Pramāṇa, Vol. 19, No. 2, August 1982, pp. 159–173. © Printed in India.

When is a hidden variable theory compatible with quantum mechanics?

M D SRINIVAS
Department of Theoretical Physics, University of Madras, Madras 600 025, India

MS received 10 March 1982

Abstract. This paper is devoted to a study of some of the basic conditions which have to be satisfied by a hidden variable theory in order that it can reproduce the quantum mechanical probabilities. Of course one such condition, which emerges from the important theorem of Bell, is that a hidden variable theory has to be non-local. It is shown that a hidden variable theory is also incompatible with the conventional interpretation of mixed states and the mixing operation in quantum theory. It is therefore concluded that, apart from being non-local, a hidden variable theory would also necessarily violate the usual assumption of quantum theory that the density operator provides an adequate characterization of any ensemble of systems, pure or mixed.

Keywords. Hidden variable theories; complete specification of the state of a system; compatibility with quantum mechanics; local causality; density operators; mixed ensembles; quantum mechanics.

1. Introduction: The framework of hidden variable theories

In conventional non-relativistic quantum theory the state of a system is represented by a density operator ρ on a Hilbert space $\mathcal{H}$. A measurement performed on the system (normally considered to be instantaneous) is characterised by a self-adjoint operator A on $\mathcal{H}$, such that the probability that the result of the measurement lies in a Borel set $\Delta \subset R$, is given by

$$\Pr_A^\rho(\Delta) = \mathrm{Tr}\,(\rho\, P^A(\Delta)), \tag{1}$$

where $\Delta \to P^A(\Delta)$ is the spectral measure associated with the self-adjoint operator A. Usually, equation (1) is interpreted as implying that if a large number of systems (*i.e.* an ensemble) is prepared according to a procedure characteristic of the ' quantum state ' ρ, and if each of them is subjected to a measurement characterised by the ' observable ' A, then $\Pr_A^\rho(\Delta)$ is the fraction of systems which yield an outcome in the set $\Delta \subset R$.

For various motivating reasons, it has been argued (ever since the inception of quantum theory) that the above specification of the state of a system in quantum theory should be ' completed ', so that the completed theory would predict the actual outcomes of experiments performed on individual systems, and not merely the statistics of the outcomes of experiments performed on ensembles. Such a ' completion ' is sought to be achieved by constructing a ' hidden variable theory ' where one in-

cludes in the characterization of the state of a system, certain additional variables or parameters, (which are ' hidden ' as far as quantum theory is concerned), over and above the quantum-theoretic state of the system. The general framework of such hidden variable theories is the following (For general reviews of the subject, we refer to Belinfante 1973; Bell 1966, 1971a; Bohm and Bub 1966; Bub 1974; Capasso *et al* 1970; Clauser and Shimomy 1978; d'Espagnat 1976, 1979; Fine 1976; Kochen and Specker 1967; Jammer 1974; Roy 1980; Selleri and Tarozzi 1981; Vandana Shiva 1978; Virendra Singh 1980 and Wigner 1980).

Associated with each quantum system is a measure space (Γ, B) such that the following conditions are satisfied:

(I) With each density operator ρ is associated a probability measure μ^ρ on (Γ, B); with each self-adjoint operator A is associated a real valued random variable $X_A: \Gamma \to R$.

(II) For each state ρ and observable A,

$$\mathrm{Pr}^\rho_A(\Delta) = \mathrm{Tr}\,[\rho P^A(\Delta)] = \mu^\rho [X_A^{-1}(\Delta)], \tag{2}$$

for all Borel sets $\Delta \subset R$.

The interpretation of the above framework is the following. Γ is the set of all ' completely-specified states ' (or the so-called ' dispersion-free states ') which are now characteristic of individual systems; in other words, each point $\lambda \in \Gamma$ is thought of as providing a ' complete-specification ' of the state of an individual system. In the general framework discussed in this paper, no further assumptions are made regarding the space Γ, though in many models it turns out that each point $\lambda \in \Gamma$ is specified by a collection of variables $(\rho, \theta_1, \theta_2, ..)$ where ρ is a density operator (*i.e.* the quantum-theoretic state of the system) and $\theta_1, \theta_2, ..$ are the additional (or ' hidden ') variables needed to complete the quantum-theoretic description. Now, if an individual system is in the completely-specified state $\lambda \in \Gamma$, then the value of an observable A, (that one obtains in a measurement of A), is the number $X_A(\lambda)$. Also, if an ensemble of systems is prepared according to a procedure characteristic of the quantum state ρ, then associated with this ensemble is a probability measure μ^ρ on Γ, which describes the way in which the various individual systems of the ensemble are distributed over the points of Γ. The requirement (II) is precisely the condition that the individual systems of the ensemble are distributed over the points of Γ in such a way that the statistical distribution of experimental outcomes is exactly the same as that predicted by quantum theory.

Apart from (I) and (II), another requirement that we need to impose on a hidden variable theory is the following spectrum rule (see Fine 1976):

(III) For each observable A,

$$X_A(\lambda) \in \sigma(A); \tag{3}$$

where $\sigma(A)$ is the spectrum of A.

This requirement follows from the generally accepted principle of quantum theory that in any experiment to measure A, the outcome is always a number belonging to the spectrum of A.

It is well-known (Bell 1966; Bub 1974; Fine 1976; Kochen and Specker 1967; Wigner 1970) that there do exist models which satisfy the requirements (I) to (III). In fact to show that (I) to (III) are consistent, we only need to consider the space

$$\Gamma = \prod_{A \text{ (all observables)}} \sigma(A) \tag{4}$$

together with the canonical projection $\prod_A : \Gamma \to \sigma(A)$. Then the random variable X_A associated with the observable A, can be defined to be

$$X_A = I_A \circ \prod_A, \tag{5}$$

where I_A is the identity map on $\sigma(A)$. If we define the measure μ_A^ρ on $\sigma(A)$, by

$$\mu_A^\rho(\Delta) = \Pr_A^\rho(\Delta), \tag{6}$$

then we can define the measure μ^ρ on Γ by

$$\mu^\rho = \prod_A \mu_A^\rho. \tag{7}$$

It is trivial to check now that conditions (I)–(III) are satisfied.

Historically, there has been quite some confusion surrounding the possibility of hidden variable theories, mainly because of the several proofs of ' impossibility ' of such theories. These proofs only demonstrated the incompatibility of (I)–(III) with some additional requirements, mostly in the nature of certain additional conditions on the map $A \to X_A$ (between the self-adjoint operators and the associated random variables) such as the following:

(a) $$g(X_A) = X_{g(A)} \tag{8}$$

for all Borel functions $g : R \to R$; or equivalently,

$$X_A^{-1}(\Delta) = X_{PA(\Delta)}^{-1}(1) \tag{9}$$

for all Borel subsets $\Delta \subset R$;

(b) $$X_{A+B} = X_A + X_B, \tag{10}$$

for all compatible observables A, B;

(c) $$X_{AB} = X_A X_B \tag{11}$$

for all compatible observables A, B.

Ever since the critique of Bell (1966) it has been generally recognised that these additional conditions (on the map $A \to X_A$) are not motivated by any physical arguments, and in fact their violation is actually understandable from the basic principles

of the quantum theory of measurement. We need to only remark in this connection that the measurement of the observables $g(A)$, $A + B$ or AB, is very different from that of A, or of A and B separately. We may also draw attention to the fact that recently it has been shown (Fine and Teller 1978) that (in a Hilbert space of dimension $\geqslant 3$) the condition (a) (Kochen-Specker condition) is incompatible with merely the requirement (I), and that the conditions (b), (c) are incompatible with merely the requirements (I), (III)—so that, there is no need even to invoke the requirement (II) (that a hidden variable theory should reproduce the quantum—theoretic probabilities) in order to rule out the additional conditions such as (a)–(c).

The rest of this paper is devoted to an investigation of whether a hidden variable theory is compatible with certain other requirements which, unlike the conditions (a)–(c), are more basic and are based upon some of the fundamental notions of relativity and quantum theory. In § 2, we first extend the framework of hidden variable theories described above, by introducing a requirement (IV)—that the theory also specify the outcomes of a sequence of experiments performed on a system. This leads us to a discussion of the important result of Bell (1964) that every hidden variable theory (which satisfies the requirements (I)–(IV)), has to be necessarily non-local—*i.e.* it violates the condition of ' local causality '. In § 3, we shall show that hidden variable theories (which satisfy the requirements (I), (II)), are also incompatible with the conventional interpretation of mixed states and the mixing operation in quantum theory. The implications of this result are discussed in § 4, wherein it is concluded that in order to have a hidden variable formulation of quantum theory, it is necessary to give up not only ' local causality ', but also the conventional interpretation of ' mixed ' ensembles based on the usual assumption of quantum theory that the density operator is sufficient to provide an adequate characterization of an ensemble.

2. Sequential measurements, local causality and Bell's theorem

Apart from specifying the probabilities (1) for the outcome of a single experiment, quantum theory also specifies the joint probabilities for the outcome of a sequence of experiments (see for example, Wigner 1963, Srinivas 1975), by means of the so-called ' collapse postulate ', which specifies the change in the state of a system in a measurement process (depending on the particular outcome that is obtained). Since the usual prescription (due to vonNeumann and Lüders) for the change of state in a measurement is applicable only to observables which have a purely discrete spectrum (see for example, Srinivas 1980), we shall restrict ourselves only to such observables. Let A, B be two such observables, with the corresponding spectral resolutions

$$A = \sum_{\alpha} \alpha P^{A}(\alpha), \tag{12a}$$

$$B = \sum_{\beta} \beta P^{B}(\beta), \tag{12b}$$

where α, β are the eigenvalues and $P^{A}(\alpha)$, $P^{B}(\beta)$ are the projectors onto the associated eigensubspaces. Let us assume that we are in the Heisenberg picture and that the

B-measurement refers to a time later than that of A. If we start with an ensemble of systems in state ρ and if this ensemble is first subjected to a measurement of A then, according to the vonNeumann-Lüders collapse postulate, the post-measurement state of that subensemble of systems which gave the outcome α is given by the density operator

$$\frac{P^A(\alpha)\,\rho\,P^A(\alpha)}{\mathrm{Tr}\,(P^A(\alpha)\,\rho)}.$$

Hence, the conditional probability for obtaining the outcome β in the later B-measurement, given that the outcome α is obtained in the earlier A-measurement (when we start with an ensemble of systems in state ρ) is given by

$$\frac{\mathrm{Tr}\,(P^A(\alpha)\,\rho\,P^A(\alpha)\,P^B(\beta))}{\mathrm{Tr}\,(P^A(\alpha)\,\rho)}.$$

Therefore, the joint probability $\mathrm{Pr}^{\rho}_{A,B}(\alpha, \beta)$ that the outcome pair α, β is obtained when an ensemble of systems in state ρ is subjected to an A-measurement followed by a B-measurement, is given by

$$\mathrm{Pr}^{\rho}_{A,B}(\alpha, \beta) = \mathrm{Tr}\,(\rho\,P^A(\alpha)\,P^B(\beta)\,P^A(\alpha)). \tag{13}$$

If a hidden variable theory is to predict the outcome also of a sequence of experiments performed on a system, then we need to add the following requirement (IV) to the ones already considered:*

(IV) For each ordered pair (A, B) of observables (with a purely discrete spectrum) there is associated a (R^2-valued) random variable

$$X_{A,B}\colon \Gamma \to \sigma(A) \times \sigma(B),$$

such that,

$$\begin{aligned}\mathrm{Pr}^{\rho}_{A,B}(\alpha, \beta) &= \mathrm{Tr}\,[\rho\,P^A(\alpha)\,P^B(\beta)\,P^A(\alpha)] \\ &= \mu^{\rho}\,[X^{-1}_{A,B}(\alpha, \beta)]\end{aligned} \tag{14}$$

for each $(\alpha, \beta) \in \sigma(A) \times \sigma(B)$, and all quantum states ρ.

We should emphasize that requirement (IV) is consistent with (I)–(III), as we can easily extend the model constructed in § 1, so as to satisfy (IV) also. If we now write for each $\lambda \in \Gamma$

$$X_{A,B}(\lambda) = (X^{(1)}_{A,B}(\lambda),\, X^{(2)}_{A,B}(\lambda)), \tag{15}$$

then clearly $X^{(1)}_{A,B}(\lambda)$ is the outcome of the A-measurement and $X^{(2)}_{A,B}(\lambda)$ is the outcome of the B-measurement, when an individual system in the completely- specified

*For the sake of simplicity, we shall restrict ourselves to sequences composed of two measurements only.

state $\lambda \in \Gamma$ is subjected to a sequence of two experiments to measure A and B, in that order. Now, the obvious requirement of causality that a performance or non-per formance of a later B-measurement should not in any way affect the outcome of the (earlier) A-measurement, will entail that

$$X^{(1)}_{A,B}(\lambda) = X_A(\lambda). \tag{16}$$

Before going into further discussion of the above causality condition, we shall first state the following result, which is of course well-known in the literature (see for example Bub 1976; Fine 1976; Lochak 1976, etc.), though not in the following form:

Theorem 2.1: Requirements (I)–(IV) are incompatible with the following condition:

$$X_{A,B}(\lambda) = (X_A(\lambda), X_B(\lambda)) \tag{17}$$

for all pairs (A, B) of observables with a purely discrete spectrum.

Proof: From (IV) and (17) it follows that

$$\mathrm{Pr}^{\rho}_{A,B}(\alpha, \beta) = \mu^{\rho}[X^{-1}_{A,B}(\alpha, \beta)] = \mu^{\rho}[X^{-1}_{A}(\alpha) \cap X^{-1}_{B}(\beta)] \tag{18}$$

From (18) and (I)-(III), it is easy to see that

$$\sum_{\alpha} \mathrm{Pr}^{\rho}_{A,B}(\alpha, \beta) = \mu^{\rho}(X^{-1}_{B}(\beta)) = \mathrm{Pr}^{\rho}_{B}(\beta) \tag{19}$$

But it is well-known that the quantum-theoretic joint probabilities (13) violate (19) when A, B are incompatible—a phenomenon which is sometimes referred to as the 'quantum interference of probabilities' (de Broglie 1948; Lochak 1976, Srinivas (1975, 1978, 1982)). Hence we have a contradiction between IV and (19), thereby etablishing the above theorem.

Theorem 2.1 essentially states that the quantum-theoretical joint probabilities are not expressible in the 'classical' form (18), whenever A, B are incompatible. This is of course quite understandable as the earlier A-measurement can be expected to alter the completely-specified state λ of the system. Therefore, while it is necessary to demand (16) for the purposes of causality, it is definitely not reasonable to suppose that (17) be satisfied; for, the latter assumption also implies that

$$X^{(2)}_{A,B}(\lambda) = X_B(\lambda), \tag{20}$$

which would definitely not be the case if the (completely-specified) state λ of the system has changed in the process of the earlier A-measurement. In fact, a physically reasonable model would specify that in an A-measurement the (completely-specified) state of a system changes from λ to $\epsilon_A(\lambda)$, where

$$\epsilon_A : \Gamma \to \Gamma, \tag{21}$$

is a measurable map, which (like the collapse postulate) characterises the way the (completely-specified) states change in an A-measurement. The difference between the collapse postulate of quantum theory and the above transformation (21) lies in the fact that there is nothing ' stochastic ' about the way $\lambda \in \Gamma$ transforms, as the outcome of the A-measurement (*i.e.* $X_A(\lambda)$) is completely specified once $\lambda \in \Gamma$ is known. We can therefore replace (17) by the following more reasonable condition:

$$X_{A,B}(\lambda) = (X_A(\lambda), X_B(\epsilon_A(\lambda))). \tag{22}$$

So far as we are aware, there has not been any general investigation as to whether the condition (22) is in general compatible with the requirements (I) – (IV). We should of course note that the model constructed by Clauser (1971) (as an extension of the hidden variable model constructed by Bell (1964, 1966)) incorporates something like the transformation (21) for the case of a spin- 1/2 system.

Perhaps the only result known so far, which shows that the requirements (I)–(IV) come into conflict with a certain important and physically-required property, is the famous theorem of Bell (Bell 1964, 1971a, 1975, 1980; see also Wigner 1970), that a hidden variable theory necessarily violates a certain (very reasonable) requirement of ' locality ' or ' local causality '. To explain this condition, let us consider a situation where the observable A, B refer to local measurements performed in two space-like separated space-time regions. Then, from the usual causality requirement of relativistic quantum theory, that the statistics of the outcomes of the A-measurement should not depend on whether or not a B-measurement is performed, it follows that A and B commute and (13) reduces to

$$\mathrm{Pr}^{\rho}_{A,B}(\alpha, \beta) = \mathrm{Tr}\,[\rho\, P^A(\alpha)\, P^B(\beta)]. \tag{23}$$

The important point is that in addition, in a hidden variable theory, the particular outcomes $X_A(\lambda)$ or $X_B(\lambda)$ of the A-measurement or B-measurement, when performed alone, should also be the same as those obtained in a situation when both A, B are measured. In other words, we should have the following property (of ' local causality '):

$$X_{A,B}(\lambda) = (X_A(\lambda), X_B(\lambda)), \tag{24}$$

whenever A, B refer to space-like measurements*. The theorem of Bell is the following:

Theorem 2.2 (*Bell's Theorem*): Requirements I–IV are incompatible with the following condition of ' local causality ':

$$X_{A,B}(\lambda) = (X_A(\lambda), X_B(\lambda))$$

for all compatible observables A, B which refer to space-like separated measurements.

*We can also arrive at (24) from the causality condition (16), by noting that when A, B refer to space-like separated measurements, either of them can be considered the prior measurement.

The proof of Bell's theorem in the above form (Wigner 1970; Bub 1974; Fine 1976, 1982) is based on considering a system of two spin-1/2 particles in singlet spin state ψ, with the particles moving away in opposite direction. When the particles are sufficiently far apart, measurements of their spin components (such as $\sigma_1 . \hat{a}$, $\sigma_2 . \hat{b}$) could be taken as space-like separated measurements. It can then be shown that the requirements (I) – (IV), together with the condition (24) lead to the following inequality due to Wigner (1970):

$$\mathrm{Pr}^{\psi}_{\sigma_1 . \hat{a}, \sigma_2 . \hat{c}} (1, 1) \leqslant \mathrm{Pr}^{\psi}_{\sigma_1 . \hat{b}\, \sigma_2 . \hat{c}} (1, 1) + \mathrm{Pr}^{\psi}_{\sigma_1 . \hat{a}, \sigma_2 . \hat{b}} (1, 1) \qquad (25)$$

where σ_1, σ_2 are the spin operators of the two particles and $\hat{a}$, $\hat{b}$, $\hat{c}$, are arbitrary directions. The fact that the quantum-theoretic joint probabilities do not obey (25) leads to the conclusion that a hidden variable theory necessarily violates the local causality condition* (24).

3. The incompatibility of hidden variable theories with quantum theory

In this section we shall show that apart from violating local causality, a hidden variable theory is also necessarily incompatible with the conventional interpretation of the mixed states and the mixing operation in quantum theory. We may note that in order to prove this we need to invoke only the requirements (I) and (II), and hence the result is independent of the spectrum rule (III) and also of any assumptions (such as (IV)) regarding the joint probabilities for sequential or simultaneous measurements.

The conventional interpretation of the mixed states and the mixing operation is the following: Suppose we have an ensemble of N systems prepared in quantum state ρ_1, and another ensemble of N systems prepared in quantum state ρ_2. By selecting a fraction ν ($0 \leqslant \nu \leqslant 1$) of systems from the first ensemble and a fraction $(1 - \nu)$ from the second ensemble, we can form a new ensemble of N systems. Then, for N large enough, this new ensemble can be considered as an ensemble of N systems in quantum state ρ given by

$$\rho = \nu\, \rho_1 + (1 - \nu)\, \rho_2. \qquad (26)$$

It is important to note that in the above statement there is no restriction that ρ_1, ρ_2 be sure states, or that they be ' orthogonal ' in some sense or the other. A well-known consequence of the above statement is the fact that there are several different ways of preparing an ensemble of systems in a mixed (quantum) state ρ—for, any mixed state can be expressed as a convex combination of other states in several different ways, say as in the following equation

$$\rho = \nu\rho_1 + (1 - \nu)\, \rho_2 = \nu'\, \rho' + (1 - \nu')\, \rho'_2, \qquad (27)$$

where ρ_1, ρ_2, ρ'_1, and ρ'_2 are all different. One of the main features of the conventional formulation of quantum theory is that all these differently prepared ensembles

*Bell (1975, 1980) has outlined an argument which seems to show that quantum theory itself (i.e. without any reference to a hidden variable formulation) is incompatible with 'local casuality'.

are to be considered *identical*, as long as they are characterised by the *same* density operator ρ.

We should emphasise that it is the above interpretation of mixed states and the mixing operation which leads to the usual requirement that the quantum theoretic probabilities (for single or sequential measurements) have to be affine functions of the density operator (see eq. (28) below). And, it is this feature of the quantum-theoretic probabilities which leads to the fact that unlike in classical probability theory, there is an 'interference of probabilities' in any sequential measurement of two incompatible observables. This point can be seen very clearly from the following result essentially due to Wigner (Wigner 1932, 1971)—see also Srinivas and Wolf (1975) and Mugur-Schächter (1977)) and Davies (1976):

Theorem 3.1 (*Wigner-Davies Theorem*): Let A, B be two self-adjoint operators and let us suppose that, for each density operator ρ, there is associated a joint probability measure (on R^2),

$$\Delta_1 \times \Delta_2 \to R^{\rho}_{A,B}(\Delta_1, \Delta_2) \in [0, 1]$$

such that the following conditions (i)—(iii) are satisfied:

$$\text{(i)} \qquad R^{\nu\rho_1+(1-\nu)\rho_2}_{A,B}(\Delta_1, \Delta_2) = \nu\, R^{\rho_1}_{A,B}(\Delta_1, \Delta_2) + (1-\nu)\, R^{\rho_2}_{A,B}(\Delta_1, \Delta_2) \tag{28}$$

for all ν $(0 \leqslant \nu \leqslant 1)$ and for all density operators ρ_1, ρ_2;

$$\text{(ii)} \qquad R^{\rho}_{A,B}(\Delta, R) = \mathrm{Pr}^{\rho}_{A}(\Delta) = \mathrm{Tr}\,(\rho\, P^{A}(\Delta)) \tag{29}$$

for each density operator ρ;

$$\text{(iii)} \qquad R^{\rho}_{A,B}(R, \Delta) = \mathrm{Pr}^{\rho}_{B}(\Delta) = \mathrm{Tr}\,(\rho\, P^{B}(\Delta)) \tag{30}$$

for each density operator ρ.

Then A and B are compatible (*i.e.* their spectral projectors commute) and

$$R^{\rho}_{A,B}(\Delta_1, \Delta_2) = \mathrm{Tr}\,[\rho\, P^{A}(\Delta_1)\, P^{B}(\Delta_2)] \tag{31}$$

The above theorem is a direct consequence of the theorem 2·1 of Chapter 3 of Davies (1976), once we recognise that conditions (i)-(iii) imply that there exists a positive-operator-valued measure (on R^2)

$$\Delta_1 \times \Delta_2 \to M(\Delta_1, \Delta_2),$$

such that

$$M(\Delta, R) = P^{A}(\Delta), \tag{32a}$$

and

$$M(R, \Delta) = P^{B}(\Delta). \tag{32b}$$

As Davies has shown, (31a) and (31b) are possible only when A and B are compatible, and then

$$M(\Delta_1, \Delta_2) = P^A(\Delta_1)\, P^B(\Delta_2), \tag{33}$$

from which the above theorem follows. We may also add here that the quantum-theoretic joint probabilities as given by (13) satisfy only the conditions (i) and (ii), but not (iii) whenever A, B are incompatible and the B-measurement is assumed to follow the A-measurement.

We shall now employ the Wigner-Davies theorem to prove the following theorem which constitutes the main result of this paper:

Theorem 3.2: Requirements (I), (II) of a hidden variable theory are incompatible with the following condition:

$$\mu^{\nu\rho_1+(1-\nu)\rho_2} = \nu\, \mu^{\rho_1} + (1-\nu)\, \mu^{\rho_2}, \tag{34}$$

for all ν $(0 \leqslant \nu \leqslant 1)$ and for all density operators ρ_1, ρ_2. As a consequence, a hidden variable theory which satisfies the requirements (I), (II) is necessarily incompatible with the conventional interpretation of mixed states and the mixing operation in quantum theory.

Proof: Let us consider two incompatible observables A, B and, for each density operator ρ, form the real valued set function (on R^2)

$$\Delta_1 \times \Delta_2 \to R^{\rho}_{A,B}(\Delta_1, \Delta_2),$$

given by

$$R^{\rho}_{A,B}(\Delta_1, \Delta_2) = \mu^{\rho}\,[X_A^{-1}(\Delta_1) \cap X_B^{-1}(\Delta_2)] \tag{35}$$

It follows from (I) that $R^{\rho}_{A,B}(\Delta_1, \Delta_2)$ given by (35) defines a joint probability measure on R^2. It is now quite straightforward to see that (34) and the requirement (II) together imply that $R^{\rho}_{A,B}(\Delta_1, \Delta_2)$ satisfies the conditions (i)–(iii) of the Wigner-Davies theorem. Hence it follows that A, B are compatible, which contradicts the assumption we made to start with, thereby establishing that the condition (34) is incompatible with the requirements (I), (II) of a hidden variable theory.

To prove the second assertion of the above theorem, we shall show that the condition (34) is actually a consequence of the requirement (I) and the conventional interpretation of mixed states and the mixing operation in quantum theory. For this purpose, let us consider two ensembles Σ_1, Σ_2 each of N systems prepared in (quantum) states ρ_1, ρ_2 respectively. When N is large enough, the individual systems of these ensembles are distributed over the space Γ as per the probability measures μ^{ρ_1} and μ^{ρ_2} respectively. Now, if we select a fraction ν $(0 \leqslant \nu \leqslant 1)$ of the systems of ensemble Σ_1, and a fraction $(1-\nu)$ of the systems of the ensemble Σ_2, then we can form a new ensemble of N systems Σ. It follows from (I) that the individual systems

of Σ (when N is large enough) are distributed over the space Γ as per the probability measure $\nu\,\mu^{\rho_1} + (1-\nu)\,\mu^{\rho_2}$. To show this, let us consider an arbitrary measurable subset $\Omega \subset \Gamma$. It follows from (I) that out of the $\nu\,N$ systems selected from Σ_1, the number of systems which lie in Ω is $\nu N\,\mu^{\rho_1}(\Omega)$, (if N, and hence, $\nu\,N$ is large enough), and similarly out of the $(1-\nu)\,N$ systems selected from Σ_2, the number of systems which lie in Ω is $(1-\nu)\,N\,\mu^{\rho_2}(\Omega)$, (if N, and hence, $(1-\nu)\,N$ is large enough). Hence, out of the N systems of the new (mixed) ensemble Σ, $N\{\nu\,\mu^{\rho_1}(\Omega)+(1-\nu)\,\mu^{\rho_2}(\Omega)\}$ systems lie in Ω. Since Ω was chosen to be an arbitrary measurable subset of Γ, it follows that the individual systems of the ensemble Σ are distributed over Γ as per the probability measure $\nu\,\mu^{\rho_1} + (1-\nu)\,\mu^{\rho_2}$. Having shown this, the crucial point now is to realise that it follows from the conventional interpretation of mixed states and the mixing operation in quantum theory that Σ can be considered as an ensemble of N systems in the (quantum) state $\rho = \nu\,\rho_1 + (1-\nu)\,\rho_2$. It then follows from (I) that the condition (34) *i.e.*

$$\mu^{\nu\,\rho_1+(1-\nu)\,\rho_2} = \nu\,\mu^{\rho_1} + (1-\nu)\,\mu^{\rho_2}$$

should be satisfied for all ν $(0 \leqslant \nu \leqslant 1)$ and all density operators ρ_1, ρ_2. Since we have already shown that the above condition is incompatible with (I) and (II), we can therefore conclude that any hidden variable theory which satisfies (I) and (II) is incompatible with the conventional interpretation of mixed states and the mixing operation in quantum theory. The same result can also be stated as follows: If the conventional interpretation of mixed states and the mixing operation is assumed, then every hidden variable theory satisfying (I), is necessarily incompatible with (II) —*i.e.* the theory will *not* reproduce the quantum mechanical probabilities.

4. Discussion

One important feature of a hidden variable theory, as defined by conditions (I)–(II), is that it specifies the actual outcome (say $X_A(\lambda)$) of any experment (say, to measure A) made on an individual system, provided we know the completely-specified state, $\lambda \in \Gamma$, of the system. It is therefore possible in such a theory to view the outcome of any experiment to be merely a property of the system (and the particular experimental arrangement) which is merely revealed in the measurement process. In other words, as Bell (1975) would say, the conventional notion of 'observables' can now be replaced by that of ' beables '. Also, as Bell (1971a, b) has emphasised, one of the important motivations for a hidden variable theory is that it seems to get rid of the entire problem of defining 'what constitutes a measurement?' and of understanding ' what happens in a measurement? '—both of which appear to have no definitive answers, so far, from within quantum theory itself.

Now, the question is whether quantum theory is consistent with a viewpoint where the outcomes of experiments are viewed as being actually the properties of the system (and the experimental arrangements) which are merely revealed in each measurement process. In this connection we should of course take into account the fact that the quantum state of a system changes in a measurement process. This however does not come into conflict with the above viewpoint, provided we accept that a measurement

process, while revealing an actual property of the system (and the apparatus), also alters the state of the system as in say, (21). A more important problem arises due to the fact that in quantum theory we have incompatible observables (like say, position and momentum) which cannot be measured simultaneously on the same system. The question now is: can we ' theoretically ' specify the values of all observables (including the ones which are mutually incompatible) for each individual system (even though they are not measurable), and still have no conflict with quantum theory? For example, can we ' theoretically ' specify simultaneous values for position and momentum, or more generally a ' path ' for a particle, and still have no conflict with quantum theory?

Both the theorem of Bell (theorem 2.2) and the theorem 3.2, imply that the answer to the above questions would be in the negative, (and that too for very different reasons), unless some basic changes are made in some of the fundamental notions of both relativity and quantum theory. Bell's theorem shows that even when one considers two experiments (say, to measure two compatible observables A and B) which are conducted in space-like separated space-time regions, the outcome (say, $X_A(\lambda)$ and $X_B(\lambda)$) assigned to each of them when it is the case that only *one* of the two experiments is performed, will not be the same as the outcomes (say $X^{(1)}_{A,B}(\lambda)$ and $X^{(2)}_{A,B}(\lambda)$) that should be assigned to them when *both* the experiments are performed together. Hence in any hidden variable theory, it is necessary that while assigning an outcome to any experiment, we have also to specify what other experiments are being performed even at space-like separated space-time regions, so that the theory becomes clearly non-local. The implications of this feature of hidden variable theories are being widely discussed in the literature.

Theorem 3.2 shows that assigning values to incompatible observables come into conflict with quantum theory itself, and in particular with the conventional interpretation of the mixed states and the mixing operation. To see this clearly, let us repeat the chain of arguments which led to the theorem 3.2, for the case of an ensemble of spin-$\frac{1}{2}$ systems in some quantum state ρ. We only need to consider two incompatible measurements say, of the spin-components $\boldsymbol{\sigma}.\hat{a}$ and $\boldsymbol{\sigma}.\hat{b}$ along two different directions (*i.e.* $\mid \hat{a}.\hat{b} \mid < 1$). If, to each member of the ensemble, we assign particular values for both the spin-components $\boldsymbol{\sigma}.\hat{a}$ and $\boldsymbol{\sigma}.\hat{b}$, then we can use these value assignments to construct a joint probability distribution $R^{\rho}_{\boldsymbol{\sigma}.\hat{a},\,\boldsymbol{\sigma}.\hat{b}}(\alpha,\beta)\,(\alpha,\beta=\pm 1)$, which satisfies (29) and (30). Now, the crucial point is that the conventional interpretation of the mixing operation implies that $R^{\rho}_{\boldsymbol{\sigma}.\hat{a},\,\boldsymbol{\sigma}.\hat{b}}(\alpha,\beta)$ is an affine function of the density operator ρ. Hence it follows as a consequence of the Wigner-Davies theorem (theroem 3.1) that $\boldsymbol{\sigma}\cdot\hat{a}$ and $\boldsymbol{\sigma}\cdot\hat{b}$ are compatible, which contradicts the supposition that we made to start with that $\boldsymbol{\sigma}\cdot\hat{a}$ and $\boldsymbol{\sigma}\cdot\hat{b}$ are incompatible.

Finally, the above analysis also indicates how the conventional formulation of quantum theory could be *modified* so that a hidden variable theory would not get ruled out. The main point would be to modify the conventional interpretation of the mixed states and the mixing operation, at least to an extent that it would affect only those predictions of a hidden variable theory which are considered to be outside the purview of quantum theory, in the sense that the latter is supposed to deal only

with the statistics of measurement results. This could be done as follows. Suppose we have distinct pure state density operators P_{ψ_1}, P_{ψ_2}, $P_{\psi_1'}$, $P_{\psi_2'}$, such that

$$\rho = \nu P_{\psi_1} + (1 - \nu) P_{\psi_2} = \nu' P_{\psi_1'} + (1 - \nu') P_{\psi_2'} \tag{36}$$

where $0 < \nu, \nu' < 1$. Then the conventional interpretation of mixing operation would imply that the ensemble which is obtained by mixing two ensembles of systems in pure states ψ_1, ψ_2 in the ratio $\nu : 1 - \nu$, should be deemed *identical* with the ensemble obtained by mixing two ensembles of systems in pure states ψ_1', ψ_2' in the ratio $\nu' : 1 - \nu'$, because both the mixed ensembles are characterised by the *same* density operator ρ. For the same reason, even while defining a hidden variable theory, we assumed (in condition (I)) that both the ensembles are distributed over the space Γ according to the *same* probability measure μ^ρ. We shall now suppose that, at least as far as hidden variable theories are concerned, the two ensembles are *not* identical, even though they are characterised by the *same* density operator ρ*. Firstly, we shall replace the condition (I) by the following:

(I') With each pure state ψ (a unit ray in Hilbert space), is associated a probability measure) μ^ψ on (Γ, B) ....(rest as in (I)).

The main point is that in (I') we no longer assume that for each density operator ρ there is a *unique* probability measure μ^ρ on Γ. In fact, if we now employ the usual interpretation that, given an ensemble of systems in a pure quantum state ψ, the various individual systems of the ensemble are distributed over the various points of Γ as per the measure μ^ψ**, then we shall see that there are several distinct ways in which an ensemble of systems, characterised by a (mixed state) density operator ρ, could be distributed over the space Γ. For example, if we consider the ensemble prepared by mixing two ensembles of systems in pure states ψ_1, ψ_2 in the ratio $\nu : 1 - \nu$, then the individual systems of this ensemble are distributed over Γ as per the probability measure $\nu\mu^{\psi_1} + (1 - \nu)\mu^{\psi_2}$. On the other hand, if we consider the ensemble of systems which is obtained by mixing two ensembles of systems in the pure states ψ_1', ψ_2' in the ratio $\nu' : 1 - \nu'$, then the individual systems of this ensemble are distributed over Γ as per the probability measure $\nu'\mu^{\psi_1'} + (1 - \nu')\mu\,\gamma^{\psi_2'}$. Theorem 3.2, now merely implies that these two probability measures are in general different, even though the two ensembles were characterised by the same density operator ρ. In conclusion we may thus state that apart from necessarily violating ' locality ' or ' local causality ', a hidden variable theory would also necessarily violate the usual assumption of quantum theory that the density operator provides an adequate characterization of any ensemble of systems, pure or mixed. As we have shown, a hidden variable theory would necessarily treat as distinct, two (mixed) ensembles characterised by the *same* density operator, but prepared *differently* by mixing ensembles of systems in different sets of pure states.

*The conventional view that the density operator provides an adequate characterization of any ensemble of systems has been questioned on various other grounds earlier in the literature—see for example, Benioff and Eckstein (1977), Ghirardi *et al* (1975, 1976) Haag and Bannier (1978), Mielnik (1974) and also Newton (1976).

**Notice that in (I') we continue to assume that *any ensemble* of systems in a *pure* state is completely characterised by the quantum mechanical state vector ψ.

Acknowledgements

The author is grateful to Professor J. S. Bell for several discussions and useful comments on this work, and to Professor J. Pasupathy for hospitability at the Centre for Theoretical Studies, Indian Institute of Science, Bangalore, which made these discussions possible.

References

Belinfante F J 1973 *A survey of hidden variable theories* (Oxford: Pergamon)
Bell J S 1964 *Physics* **1** 195
Bell J S 1966 *Rev. Mod. Phys.* **38** 447
Bell J S 1971a in *Foundations of quantum mechanics*, (ed.) B d' Espagnat (New York: Academic Press)
Bell J S 1971b On the hypothesis that the Schrodinger equation is exact (CERN Preprint Ref. TH 1424)
Bell J S 1975 The theory of local beables (CERN Preprint Ref. TH 2053)
Bell J S 1980 Bertlmann's socks and the nature of reality (CERN Preprint Ref. TH 2926)
Benioff P A and Eckstein H 1977 *Nuovo Cimento* **B40** 9
Bohm D and Bub J 1966 *Rev. Mod. Phys.* **38** 470
Bub J 1974 *The interpretation of quantum mechanics* (Dordrecht: Reidel)
Bub J 1976 *Found. Phys.* **6** 511
Capasso V, Fortunato D and Selleri F 1970 *Rev. del Nuovo Cimento* **2** 149
Clauser J F 1971 *Am. J. Phys.* **39** 1095
Clauser J F and Shimomy A 1978 *Rep. Prog. Phys.* **41** 1881
Davies E B 1976 *Quantum theory of open systems* (New York: Academic Press)
de Broglie L 1948, La Revue Scientifique No. 3292 fasc. 5, **87** 259
d'Espagnat B 1976 *Conceptual foundations of quantum mechanics* 2nd Ed. (Reading: Benjamin)
d'Espagnat B 1979 *Sci. Am.* **241** 128
Fine A 1976 in *Logic and probability in quantum mechanics* (ed.)P Suppes (Dordrecht: Reidel)
Fine A 1982 *Phys. Rev. Lett.* **48** 291
Fine A and Teller P 1978 *Found. Phys.* **8** 629
Ghirardi G C, Rimini A and Weber T 1975 *Nuovo Cimento* **B29** 135
Ghirardi G C, Rimini A and Weber T 1976 *Nuovo Cimento* **B33** 457
Haag R and Bannier U 1978 *Commun. Math. Phys.* **60** 1
Jammer M 1974 *The philosophy of quantum mechanics* (New York: John Wiley)
Kochen S and Specker E P 1967 *J. Math. Mech.* **17** 59
Lochak G 1975 *Found. Phys.* **6** 173
Mielnik B 1974 *Commun. Math. Phys.* **37** 221
Mugur-Schächter M 1977 in *Quantum mechanics a half century later* (eds.) J Leite-Lopes and M Paty (Dordrecht: Reidel)
Newton R G 1976 *Nuovo Cimento* **B33** 454
Roy S M 1980 *Phys. News* **11**
Selleri F and Tarozzi G 1981 *Rev. del Nuovo Cimento* **4** 1
Srinivas M D 1975 *J. Math. Phys.* **16** 1672
Srinivas M D 1978 *J. Math. Phys.* **19** 1705
Srinivas M D 1980 *Commun. Math. Phys.* **71** 131
Srinivas M D 1982, *The Wave-particle Dualism* (eds.) S Diner *et al* (Dordrecht: Reidel)
Srinivas M D and Wolf E 1975 *Phys. Rev.* **D11** 1477
Vandana Shiva 1978 *Hidden variables and locality in quantum mechanics* Ph.D. Thesis (unpublished) (University of Western Ontario)
Virendra Singh 1980 in *Gravitation, quanta and the universe*, (eds) A R Prasanna, J V Narlikar and C V Vishveshvara (New Delhi: Wiley Eastern)

Wigner E P 1932 *Phys. Rev.* **40** 749
Wigner E P 1963 *Am. J. Phys.* **31** 6
Wigner E P 1970 *Am. J. Phys.* **38** 1005
Wigner E P 1971 in *Perspectives in quantum theory* (eds) W Yourgrau and T van der Merwe (Cambridge: MIT Press)