Foundations of Physics, Vol. 13, No. 5, 1983

Nonlocality and the Kochen–Specker Paradox

Peter Heywood[1] **and Michael L. G. Redhead**[1]

Received January 27, 1982

A new proof of the impossibility of reconciling realism and locality in quantum mechanics is given. Unlike proofs based on Bell's inequality, the present work makes minimal and transparent use of probability theory and proceeds by demonstrating a Kochen–Specker type of paradox based on the value assignments to the spin components of two spatially separated spin-1 systems in the singlet state of their total spin. An essential part of the argument is to distinguish carefully two commonly confused types of contextuality; we call them ontological and environmental contextuality. These in turn are associated with two quite distinct senses of nonlocality. We indicate the relevance of our treatment to other related discussions in recent literature on the philosophy of quantum mechanics.

1. INTRODUCTION

Since the inception of quantum mechanics (QM), realism has been a prominent issue in its interpretation. Of late, particularly after the work of Kochen and Specker[(1)] and Bell,[(2)] two serious allegations have been brought against realism in quantum mechanics. First, that a realistic interpretation involves an algebraic contradiction—this charge came from the work of Kochen and Specker. Second, that realism entails nonlocality—this came from Bell's work. In this paper, by showing that any local realism leads to a Kochen–Specker type of contradiction, we shall show how the results of Kochen–Specker and of Bell are linked.[2]

[1] Department of History and Philosophy of Science, Chelsea College, University of London, Manresa Road, London England.

[2] The connection between nonlocality and the Kochen–Specker paradox appears to have been suggested by Simon Kochen in an unpublished comment to A. Shimony. (We are grateful to Professor Shimony for this information.) The idea was developed in a rather different context by Stairs.[(3)] Stairs was attempting to criticize Fine's defense of (effectively) what we call CVR below. By assuming locality, Stairs showed that CVR led to a Kochen–Specker

0015-9018/83/0500-0481$03.00/0

Let us explain some simple notation.[3] The sentence $[Q]^{\phi} = \lambda$, where Q is any physical magnitude,[4] ϕ is any quantum state, and λ any point on the real line, shall mean that in state ϕ the physical magnitude Q has the value λ.[5] It should be noted that these are assertions about values *possessed* by physical magnitudes; they say nothing of the values that will be found upon measurement. Notice also that in a realistic interpretation of QM, $[Q]^{\phi}$ is well-defined whether or not ϕ is an eigenstate of the corresponding operator $\hat{Q}$. We use the term "physical magnitude" rather than "observable" because we wish to distinguish sharply between physical magnitudes and self-adjoint operators. It is perhaps easier to slide from observables to operators than from physical magnitudes to operators. "Physical magnitude" also benefits from the lack of any anti-realist connotations that might attach to the "observable." At present we are following the common assumption that to every self-adjoint operator there corresponds a unique physical magnitude—we will revise this shortly. The sentence $P_Q^{\phi}(\lambda) = r$ for any quantum state ϕ and physical magnitude Q, and for any point r in $[0, 1]$, shall mean that the quantum mechanical probability that a system in state ϕ on measurement of Q will be found to have the value λ, is r.

We shall be concerned with the sort of realism which asserts at least that at all times and in all states every physical magnitude which pertains to the system has some value.[6] However, we must have a more detailed specification of the realist thesis before we can assess its merits. There is in the literature a profusion of "rules"—essentially constraints restricting the possible value assignments—each of which, it is supposed, must be accepted by any realist. We shall use several and we beg the readers forgiveness for the introduction of still others.

The following rule has been widely accepted as central to any realist interpretation of *QM* and will be important in our proofs:

Value Rule (VR): $P_Q^{\phi}(\lambda) = 0 \Rightarrow [Q]^{\phi} \neq \lambda$

---

contradiction and concluded that CVR must be false. We show below that CVR is derivable from the generally accepted VR and FUNC* and hence deduce a different conclusion from Stair's work, namely a demonstration of nonlocality. We acknowledge our debt to Stairs, although the details of our arguments are quite different.

[3] Our notation is simple initially, it gets a little more complicated as we adapt it to cope with problems as they arise.

[4] A physical magnitude is a property of an object or a physical system, the associated self-adjoint operator $\hat{Q}$ is of course a mathematical entity; let's not confuse the two.

[5] This is our adaptation of a notation due to Fine.[(4)] We confine our discussion to physical magnitudes whose associated self-adjoint operators have a discrete spectrum.

[6] This, of course, excludes what might be called "dispositional realism," according to which systems have dispositions (propensities) to manifest properties under certain conditions, though they need not possess these properties at all times.

Essentially this rule constrains a physical magnitude's value assignments to its eigenvalue spectrum and for any state ϕ to those eigenvalues which have a nonzero probability of being found on measurement.[7]

VR is usually justified by appeal to another rule:

Faithful Measurement (FM): Any measurement of a physical magnitude Q reveals the value which Q had immediately prior to the measurement.

This rule allows us to move freely between measured and possessed values. We cannot actually use FM to *prove* VR since we must in addition assume the identity of probability distributions for measured and possessed values.[8] But it would demonstrate a remarkable conspiracy on Nature's part if possessed values were always revealed on measurement, but repeated measurement did not uncover *all* those values which were possessed. If we reject such a conspiracy then FM may indeed be regarded as justifying VR.

Another rule much discussed since the work of Kochen and Specker is Functional Composition, commonly known as FUNC.

FUNC: If $\hat{A}$ and $\hat{B}$ are two operators and there exists a function f such that $\hat{B} = f(\hat{A})$, then

$$[B]^{\phi} = f([A]^{\phi})$$

for any state ϕ, where A and B are the (unique) physical magnitudes corresponding to $\hat{A}$ and $\hat{B}$, respectively. We shall sometimes express this result in the form

$$[f(A)]^{\phi} = f([A]^{\phi})$$

where the symbol $f(A)$ derives its significance from the fact that it is the physical magnitude whose associated self-adjoint operator is $f(\hat{A})$. Essentially the idea of FUNC is that the algebraic structure of the operators should be mirrored in the algebraic structure of the possessed values of the physical magnitudes. Kochen and Specker[(1)] showed that if FUNC constrains the value assignments a contradiction will result provided that the Hilbert space of state vectors has a dimension greater than two. Their important result is purely algebraic.[9]

[7] VR only makes sense for physical magnitudes associated with a discrete spectrum. For physical magnitudes associated with a continuous spectrum it would imply that no precise value was possessed by the physical magnitude in question. Compare Teller[(6)] for further discussion of this point.

[8] See Ref. 5.

[9] Kochen and Specker's result is in fact a corollary of Gleason's[(7)] theorem, and is also implicit in the work of Bell.[(8)] For a comprehensive discussion see Fine and Teller.[(9)] The question of submitting FUNC to experimental test is examined in Redhead.[(10)]

Kochen and Specker believed that they had shown the impossibility of any realist interpretation of QM (except for the special case of spin-$\frac{1}{2}$ systems). Naturally there has been much discussion as to whether a commitment to FUNC is necessary for realism. It is undisputed however that the Kochen and Specker result relies on FUNC holding in the following situation:

$$\hat{A} = f(\hat{B}), \qquad \hat{A} = g(\hat{C}), \qquad [\hat{B}, \hat{C}] \neq 0$$

for some functions f and g. In this situation FUNC constrains the values of incompatible magnitudes:

$$[A]^{\phi} = f([B]^{\phi})$$

and

$$[A]^{\phi} = g([C]^{\phi})$$

Hence

$$f([B]^{\phi}) = g([C]^{\phi})$$

As is well-known this situation can only arise if f and g are many-one mappings, so in particular if $\hat{B}$ and $\hat{C}$ are maximal[10] (nondegenerate) then $\hat{A}$ must be a degenerate operator having some of its eigenvalues equal.

This reasoning and with it the Kochen–Specker proof can be blocked by restricting FUNC. The restriction is that FUNC applies only within the set of functions of a single maximal operator. Thus we will break the relationship between the algebraic structure of operators and the algebraic structure of physical magnitudes just where Kochen and Specker used it to get their contradiction. Let us make this clearer. So far we have assumed that to each operator there corresponds one and only one physical magnitude, in the future we shall assume this only for maximal operators. So we will still write $[R]^{\phi}$ for the possessed value of a physical magnitude whose corresponding operator $\hat{R}$ is maximal. But corresponding to each nonmaximal operator we shall assume there are many physical magnitudes. Furthermore these many physical magnitudes are distinguished from one another by the functional relationship between their values and the values of physical magnitudes corresponding to maximal operators. Thus, suppose as above that

$$\hat{A} = f(\hat{B}), \qquad \hat{A} = g(\hat{C}), \qquad [\hat{B}, \hat{C}] \neq 0$$

[10] Operators are, of course, maximal or nonmaximal. When we describe physical magnitudes as maximal or otherwise, we mean to impute the property to their associated operators.

where $\hat{B}$ and $\hat{C}$ are maximal so we can unambiguously associate physical magnitudes B and C with them. But A is nonmaximal and with it we associate two physical magnitudes—there may be more—A_B and A_C and these physical magnitudes are identified by their functional relations with B and C in respect to the value assignments—specifically

$$[A_B]^\phi = f([B]^\phi) \qquad \text{and} \qquad [A_C]^\phi = g([C]^\phi)$$

In a sense physical magnitudes corresponding to maximal operators are ontologically prior to those which correspond to nonmaximal operators. Knowing to which self-adjoint operator it corresponds is not sufficient to identify unambiguously a nonmaximal physical magnitude, we must know also to which maximal physical magnitude its values are related. It requires a "context"—at least this is, we believe, one idea that has been in the minds of those who speak of contextuality. Every nonmaximal self-adjoint operator now corresponds not only to one, but to many different physical magnitudes. In fact this is to de-Ockhamize QM *à la* van Fraassen.[11] We shall refer to it as *Ontological Contextuality*. Of course this blocks the Kochen–Specker proof because we have *no* reason to think, as they require, that

$$[A_B]^\phi = [A_C]^\phi$$

when $\hat{B}$ and $\hat{C}$ does not commute. Notice that in order to *measure* A_B for example we proceed by measuring B and applying the function f to the result. Similarly A_C is measured by measuring C and applying the function g to the result, and so on.

We accept then a restricted version of FUNC—call it FUNC*—which follows directly from the *definition* of $[A_B]^\phi$ given above.

FUNC*: Let $\hat{B}$ be a maximal self-adjoint operator and $\hat{A}$ and $\hat{D}$ self-adjoint operators such that for functions h, f, and g we have the relations

$$\hat{A} = f(\hat{B}), \qquad \hat{D} = g(\hat{B}), \qquad \hat{A} = h(\hat{D})$$

then

$$[A_B]^\phi = h([D_B]^\phi)$$

In particular, if $\hat{D}$ is also maximal, so we can identify D_B with the unique physical magnitude D,

$$[A_B]^\phi = [A_D]^\phi$$

So the value of A_B in a given state ϕ does not depend on the maximal physical magnitudes of which it is considered a function in the equivalence class of 1 : 1 functions of a given maximal physical magnitude. Thus if we denote the equivalence class generated by B, by $\{B\}$ our notation can be conveniently modified be *defining* a new symbol

$$[A]^{\phi}_{\{B\}} \underset{Df}{=} [A_B]^{\phi}$$

which serves to stress that $[A_B]^{\phi}$ depends only on the equivalence class $\{B\}$, or, which comes to the same thing, the complete orthonormal basis of eigenvectors for $\hat{B}$ (in the case of the discrete spectrum we are considering for the purposes of this paper).

With regard to our avowal of FUNC* two questions immediately arise. Is it consistent with and is it independent of VR? Both questions can be answered in the affirmative. Let R be any maximal physical magnitude with a unique (up to phase) orthonormal basis of eigenvectors. Order this set in some way $\phi_1, \phi_2 \ldots$. Then let n_0 be the first n such that $|\langle \phi_n \mid \phi \rangle|^2 \neq 0$ and define

$$[f(R)]^{\phi}_{\{R\}} = f(\lambda_{n_0}) \qquad (1)$$

where λ_{n_0} is the eigenvalue corresponding to ϕ_{n_0}. By choice of n_0 this satisfies VR and also agrees with FUNC*. Hence we have established consistency. But now suppose (1) holds only if f is *not* the identity function and define, $[R]^{\phi}_{\{R\}} = \lambda_{n_1}$, where n_1 is the second n such that $|\langle \phi_n \mid \phi \rangle|^2 \neq 0$ and λ_{n_1} is the eigenvalue corresponding to ϕ_{n_1}, then VR still holds but FUNC* is violated. Hence we have established independence.[11] There is also the question of whether FUNC* is consistent with the algebraic structure of *maximal* physical magnitudes. Here we can cite a theorem of Maczynski[(12)] to the effect that it is indeed possible to assign values to the set of maximal physical magnitudes consistently with maintaining functional relationships between them.

With regard to the proliferation of physical magnitudes envisaged by the van Fraassen solution to the Kochen–Specker paradox, an interesting question arises in the case of two spatially separated systems.

Let S_1 and S_2 be two such systems which may or may not have interacted in the past and let their associated Hilbert space H_1 and H_2 be of arbitrary dimension. If we so choose we may describe the combined system $(S_1 + S_2)$ in the product space $H_1 \otimes H_2$. Let $\hat{A}$ be a maximal operator in the space H_1 and let $\hat{B}$ be a maximal operator in the space H_2. The following

[11] The above argument is due to Professor Fine. We are grateful for his advice on the question of consistency and independence of FUNC* and VR.

problem occurs: since $(A \otimes 1)$, the magnitude associated with the operator $(\hat{A} \otimes \hat{1})$, is not a maximal magnitude in the product space, should $(A \otimes 1]^{\phi}_{\{X\}}$ and $[A \otimes 1]^{\phi}_{\{Y\}}$, where X and Y are maximal incompatible magnitudes *on the product space*, be treated as different physial magnitudes? A straightforward application of ontological contextuality would lead us to do so. The intuitive consequences of doing so are peculiar. A sort of holism is involved, physical magnitudes maximal in the product space are somehow prior to physical magnitudes maximal locally in the factor spaces. Perhaps this might be called "nonseparability," but we shall refer to it as failure of "ontological locality." Let us state the principle of Ontological Locality.

Ontological Locally (OLOC): If H_1 and H_2 are the Hilbert spaces for two spatially separated systems and $(\hat{A} \otimes \hat{1})$ is a locally maximal operator then

$$[A \otimes 1]^{\phi}_{\{X\}} = [A \otimes 1]^{\phi}_{\{Y\}}$$

for any state ϕ of the joint system where $\hat{X}$ and $\hat{Y}$ are both maximal operators in $H_1 \otimes H_2$ and $[\hat{X}, \hat{Y}] \neq 0$. In other words, locally maximal physical magnitudes on either of two spatially seperated systems are not "split" by ontological contextuality relative to the specification of different maximal physical magnitudes for the joint system.

There have been attempts in the literature to find a purely algebraic proof of nonlocality. Thus Demopoulos[(13)] has attempted to prove that the partial algebra of locally maximal and maximal magnitudes on two separated systems cannot be embedded in a commutative algebra. If successful this would demonstrate that we must treat $(A \otimes 1)_{\{X\}}$ and $(A \otimes 1)_{\{Y\}}$ as different magnitudes. In fact the proof fails.[12] This means that we do not need to violate OLOC to rescue realism from inconsistency. It does not follow of course that OLOC is not violated.

But locality was an issue brought to the fore in the foundations of quantum mechanics by Bell and in the proofs of nonlocality which proceed *via* the violation of some form of the Bell inequality there is another sense of locality involved.[13] Associated with this is another form of contextuality. We have ontological contextuality, the further form of contextuality which we will call *Environmental Contextuality* is totally different in origin though, as

[12] See papers by Humphreys[(14)] and Bub[(15)] and a forthcoming review by Heywood.[(16)] The algebraic approach appears to have originated with Bub[(17)] who posed the problem of extending Maczynski's[(12)] theorem on the Boolean representability of maximal magnitudes to locally maximal magnitudes.

[13] See Clauser and Shimony[(18)] and Redhead[(19)] for a detailed discussion of this approach. The effective tacit assumption of OLOC in this work should be noted.

we shall see, it is *experimentally* indistinguishable from ontological contextuality. Environmental contextuality involves the idea that there is some nonquantum interaction between the system of interest and its surroundings which occurs before the act of measurement and alters the values of the magnitudes of the system. These interactions are invoked to explained the failure of the Bell inequalities and are thus supposed to occur just before a measurement takes place, but presumably there is no reason why they should only occur immediately prior to a measurement which is why we call this kind of contextuality 'environmental.' By their very nature we know next to nothing about these supposed interactions, but we do presume that when they occur just before a measurement they are in fact a nonquantum mechanical interaction taking place between the measured system and the measuring apparatus depending on among other things the maximal magnitude on the measured system which the apparatus is set up to measure. Expanding our notation a little further we write $[A]^{\phi}_{\{B\}}(B)$ which is the value that the magnitude A_B takes after the interaction between the system and an apparatus set to measure B, but before the actual measurement takes place. The letter in parentheses labels the magnitude which the apparatus *is set* to measure. To make the meaning of this symbol clearer consider $[A]^{\phi}_{\{B\}}(C)$ where B and C are incompatible maximal magnitudes. This simply means the *value* that the magnitude A_B would take if the measuring apparatus were set to measure C, of course we can never know what this value is, because we cannot measure A_B and leave the apparatus set to measure C—B and C are incompatible. When we apply *this* idea of contextuality to spatially separated systems we come up with another form of locality, we shall call it Environmental Locality.

Environmental Locality (ELOC): If S_1 and S_2 are two spatially separated systems, Q, a physical magnitude for S_1, X and Y maximal magnitudes for the joint system $S_1 + S_2$, then if the difference between an apparatus set to measure X and one set to measure Y is only in the setting of that part of it at S_2

$$[Q \otimes 1]^{\phi}_{\{X\}}(X) = [Q \otimes 1]^{\phi}_{\{X\}}(Y)$$

In other words, the value possessed by a local physical magnitude cannot be changed by altering the arrangement of a *remote* piece of apparatus which forms *part* of the measurement context for the combined system. Note that we have *not* presumed OLOC in the spcification of ELOC, although as we shall discuss later, it is only when OLOC obtains that ELOC can *properly* be called a locality principle at all.

What we shall do in this paper is to link these discussions of locality in the following way. We shall show for an appropriate physical system that

$$\text{FUNC}^* + \text{VR} + \text{ELOC} + \text{OLOC} \Rightarrow \text{Contradiction}$$

where the contradiction is derived from a connection between value assignments that is very closely connected with FUNC although as we shall see differs from FUNC in a rather subtle way. Hence, FUNC* + VR ⇒ (~ELOC) or (~OLOC).[14] So we are presented with a dilemma. If we hold on to FUNC* and VR we must violate either ELOC or OLOC (or both). The implications of grasping either horn of the dilemma will be discussed in section 4. We turn now to the formal proof of our main result.

2. THE COMEASURABLE VALUE RULE

We begin by slightly amending VR to take proper account of environmental contextuality. Considering the case of a maximal magnitude R, we write

$$P_R^\phi(\lambda) = 0 \Rightarrow [R]^\phi(R) \neq \lambda$$

i.e., the vanishing probability of observing measurement results constrains the values R may have *in the measurement context of R*, and is silent about the quantity $[R]^\phi(P)$ for example, where P is an incompatible maximal magnitude.

Similarly we adapt FUNC* to take account of environmental contextuality: for any environmental context C

$$[A]_{\{B\}}^\phi(C) = h([D]_{\{B\}}^\phi(C))$$

where $\hat{B}$ is maximal and $\hat{A} = h(\hat{D})$,

$$\hat{A} = f(\hat{B}) \qquad \text{and} \qquad \hat{D} = g(\hat{B})$$

In particular we have the result

$$[A]_{\{B\}}^\phi(B) = f([B]^\phi(B))$$

which is the form in which we shall employ FUNC* later in this section.

From VR and FUNC* we shall now derive what we call the Comeasurable Value Rule (CVR). In his paper "On the Completeness of

[14] We are indebted to the eagle-eyed Arthur Fine for pointing out to us that FUNC* should be included among the assumptions of our proof.

Quantum Theory"[4] Fine moots, but quickly rejects, a constraint on value assignments which we call the Extended Value Rule. It is as follows:

Extended Value Rule (EVR): If $\hat{Q}_1$ and $\hat{Q}_2$ commute, $P^{\phi}_{Q_1,Q_2}(\lambda,\mu)=0 \Rightarrow$ either $[Q_1]^{\phi}\neq\lambda$ or $[Q_2]^{\phi}\neq\mu$ where $P^{\phi}_{Q_1,Q_2}(\lambda,\mu)$ denotes the quantum mechanical joint probability of finding measurement results λ for the physical magnitude Q_1 and μ for the physical magnitude Q_2 in the state ϕ. The reason he so swiftly rejected this rule is because it is easily shown to imply FUNC. Thus we have, from the statistical algorithm of QM, if W_ϕ is the density operator associated with the state ϕ and χ_Δ denotes the usual characteristic function associated with the set Δ,

$$\begin{aligned} P^{\phi}_{Q,f(Q)}(\lambda,\mu) &= \mathrm{Tr}\ W_\phi\cdot\chi_\lambda(Q)\cdot\chi_\mu(f(Q)) \\ &= \mathrm{Tr}\ W_\phi\cdot\chi_\lambda(Q)\cdot\chi_{f^{-1}(\mu)}(Q) \\ &= \mathrm{Tr}\ W_\phi\cdot\chi_{\{\lambda\}\cap f^{-1}(\mu)}(Q) \end{aligned}$$

Hence if $\lambda\notin f^{-1}(\mu)$ i.e., if $\mu\neq f(\lambda)$ it follows that

$$P^{\phi}_{Q,f(Q)}(\lambda,\mu)=0$$

and hence by EVR

$$[Q]^{\phi}\neq\lambda \qquad \text{or} \qquad [f(Q)]^{\phi}\neq\mu$$

But suppose

$$[Q]^{\phi}=\lambda$$

then it follows that

$$[f(Q)]^{\phi}\neq\mu, \qquad \forall p\neq f(\lambda)$$

i.e.,

$$[f(Q)]^{\phi}=f(\lambda)=f([Q]^{\phi})$$

which is FUNC.

We shall now introduce our CVR as a restriction of EVR to certain comesurable magnitudes.

It is a necessary condition for two magnitudes to be comeasurable that their associated self-adjoint operators commute, but this is not sufficient when account is taken of contextuality. It is *certainly* a sufficient condition for genuine comeasurability that the two magnitudes Q_1 and Q_2 are defined in the context of the *same* maximal magnitude R.

We now introduce our new rule.

The Comeasurable Value Rule (CVR): If $\hat{Q}_1$ and $\hat{Q}_2$ commute, and $\hat{R}$ is a maximal operator such that $\hat{Q}_1 = f(\hat{R})$ and $\hat{Q}_2 = g(\hat{R})$ for functions f and g, then $P^{\phi}_{Q_1, Q_2}(\lambda, \mu) = 0 \Rightarrow$ either $[Q_1]^{\phi}_{\{R\}}(R) \neq \lambda$ or $[Q_2]^{\phi}_{\{R\}}(R) \neq \mu$. We shall now show that CVR follows from VR and FUNC*. It is easy to see that

$$\begin{aligned} P^{\phi}_{Q_1, Q_2}(\lambda, \mu) &= \text{Tr } W_{\phi} \cdot \chi_{\lambda}(f(R)) \cdot \chi_{\mu}(g(R) = \text{Tr } W_{\phi} \cdot \chi_{f^{-1}(\lambda)}(R) \cdot \chi_{g^{-1}(\mu)}(R)) \\ &= \text{Tr } W_{\phi} \cdot \chi_{f^{-1}(\lambda)) \cap g^{-1}(\mu)}(R) \\ &= P^{\phi}_{R}(f^{-1}(\lambda) \cap g^{-1}(\mu)) \end{aligned}$$

Hence from the Value Rule

$$P^{\phi}_{Q_1, Q_2}(\lambda, \mu) = 0 \Rightarrow [R]^{\phi}(R) \notin f^{-1}(\lambda) \cap g^{-1}(\mu)$$

But suppose

$$[Q_1]^{\phi}_{\{R\}}(R) = \lambda \qquad \text{and} \qquad [Q_2]^{\phi}_{\{R\}}(R) = \mu$$

Then by FUNC*

$$[R]^{\phi}(R) \in f^{-1}(\lambda) \qquad \text{and} \qquad [R]^{\phi}(R) \in g^{-1}(\mu)$$

Thus,

$$[R]^{\phi}(R) \in f^{-1}(\lambda) \cap g^{-1}(\mu)$$

Hence,

$$\begin{aligned} P^{\phi}_{Q_1, Q_2}(\lambda, p) = 0 \Rightarrow & \sim([Q_1]^{\phi}_{\{R\}}(R) = \lambda \quad \text{and} \quad [Q_2]^{\phi}_{\{R\}}(R) = \mu) \\ & \Rightarrow \text{either } [Q_1]^{\phi}_{\{R\}}(R) \neq \lambda \text{ or } [Q_2]^{\phi}_{\{R\}}(R) \neq \mu \end{aligned}$$

which is just our CVR. Conversely from CVR it is clear that we can derive only the restricted, acceptable version of FUNC, viz FUNC*.

Let us develop a particular case of CVR which will interest us especially. Suppose $\hat{Q} \otimes \hat{I}$ and $\hat{I} \otimes \hat{Q}'$ are self-adjoint operators in some product space $H_1 \otimes H_2$ describing the states of two spatially separated systems S_1 and S_2. We use the convention that in the tensor product $\hat{A} \otimes \hat{B}$ of two operators the left-hand operator $\hat{A}$ acts on the space H_1 and the right-hand operator $\hat{B}$ on the space H_2. Let $\hat{Q} = h(\hat{A})$ and $\hat{Q}' = k(\hat{B})$ for functions h and k. Then $\hat{Q} \otimes \hat{I} = h(\hat{A}) \otimes \hat{I} = h(\hat{A} \otimes \hat{I})$. Similarly $\hat{1} \otimes \hat{Q}' = k(\hat{1} \otimes \hat{B})$. We suppose $\hat{A}$ and $\hat{B}$ are nondegenerate on their respective components of the product space so $A \otimes 1$ and $1 \otimes B$ are locally maximal physical magnitudes. Let $\hat{A} = \sum_i \alpha_i \hat{P}_i$ and $\hat{B} = \sum_i \beta_i \hat{P}'_i$ be the spectral resolutions of

$\hat{A}$ and $\hat{B}$. Now any physical magnitude with associated self-adjoint operator $\hat{0} = \sum_{ij} c_{ij} \hat{P}_i \otimes \hat{P}'_j$ where $c_{ij} = F(\alpha_i, \beta_j)$ and F: $\mathbb{R}^2 \to \mathbb{R}$ is $1-1$, is sufficient to show that CVR will apply to $Q \otimes I$ and $I \otimes Q'$. Since F is 1–1 there are functions f and g such that $f(c_{ij}) = \alpha_i$ and $g(c_{ij}) = \beta_j$, so

$$\begin{aligned} hf(\hat{0}) &= h\left(\sum_{ij} f(c_{ij}) \hat{P}_i \otimes \hat{P}'_j\right) \\ &= h\left(\sum_i \alpha_i \hat{P}_i \otimes \sum_j P'_j\right) \\ &= h(\hat{A} \otimes \hat{I}) \\ &= \hat{Q} \otimes \hat{I} \end{aligned}$$

Similarly

$$kg(\hat{0}) = \hat{I} \otimes \hat{Q}'$$

Thus in a product space $H_1 \otimes H_2$

$$P^{\phi}_{Q \otimes I, I \otimes Q'}(X, Y) = 0$$

By CVR implies

$$\text{either } [Q \otimes I]^{\phi}_{\{0\}}(0) \neq X \text{ or } [I \otimes Q']^{\phi}_{\{0\}}(0) \neq Y \tag{2}$$

Before applying this result we want to give a succint formulation of ELOC and OLOC. To this end we first *define* the symbol $\langle A, B \rangle$ to mean just the maximal physical magnitude 0 for the joint system whose associated operator $\hat{0}$ is constructed from the operators $\hat{A}$ and $\hat{B}$ associated with the component systems in the way described. Note that the bracket symbol denotes a *function* defined on the ordered pair (A, B) which is (partly) specified by the function F.

Now in order to measure 0 we connect up two pieces of apparatus, one interacting with S_1 and adjusted to measure A on S_1, the other interacting with S_2 and adjusted to measure B on S_2. The ordered pair of these measurement results is then subjected to the function F. The resulting number is a measurement of 0. The environmental context referring to a measurement of 0 can thus be spelled out as the ordered pair (A, B), indicating that the apparatus interacting with S_1 is set to measure A and the apparatus interacting with S_2 is set to measure B.

Our result (2) can now be expressed in the following form:

$$\text{If } P^{\phi}_{Q \otimes I, I \otimes Q'}(X, Y) = 0 \text{ then either } [Q \otimes I]^{\phi}_{\{\langle A, B\rangle\}}(A, B) \neq X \text{ or } [I \otimes Q']^{\phi}_{\{\langle A, B\rangle\}}(A, B) \neq Y \tag{3}$$

The suffix $\{\langle A, B\rangle\}$ shows that it does not matter which $1-1$ function we choose for F in the specification of $\langle A, B\rangle$. We shall follow the convention that in an ordered pair of symbols the first member always refers to system S_1 and the second member to system S_2.

We are now in a position to state our two locality principles (one of them, OLOC, in a slightly extended form) using our new notation. For all Q, A, B, C, D, and E, where $Q = h(A)$ for some function h, and A, B, C, D, and E are all maximal

OLOC:

$$[Q \otimes I]^{\phi}_{\{\langle A,B\rangle\}}(D, E) = [Q \otimes I]^{\phi}_{\{\langle A,C\rangle\}}(D, E) \tag{4}$$

ELOC:

$$[Q \otimes I]^{\phi}_{\{\langle A,B\rangle\}}(D, E) = [Q \otimes I]^{\phi}_{\{\langle A,B\rangle\}}(D, C) \tag{5}$$

Notice that in the formulation (4) of OLOC we have employed FUNC* to write

$$[Q \otimes I]^{\phi}_{\langle\{A,B\rangle\}}(D, E) = h([A \otimes I]^{\phi}_{\{\langle A,B\rangle\}}(D, E))$$

and

$$[Q \otimes I]^{\phi}_{\{\langle A,C\rangle\}}(D, E) = h([A \otimes I]^{\phi}_{\{\langle A,C\rangle\}}(D, E))$$

Eq. (4) then follows from our previous formulation of OLOC in terms of locally maximal magnitudes such as $A \otimes I$. *Mutatis mutandis* we can apply (4) and (5) to physical magnitudes associated with operators $I \otimes Q'$ which are local for S_2.

3. THE INCOMPATIBILITY OF CVR AND LOCALITY

Now we shall employ the results (3), (4), and (5) to derive a contradiction. In the proof we shall again consider as one system $S_1 + S_2$, two spatially separated systems S_1 and S_2. We suppose for simplicity that each system is associated with a Hilbert space of N-dimension and consider two locally maximal self-adjoint operators $\hat{A} \otimes \hat{I}$ and $\hat{I} \otimes \hat{B}$. Let $\hat{A}$ have N distinct eigenvalues written in some arbitrary order as $a_1, \ldots, a_N$ and $\hat{B}$ have N distinct eigenvalues written in some arbitrary order as $b_1, \ldots, b_N$. Let the state of the combined system be

$$\Psi = \sum_{m=1}^{N} C_m \, |a_m\rangle \otimes |b_m\rangle \tag{6}$$

where C_m are unspecified complex coefficients and $|a_m\rangle$ denotes the eigenvector of $\hat{A}$ associated with the eigenvalue a_m and $|b_m\rangle$ the eigenvector of $\hat{B}$ associated with the eigenvalue b_m. Ψ is thus a linear combination of simultaneous eigenvectors of $\hat{A} \otimes \hat{I}$ and $\hat{I} \otimes \hat{B}$, the mth eigenvalue of $\hat{A}$ being correlated with the mth eigenvalue of $\hat{B}$.

Now consider some nonmaximal self-adjoint operator $\hat{Q}$ such that $\hat{Q} = f(\hat{A}) = g(\hat{A}')$ for suitable functions f and g, where $\hat{A}'$ is another maximal operator which does not commute with $\hat{A}$. By construction we have the result

$$P^{\Psi}_{A\otimes I, I\otimes B}(a_m, y) = 0 \qquad \forall y \neq b_m \tag{7}$$

Hence applying (3) we obtain

$$\begin{aligned}
[A \otimes I]^{\psi}_{\{\langle A,B\rangle\}}(A, B) &= a_m \\
&\Rightarrow [I \otimes B]^{\Psi}_{\{\langle A,B\rangle\}}(A, B) \neq y \qquad \forall y \neq b_m \\
&\Rightarrow [I \otimes B]^{\Psi}_{\{\langle A,B\rangle\}}(A, B) = b_m
\end{aligned} \tag{8}$$

Also

$$P^{\Psi}_{f(A\otimes I), I\otimes B}(f(x), b_m) = P^{\psi}_{A\otimes I, I\otimes B}(f^{-1}(f(x)), b_m)$$

by an obvious extension of our notation to allow for the value of $A \otimes I$ revealed by measurement to be a member of the set $f^{-1}(f(x))$. But

$$\begin{aligned}
P^{\Psi}_{A\otimes I, I\otimes B}(f^{-1}(f(x)), b_m) = 0 \qquad &\forall x \text{ such that } \quad a_m \notin f^{-1}(f(x)) \\
&\text{i.e., such that } \quad f(a_m) \neq f(x)
\end{aligned}$$

Hence again using (3)

$$\begin{aligned}
[I \otimes B]^{\Psi}_{\{\langle A',B\rangle\}}(A', B) = b_m &\Rightarrow [f(A \otimes I)]^{\Psi}_{\{\langle A',B\rangle\}}(A', B) \neq f(x) \\
&\qquad \forall x \text{ such that } f(a_m) \neq f(x) \\
&\Rightarrow [f(A \otimes I)]^{\Psi}_{\{\langle A',B\rangle\}}(A', B) = f(a_m)
\end{aligned} \tag{9}$$

We now apply OLOC and ELOC in the form (4) and (5), suitably transposed to system S_2, to give

$$[I \otimes B]^{\Psi}_{\{\langle A,B\rangle\}}(A, B) = [I \otimes B]^{\Psi}_{\{\langle A',B\rangle\}}(A', B)$$

So from (8) and (9) we obtain

$$[A \otimes I]^{\Psi}_{\{\langle A,B\rangle\}}(A, B) = a_m \Rightarrow [f(A \otimes I)]^{\Psi}_{\{\langle A',B\rangle\}}(A', B) = f(a_m) \tag{10}$$

or more succintly

$$[f(A \otimes I)]^{\Psi}_{\{\langle A',B\rangle\}}(A', B) = f([A \otimes I]^{\Psi}_{\{\langle A,B\rangle\}}(A, B)) \tag{11}$$

Now (11) is not quite FUNC, since the latter principle preserves functional relationships between the *coexisting* values attributed to physical magnitudes, whereas in (11) we have *different* environmental contexts on the two sides of the equation.

Nevertheless (11) can be used to demonstrate a contradiction just as well as the noncontextualized form of FUNC. To see this let us specialize to the case of two spin-1 systems prepared so that the combined system $S_1 + S_2$ is in the singlet state of the total spin. If α is any direction and $|\alpha_+\rangle$, $|\alpha_-\rangle$, and $|\alpha_0\rangle$ are eigenvectors of the spin component of either system projected in the α – direction with eigenvalues 1, -1, and 0 (in units of $\hbar$), respectively, then Ψ is given by

$$\Psi = 1/\sqrt{3}\ (|\alpha_+\rangle \otimes |\alpha_-\rangle + |\alpha_-\rangle \otimes |\alpha_+\rangle - |\alpha_0\rangle \otimes |\alpha_0\rangle) \tag{12}$$

We now take for the operators $\hat{A}$ and $\hat{B}$ in our general analysis the single operator

$$\hat{H}_s = a\hat{S}_x^2 + b\hat{S}_y^2 + c\hat{S}_z^2 \tag{13}$$

considered by Kochen and Specker[1] where $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$ are the operators associated with the x, y, and z components of spin of either system. $\hat{H}_s$ has eigenvalues $(a+b)$, $(a+c)$, and $(b+c)$ and we assume a, b, and c are distinct real numbers so that these eigenvalues are all unequal. Hence $\hat{H}_s$ is indeed maximal. Ψ can be expressed in terms of its eigenvectors. If α is identified with the z-direction, given the relations

$$\begin{aligned} |\alpha_+\rangle &= 1/\sqrt{2}\,(|a+c\rangle - |b+c\rangle) \\ |\alpha_-\rangle &= 1/\sqrt{2}\ (|a+c\rangle + |b+c\rangle) \\ |\alpha_0\rangle &= |a+b\rangle \end{aligned} \tag{15}$$

where $|a+b\rangle$, $|a+c\rangle$, and $|b+c\rangle$ are eigenvectors of $\hat{H}_s$ with the indicated eigenvalues, it is easily verified that

$$\Psi = 1/\sqrt{3}\ (|a+c\rangle \otimes |a+c\rangle - |b+c\rangle \otimes |b+c\rangle - |a+b\rangle \otimes |a+b\rangle) \tag{15}$$

This is of the form (6) to which our general analysis applies. Take now for $\hat{A}'$ the operator

$$\hat{H}'_s = a\hat{S}_{x'}^2 + b\hat{S}_{y'}^2 + c\hat{S}_z^2 \tag{16}$$

where the new set of orthogonal directions denoted by the labels $x'y'$, and z are obtained from those denoted by x, y, z by a rotation about the z-axis. Finally take for $\hat{Q}$ the operator

$$S_z^2 = f(\hat{H}_s) = f(\hat{H}_s') \tag{17}$$

where f: $\mathbb{R} \to \mathbb{R}$ is defined by

$$f(x) = (c-a)^{-1}(b-c)^{-1}(x-(a+b))(x-2c) \tag{18}$$

(Note that since f does not possess an inverse (17) does not imply $\hat{H}_s = \hat{H}_s'$!) Since $f(\hat{H}_s \otimes I) = f(\hat{H}_s) \otimes \hat{I} = \hat{S}_z^2 \otimes \hat{I}$ our result (11) can be expressed in an abbreviated notation in the form

$$[\hat{S}_z^2 \otimes I]^{\Psi}_{\{H_s' \otimes I\}}(H_s' \otimes I) = f([H_s \otimes I]^{\Psi}(H_s \otimes I)) \tag{19}$$

where we have suppressed any contextuality parameter either ontological or environmental on which the values of the physical magnitudes do not depend in the light of OLOC and ELOC. (We write the contextuality parameters as $\hat{H}_s' \otimes I$ and $H_s \otimes I$ to indicate, in accordance with our convention for the order of factors in a tensor product, which system we are referring to. Having dropped the ordered pair notation used in (11) we need some other method of distinguishing which system H_s or H_s' refers to.)

Now (19) assigns to $[S_z^2 \otimes I]_{\{H_s' \otimes I\}}(H_s' \otimes I)$ and, most importantly, to the direction labeled z, a number with the following two properties:

1. It is independent of the orientation of the x' and y' axes used in specifying $\hat{H}_s'$.
2. It has the value 0 or 1.

Condition (2) follows from the fact that $H_s \otimes I$ must be assigned one of its eigenvalues and the action of f as specified in (18) is to project the set of eigenvalues onto the set $\{0, 1\}$. Now we can repeat the argument which led to (19) using instead an orthogonal triad x', y, and z', so that we find that

$$[S_y^2 \otimes I]^{\Psi}_{\{H_s'' \otimes I\}}(H_s'' \otimes I)$$

and hence the direction y is assigned some number which is a different function, say g, of $[H_s \times I]^{\Psi}(H_s \otimes I)$ where

$$\hat{H}_s'' = a\hat{S}_{x'}^2 + b\hat{S}_y^2 + c\hat{S}_{z'}^2$$

This number again has the properties:

1. It is independent of the x' and z' axes used in specifying $\hat{H}_s''$.
2. It has the value 0 or 1.

The function g is given by

$$g(x)=(b-c)^{-1}(a-b)^{-1}(x-(c+a))(x-2b) \tag{20}$$

Finally the argument is repeated for $S_x^2 \otimes I$ and the number assigned to $[S_x^2 \otimes I]^{\Psi}_{\{H_s'''\otimes I\}}(H_s''' \otimes I)$, and hence the direction x, is a new function, say h, of $[H_s \otimes I]^{\Psi}(H_s \otimes I)$ with the properties:

1. It is independent of the y' and z' axes used in specifying a rotated operator $\hat{H}_s'''=a\hat{S}_x^2+b\hat{S}_{y'}^2+c\hat{S}_{z'}^2$.
2. It has the value 0 or 1.

The function h is given by

$$h(x)=(a-b)^{-1}(c-a)^{-1}(x-(b+c))(x-2a) \tag{21}$$

But the three functions f, g, and h have the property that, acting on any eigenvalue of $H_s \otimes \hat{I}$, the sum of the three values is always 2. This can be checked at once from the equations (18), (20), and (21) with x given the value $a+b$, $a+c$, or $b+c$.

So our final result is to assign to three arbitrarily chosen orthogonal directions x, y, z three unique numbers, each of which is 0 or 1, and whose sum is 2. But such an assignment of numbers is known to be impossible for an appropriately chosen finite set of orthogonal triads of directions in Euclidean 3-space. This is what Kochen and Specker[1] showed explicitly.

But notice that the numbers $[S_z^2 \otimes I]^{\Psi}_{\{H_s'\otimes I\}}(H_s' \otimes I)$,

$$[S_y^2 \otimes I]^{\Phi}_{\{H_s''\otimes I\}}(H_s'' \otimes I) \qquad \text{and} \qquad [S_z^2 \otimes I]^{\Psi}_{\{H_s'''\otimes I\}}(H_s''' \otimes I)$$

we use to get the contradiction, not only cannot in general be *measured* simultaneously, since $\hat{H}_s'$, $\hat{H}_s''$, and $\hat{H}_s'''$ do not commute, but cannot even be said to *coexist* simultaneously, owing to the differing environmental contexts. In contrast with the original Kochen–Specker paradox we are dealing now not with simultaneously existing value assignments (albeit not simultaneously measurable) but with numbers which *would* be assigned on the assumption of *differing* environmental contexts.

4. CONCLUSION

The upshot of our argument is that a quantum realist who wishes to impose the Value Rule and does not want to deny the innocent-looking FUNC* is bound to accept some form of nonlocality, that is either he must deny OLOC or ELOC.

Let us look at the two horns of the dilemma in turn. Violating OLOC means that we can no longer specify the properties of one system independently of a specification of properties relating to the whole combined system. This leads to an ontological holism in which it is impossible to make sense of a realist construal of quantum mechanics which associates properties (physical magnitudes) independently with each of two separated systems. If OLOC is violated the physical magnitudes associated with locally maximal operators are not themselves "local" at all. Hence the question whether such magnitudes can have their values changed by an environmental change in the way denied by imposing ELOC, is not really a locality issue. The violation of ELOC becomes a locality issue only if OLOC obtains. In such a case violation of ELOC shows that the value of a physical magnitude that may properly be said to pertain to one of two separated systems can have its value changed by altering the setting of an apparatus interacting with the other system. The consequences for a realist construal of quantum mechanics of such a situation have been spelled out by one of us elsewhere.[(19)] The reason why ELOC and OLOC, although conceptually quite distinct, have been conflated in the literature[15] is that violation of either principle when expressed in terms of *measurement* results demonstrates a dependence of the outcome recorded by the apparatus connected to one system on the setting of the apparatus connected to the other (remote) system.

If we decide to retain OLOC then our work provides a demonstration that ELOC is violated of quite a different character from that involved in discussions of the Bell inequality. Here the very considerable literature has concentrated on the assumptions implicit in the derivation of various forms of the Bell inequality. Fine[(4)] in particular has argued that all such proofs invoke a hidden assumption of the existence of joint probability distributions for incompatible magnitudes. Our own approach makes minimal and transparent use of probability theory, but is certainly free of any joint distribution assumption for physical magnitudes associated with noncommuting operators. In short it really does appear to us that realism in QM and locality just cannot reasonably be reconciled.[16]

[15] See in particular the otherwise excellent discussion by van Fraassen.[(20)]

[16] Our work also has relevance to another rule introduced by Fine.[(21)] This is his Correlation Rule, which he seeks to show implies FUNC, and hence is contradictory. But the

ACKNOWLEDGMENTS

We are grateful to Arthur Fine, Paul Teller, Abner Shimony, Roger Cook, and Jon Dorling for their attentive reading of an earlier draft of this paper. Their efforts greatly improved the presentation of our results, although none of the views expressed should be imputed to them. One of us, Peter Heywood, acknowledges the award of a Chelsea College Research Studentship during the tenure of which this work was carried out.

REFERENCES

1. S. Kochen and E. P. Specker, *J. Math. Mech.* **17**, 59 (1967).
2. J. S. Bell, *Physics* **1**, 195 (1964).
3. A. Stairs, *PhD Thesis*, University of Western Ontario, 1978.
4. A. Fine, *Synthese* **29**, 257 (1974).
5. A. Fine, *Synthese* **42**, 145 (1979).
6. P. Teller, *Phil. Sci.* **44**, 475 (1977).
7. A. M. Gleason, *J. Math. Mech.* **6**, 885 (1957).
8. J. S. Bell, *Rev. Mod. Phys.* **38**, 447 (1966).
9. A. Fine and P. Teller, *Found. Phys.* **8**, 629 (1978).
10. M. L. G. Redhead, *Phil. Sci.* **48**, 50 (1981).
11. B. C. van Fraassen, in *Contemporary Research in the Foundations and Philosophy of Quantum Theory*, C. A. Hooker, ed. (Reidel, Dordrecht, 1973), p. 80.
12. M. J. Maczynski, *Rep. Math. Phys.* **2**, 135 (1971).
13. W. Demopoulos, in *Studies in the Foundations of Quantum Mechanics*, P. Suppes, ed. (PSA, East Lansing, 1980), p. 119.
14. P. Humphreys, in *Studies in the Foundations of Quantum Mechanics*, P. Suppes, ed. (PSA, East Lansing, 1980), p. 145.
15. J. Bub, in *Studies in the Foundations of Quantum Mechanics*, P. Suppes, ed. (PSA, East Lansing, 1980), p. 149.
16. P. Heywood, *Studies in the Foundations of Quantum Mechanics*, P. Suppes, ed. (PSA, East Lansing, 1980); *Brit. J. Phil. of Sci.* (to appear).
17. J. Bub, *Found. Phys.* **6**, 511 (1976).
18. J. Clauser and A. Shimony, *Rep. Prog. Phys.* **41**, 1881 (1978).
19. M. L. G. Redhead, "Causality, Relativity and the Einstein-Podolsky-Rosen Paradox," in *Space, Time and Causality*, R. Swinburne, ed. (Reidel, Dordrecht, 1983), p. 151.
20. B. C. van Fraassen, *Synthese* **42**, 155 (1979).
21. A. Fine, *Phil. Sci.* **44**, 95 (1977).

Correlation Rule in the case of two separated systems is essentially the noncontextualized version of our (8). Fine concludes that his Correlation Rule cannot be maintained. This only follows if we do not allow for contextuality and possible locality violation in the way we have described.