alpha\rangle$. In this formalism a general density operator reads

$$\rho=\int P(\alpha)|\alpha\rangle\langle\alpha|d^2\alpha, \tag{3}$$

with $d^2\alpha=d(\mathrm{Re}\ \alpha)d(\mathrm{Im}\ \alpha)$. The operators are determined by their diagonal matrix elements, e.g., $\langle\alpha|A|\alpha\rangle$, which are usually denoted as $A_Q(\alpha,\alpha^*)$. The averages are given by

$$\mathrm{Tr}\ A\rho=\int A_Q(\alpha,\alpha^*)P(\alpha)|\alpha\rangle\langle\alpha|d^2\alpha. \tag{4}$$

That is the operators are represented by functions, and therefore the picture apparently looks classical. However, nonclassical phenomena occur when $P(\alpha)$ are not non-negative or are more singular than a delta function. It is well known that superpositions of two Fock states are highly singular, thus suitably selected observables must reveal the impossibility of treating relation (4) as a model employing probabilistic distribution of classical amplitudes and a functional representation of the observables.

Thus the criterion in [1,2] belongs to the same class of nonclassical effects as antibunching or 100% interferometric contrast of products of intensity fluctuations observed by multiple detectors, etc. More general claims based on the criterion are unfounded.

These remarks on the example of Alicki and Van Ryn [1,2] given above were tailored such that the *best* features of the criterion were stressed. However, one can also present a "complementary" discussion of the specific example given in [1,2]. It will be shown that the example has not only a classical model but even a classical realization. Let us take a balanced Mach-Zehnder interferometer. It is well known that such a device is capable of performing any unitary $\mathcal{U}(2)$ transformation on a photonic qubit (with the two distinguishable states of being in the upper beam and in the lower beam). But it also well known that there is no distinction between second-order interference (in terms of fields) in the quantum and classical realm; see [15]. Thus if one takes as the inputs to the two entry ports of the interferometer, 1 and 2, two classical analytic signals $I(t)_i=a_iI(t)$ with amplitudes, a_i, as in $|\varphi\rangle$ [and both signals following the same temporal behavior, $I(t)$], then in the output ports one receives $a_i'I(t)$, with $a_i'=\Sigma_{i,j=1,2}U_{ij}a_j$ (we skip the retardation effects). The response of a detector is proportional to the intensity of the field impinging on it. Thus, the probability to register a count in output i is given by $\langle\langle|a_i'I(t)|^2\rangle\rangle$, where $\langle\langle\ \rangle\rangle$ denotes some time integration over the detector's time resolution.

Therefore one can tune the interferometer in such a way that it unitarily transforms qubits of amplitudes (1,0) and (0,1) into two basis states of the observable A; i.e., it performs a transformation $U(A)_{ij}$ (the other tunings will be for B and $B-A$ of the example). The experimentally observable averages are given by

$$\langle X\rangle_{\mathrm{av}}=\frac{\sum_{i=1,2}\lambda(X)_i\langle\langle|a_i'I(t)|^2\rangle\rangle}{\langle\langle|I(t)|^2\rangle\rangle}, \tag{5}$$

where $\lambda(X)_i$ are equal to the eigenvalues of the observables $X=A$, B, or C of the example and a_i' now stands for $\Sigma_{i,j=1,2}U_{ij}(X)a_j$. Just a glance reveals that the predictions of the quantum example and these classical models are identical. Even $\langle A-B\rangle_{\mathrm{av}}$ is always positive [16].

The same algebra holds for the polarization version of the experiment, as in [6]. Nonclassicality may be shown only if one considers that photonic nature of light may introduce antibunching effects at the detection stations or nonclassical correlations between the trigger (idler) and the detectors measuring the polarization of the signal photon (as it was done in [6]). Thus in the experiment [6] the sole nonclassicality is due to the statistical properties of the PDC process [17] and cannot be analyzed directly using only the observables of the example in [1,2].

The presented results of this section, despite claims in [10], are concurrent with those presented in [9,10] which pertain to analysis of optical experiments, which involves the field-theoretic approach. The conclusion in [10] is that in this context "only experiments with single photons can show directly deviations from classical probabilistic model." This exactly why statistical properties of the PDC process were used in the discussed experiments [6], as they allow one to have a heralded ("event ready" [11]) single photons. Such a state preparation gives nonclassical phenomena. It is well known that, e.g., very weak coherent pulses do not reveal any nonclassicality neither in interference nor in correlations (e.g., antibunching is impossible no matter how weak is the pulse). As it was argued above this is the sole nonclassical feature of the experiments, which is not reflected in the probabilities required to compute the average values of the considered observables [Eq. (5)]. Note that the fact that the nonclassicality criterion is equivalent to von Neumann's theorem is now admitted by Alicki ([10], Sec. 3).

ACKNOWLEDGMENTS

This work received support from the EU 6FP program SCALA. It was done at National Centre for Quantum Information of Gdansk. Discussions with Professor Alicki and Professor Genovese are acknowledged.

[1] R. Alicki and N. Van Ryn, e-print arXiv:0704.1962.
[2] R. Alicki and N. Van Ryn, J. Phys. A: Math. Theor. **41**, 062001 (2008).
[3] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932); *Mathematical Foundations of Quantum Mechanics* (Princeton University, Princeton, NJ, 1955).
[4] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[5] N. D. Mermin, Rev. Mod. Phys. **65**, 803 (1993).
[6] G. Brida, I. Degiovanni, M. Genovese, V. Schettini, S. Polyakov, and A. Migdall, Opt. Express **16**, 11750 (2008); for the latest version of such an experiment see G. Brida, I. P. Degiovanni, M. Genovese, F. Piacentini, V. Schettini, N. Gisin, S. V. Polyakov, and A. Migdall, e-print arXiv:0811.3376.
[7] J. S. Bell, Physics **1**, 195 (1964); for a review see M. Genovese, Phys. Rep. **413**, 319 (2005).
[8] R. Gill, G. Weihs, A. Zeilinger, and M. Zukowski, Proc. Natl. Acad. Sci. U.S.A. **99**, 14632 (2002); Europhys. Lett. **61**, 282 (2003).
[9] R. Alicki, M. Piani, and N. Van Ryn, J. Phys. A: Math. Theor. **41**, 495303 (2008).
[10] R. Alicki, e-print arXiv:0810.4708.
[11] M. Zukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, Phys. Rev. Lett. **71**, 4287 (1993).
[12] Because none of the pairs of out of these three observables commutes. In the laboratory this implies different operational situations.
[13] The inequality is $0 \leq A(x) \leq B(x)$, where, in the language of this comment, x is a hidden variable. This is an assumption of a specific distinguishing property of the minimal models, which is not derivable from the properties of the observed averages.
[14] Here the measurement outcomes, X_k, play the role of hidden variables.
[15] In simpler words there is no distinction in the Young-type interference between classical fields and those revealing quantum nature.
[16] One could argue here that one might split the incoming classical beams into two pairs and send them into two differently tuned Mach-Zehnder interferometers and then observe fluctuations giving negative values, but such an experiment does not model the two output situation assumed in the example of [2].
[17] The statistics of PDC radiation enables the heralding employed in the experiment: upon registration of an idler photon one can expect just one click (single photon) at just one of the detectors in the observation station behind the polarization analyzers. This feature is missing in the classical model presented here. For the experiment one could use other sources of single photons, such as quantum dots, color centers in diamonds, etc. In such a case the single photon, antibunched, nature of emissions is warranted by the workings of the source.