Hidden variable theories: contextuality and nonlocality

J.E. Sipe

March 2019

In the last chapter we exhibited PVM-faithful models for a two-dimensional Hilbert space. The structure of operators acting over such a space is particularly simple, for any projector can only be part of one PVM; for example, of all the PVMs possible, the projector $P_{+\hat{\mathbf{a}}}$ considered in the last chapter is an element only of $\{P_{+\hat{\mathbf{a}}}, P_{-\hat{\mathbf{a}}}\}$. For Hilbert spaces of larger dimension the situation is more complicated, and new issues arise. For example, in our discussion of a spin-1 particle – involving a three-dimensional Hilbert space – we considered the PVM $\{P_a, P_b, P_c\}$ and the PVM $\{P_{a'}, P_{b'}, P_c\}$; both of these share the projector P_c. In an PVM-faithful model this has important consequences. For consider the sets of characteristic functions associated with these PVMs, $\{P_a(\lambda), P_b(\lambda), P_c(\lambda)\}$ and $\{P_{a'}(\lambda), P_{b'}(\lambda), P_c(\lambda)\}$. In each set for each λ two of the three functions must vanish, and the other equal unity. So for example if $P_c(\lambda) = 1$ for a particular λ, then $P_a(\lambda) = P_b(\lambda) = P_{a'}(\lambda) = P_{b'}(\lambda) = 0$ for that λ as well.

1 Disaster again

Even stronger constraints arise. For example, from our original basis $\{|a\rangle, |b\rangle, |c\rangle\}$ construct a range of kets

$$|v\rangle = x\,|a\rangle + y\,|b\rangle + z\,|c\rangle\,,$$

where (x, y, z) are all real, and

$$x^2 + y^2 + z^2 = 1.$$

We can formally identify each $|v\rangle$ (only a small fraction of the states that can be considered), with a point on the surface of a unit sphere. So we can associate a vector in three-space $\mathbf{v} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$ with the ket $|v\rangle$, $\mathbf{v} \leftrightarrow |v\rangle$.

Now for any two such vectors in Hilbert space,

$$\begin{aligned} |v_1\rangle &= x_1\,|a\rangle + y_1\,|b\rangle + z_1\,|c\rangle\,, \\ |v_2\rangle &= x_2\,|a\rangle + y_2\,|b\rangle + z_2\,|c\rangle\,, \end{aligned}$$

it is easy to confirm that $\langle v_1|v_2\rangle = 0$ if and only if the associated vectors in three-space are orthogonal, $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. So any orthogonal triad of unit vectors $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ can be associated with a basis of vectors in the Hilbert space, $\{|v_1\rangle, |v_2\rangle, |v_3\rangle\}$, or alternately with a PVM $\{P_{v_1}, P_{v_2}, P_{v_3}\}$, where $P_{v_1} = |v_1\rangle\langle v_1|$, etc. And, in an PVM-F model, we can associate any orthogonal triad of unit vectors $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ with a set of characteristic functions $\{P_{v_1}(\lambda), P_{v_2}(\lambda), P_{v_3}(\lambda)\}$; for a given λ, two of these functions must vanish, and one must equal unity.

To help picture things, associate the value of zero with the colour yellow, and the value of one with the colour green. Now we can consider the whole range of orthogonal triads of unit vectors, each unit vector in each triad identifying a point on the unit sphere, and with each orthogonal trial associate a set of characteristic functions. Then to be able to build a PVM-F model for a spin 1 system, it must be possible to paint the surface of a unit sphere in such a way that, of the vectors in *any* orthogonal triad, two point to yellow spots and one to a green spot.

This is impossible, as we sketch in the Appendix of this chapter. The result is known as the Bell-Kochen-Specker theorem. The proof in the Appendix is a generalization of an argument by Bell; in a different approach, Kochen and Specker proved that such a hidden variable model is impossible by considering a finite number of rays, rather than the infinite number that are invoked by the "colouring" argument. In their original argument in 1967, Kochen and Specker showed that an inconsistency could be demonstrated using 117 rays. For a while a "cottage industry" emerged, which aimed at demonstrating an inconsistency with fewer and fewer rays; see the review in the book by Bub. In published work in 1976 Jost demonstrated an inconsistency with 109 rays, while in even earlier, unpublished work Schütte demonstrated an inconsistency with 33 rays. In other unpublished work, Conway and Kochen demonstrated an inconsistency with 31 rays.

These arguments address a three-dimensional Hilbert space. But if an PVM-F model cannot be constructed in a three-dimensional Hilbert space – that is, there would be contradictions with the predictions of operational quantum mechanics – it is also impossible in higher dimensional spaces, because projectors in a three-dimensional subspace that correspond to those considered here can be found, and the inconsistency demonstrated for them. In fact, if one considers explicitly higher-dimensional Hilbert spaces even simpler proofs of inconsistency can be found. Here is one due to David Mermin for a four-dimensional Hilbert space, which we can consider as the direct product of Hilbert spaces describing spin$-\frac{1}{2}$ systems. For each spin introduce projectors associated with all directions $\hat{\mathbf{a}}$, $P^{(1)}_{+\hat{\mathbf{a}}}$ for the first spin and $P^{(2)}_{+\hat{\mathbf{a}}}$ for the second spin. Then for the first spin, for example, the component of the spin operator $\mathbf{S}^{(1)}$ in the $\hat{\mathbf{a}}$ direction can be written as $\left(\mathbf{S}^{(1)} \cdot \hat{\mathbf{a}}\right) = (/2)\,\sigma^{(1)}_{\hat{\mathbf{a}}}$, where the Hermitian operator $\sigma^{(1)}_{\hat{\mathbf{a}}} = P^{(1)}_{+\hat{\mathbf{a}}} - P^{(1)}_{-\hat{\mathbf{a}}}$. In the full Hilbert space we denote the direct product $\sigma^{(1)}_{\hat{\mathbf{a}}} \otimes \sigma^{(2)}_{\hat{\mathbf{b}}}$ by simply $\sigma^{(1)}_{\hat{\mathbf{a}}}\sigma^{(2)}_{\hat{\mathbf{b}}}$, and the direct product $\sigma^{(1)}_{\hat{\mathbf{a}}} \otimes \mathcal{I}^{(2)}$, where $\mathcal{I}^{(2)}$ is the identity operator over the Hilbert space of the second spin, by simply $\sigma^{(1)}_{\hat{\mathbf{a}}}$.

Then consider the following nine operators

$$\begin{array}{ccc}
\sigma_{\hat{\mathbf{x}}}^{(1)} & \sigma_{\hat{\mathbf{x}}}^{(2)} & \sigma_{\hat{\mathbf{x}}}^{(1)}\sigma_{\hat{\mathbf{x}}}^{(2)} \\
\sigma_{\hat{\mathbf{y}}}^{(2)} & \sigma_{\hat{\mathbf{y}}}^{(1)} & \sigma_{\hat{\mathbf{y}}}^{(1)}\sigma_{\hat{\mathbf{y}}}^{(2)} \\
\sigma_{\hat{\mathbf{x}}}^{(1)}\sigma_{\hat{\mathbf{y}}}^{(2)} & \sigma_{\hat{\mathbf{y}}}^{(1)}\sigma_{\hat{\mathbf{x}}}^{(2)} & \sigma_{\hat{\mathbf{z}}}^{(1)}\sigma_{\hat{\mathbf{z}}}^{(2)}
\end{array}$$

It is easy to show that the operators in each of the three rows and each of the three columns are mutually commuting. As well, the product of the operators in any row is the identity $\mathcal{I} = \mathcal{I}^{(1)} \otimes \mathcal{I}^{(2)}$, while the product of the operators in each of the first two columns is $\mathcal{I}$, while the product of the operators in the third column is $-\mathcal{I}$. Now in a PVM-F model each of these operators will have a variable associated with it; write those variables as

$$\begin{array}{ccc}
\sigma_{\hat{\mathbf{x}}}^{(1)}(\lambda) & \sigma_{\hat{\mathbf{x}}}^{(2)}(\lambda) & \sigma_{\hat{\mathbf{x}}}^{(1)}\sigma_{\hat{\mathbf{x}}}^{(2)}(\lambda) \\
\sigma_{\hat{\mathbf{y}}}^{(2)}(\lambda) & \sigma_{\hat{\mathbf{y}}}^{(1)}(\lambda) & \sigma_{\hat{\mathbf{y}}}^{(1)}\sigma_{\hat{\mathbf{y}}}^{(2)}(\lambda) \\
\sigma_{\hat{\mathbf{x}}}^{(1)}\sigma_{\hat{\mathbf{y}}}^{(2)}(\lambda) & \sigma_{\hat{\mathbf{y}}}^{(1)}\sigma_{\hat{\mathbf{x}}}^{(2)}(\lambda) & \sigma_{\hat{\mathbf{z}}}^{(1)}\sigma_{\hat{\mathbf{z}}}^{(2)}(\lambda)
\end{array} \tag{1}$$

In fact for a given λ each of these must be ± 1. Now fixing λ and forming the relations that hold because of the fact that the operators in each row commute, we must have

$$\begin{aligned}
\left(\sigma_{\hat{\mathbf{x}}}^{(1)}(\lambda)\right)\left(\sigma_{\hat{\mathbf{x}}}^{(2)}(\lambda)\right)\left(\sigma_{\hat{\mathbf{x}}}^{(1)}\sigma_{\hat{\mathbf{x}}}^{(2)}(\lambda)\right) &= 1 \\
\left(\sigma_{\hat{\mathbf{y}}}^{(2)}(\lambda)\right)\left(\sigma_{\hat{\mathbf{y}}}^{(1)}(\lambda)\right)\left(\sigma_{\hat{\mathbf{y}}}^{(1)}\sigma_{\hat{\mathbf{y}}}^{(2)}(\lambda)\right) &= 1, \\
\left(\sigma_{\hat{\mathbf{x}}}^{(1)}\sigma_{\hat{\mathbf{y}}}^{(2)}(\lambda)\right)\left(\sigma_{\hat{\mathbf{y}}}^{(1)}\sigma_{\hat{\mathbf{x}}}^{(2)}(\lambda)\right)\left(\sigma_{\hat{\mathbf{z}}}^{(1)}\sigma_{\hat{\mathbf{z}}}^{(2)}(\lambda)\right) &= 1,
\end{aligned} \tag{2}$$

and from the fact that the operators in each column commute we must have

$$\begin{aligned}
\left(\sigma_{\hat{\mathbf{x}}}^{(1)}(\lambda)\right)\left(\sigma_{\hat{\mathbf{y}}}^{(2)}(\lambda)\right)\left(\sigma_{\hat{\mathbf{x}}}^{(1)}\sigma_{\hat{\mathbf{y}}}^{(2)}(\lambda)\right) &= 1, \\
\left(\sigma_{\hat{\mathbf{x}}}^{(2)}(\lambda)\right)\left(\sigma_{\hat{\mathbf{y}}}^{(1)}(\lambda)\right)\left(\sigma_{\hat{\mathbf{y}}}^{(1)}\sigma_{\hat{\mathbf{x}}}^{(2)}(\lambda)\right) &= 1, \\
\left(\sigma_{\hat{\mathbf{x}}}^{(1)}\sigma_{\hat{\mathbf{x}}}^{(2)}(\lambda)\right)\left(\sigma_{\hat{\mathbf{y}}}^{(1)}\sigma_{\hat{\mathbf{y}}}^{(2)}(\lambda)\right)\left(\sigma_{\hat{\mathbf{z}}}^{(1)}\sigma_{\hat{\mathbf{z}}}^{(2)}(\lambda)\right) &= -1,
\end{aligned} \tag{3}$$

and we have used the fact that in a PVM-F model we have $\mathcal{I}(\lambda) = 1$ (perhaps obvious, but how would you prove it?) Now for a fixed λ the form the product of the nine quantities (1). According to (2) it should be 1, but according to (3) it is -1, giving a contradiction.

Mermin's example makes clear that the problem is in a kind of "mismatch" between the relations between operators over a Hilbert space and the simpler relations between variables assumed to be associated with these operators that take values equal to the eigenvalues of the operators. The structure of the algebra of the operators is just much richer than the structure of the algebra of the variables.

2 Alternatives

To review the situation: Consider the two PVMs $\mathcal{P}_I = \{P_a, P_b, P_c\}$ and $\mathcal{P}_{II} = \{P_{a'}, P_{b'}, P_c\}$, one or the other implemented on an ensemble of systems following a preparation characterized by the ray ψ. Operational quantum mechanics guarantees that the fraction of times outcome c results would be the same for both PVMs, and given by $\langle\psi|P_c|\psi\rangle$. Yet it is impossible to construct functions $P_c(\lambda)$ for all c such if the appropriate principal variables λ were such that $P_c(\lambda) = 1$, then in either experiment outcome c would result, and if they were such that $P_c(\lambda) = 0$ then in neither experiment would outcome c result. Measurements associated with Hermitian operators cannot be thought of as simply revealing the value of a variable associated with the operator; partner noncontextuality is not possible. It is clear that this did not plague the PVM-F models for two-dimensional Hilbert spaces presented in the last chapter simply because the feature of partner noncontextuality does not arise for two-dimensional Hilbert spaces; in a two-dimensional Hilbert space, each projector can have only one partner.

Stepping back, from one perspective the problem in higher dimensional Hibert spaces is perhaps not surprising. The gadgets implementing $\mathcal{P}_I$ and $\mathcal{P}_{II}$ would be different. Even if we maintain the realist assumption that there are a set of principal variables λ characterizing the state of the spin 1 particle, one could imagine that whether outcome c resulted or not might depend on the nature of the measurement gadget, which would be quite different for $\mathcal{P}_I$ and $\mathcal{P}_{II}$; after all, the measurements are different even at the operational level. So for a given set of principal variables λ one could reasonably imagine that outcome c results for one measurement, and not for the other. Generalizing the idea behind the PVM-F model, we might consider introducing different functions $P_c^{\mathcal{P}_I}(\lambda)$ and $P_c^{\mathcal{P}_{II}}(\lambda)$. Each would equal unity if the values of the principal variables were such that outcome c would result and zero if it would not – if a measurement characterized by the indicated PVM were performed – but in general $P_c^{\mathcal{P}_I}(\lambda) \neq P_c^{\mathcal{P}_{II}}(\lambda)$, because how the spin 1 particle behaved during the measurement process could be quite different for the gadgets implementing the two PVMs. Of course, for any preparation procedure the average over the probability distribution $\mu_\psi(\lambda)$ of $P_c^{\mathcal{P}_I}(\lambda)$ and $P_c^{\mathcal{P}_{II}}(\lambda)$ would have to be the same were the results of operational quantum mechanics to be recovered,

$$\begin{aligned} \langle\psi|P_c|\psi\rangle &= \int \mu_\psi(\lambda) P_c^{\mathcal{P}_I}(\lambda) d\lambda \\ &= \int \mu_\psi(\lambda) P_c^{\mathcal{P}_{II}}(\lambda) d\lambda. \end{aligned} \tag{4}$$

It would just be that for some particular values λ we would have $P_c^{\mathcal{P}_I}(\lambda) \neq P_c^{\mathcal{P}_{II}}(\lambda)$.

This kind of hidden variable model certainly exhibits partner contextuality. As well, imagine two gadgets, one characterized by $\mathcal{P}_I = \{P_a, P_b, P_c\}$ and the other by $\mathcal{P}_{II} = \{P_{a'}, P_{b'}, P_c\}$. Perhaps the outcome in each is indicated by the lighting of one of three light bulbs, where the bulbs are located far from the region of the gadget that interacts with the spin 1 system. Now imagine "rewiring" each gadget, in the first removing the bulbs associated with a and b and replacing them by a new bulb that will light whenever each a or b would have lit, and in the second removing the bulbs associated with a' and b' and replacing them by a new bulb that will light whenever a' or b' would have lit. With this rewiring, both gadgets can be thought of as implementing the PVM $\{P_c, \mathcal{I} - P_c\}$, with outcomes c and "not c." Yet there will be experimental runs where λ will be such that in one gadget outcome c will result and in the other gadget it will not, since for some values λ we have $P_a^{\mathcal{P}_I}(\lambda) \neq P_a^{\mathcal{P}_{II}}(\lambda)$. Thus this kind of model also exhibits gadget contextuality.

This suggests that the labeling of the functions $\left\{P_k^{\mathcal{P}}(\lambda)\right\}$ simply by the PVM $\mathcal{P}$ would not even be sufficient; the set of functions would have to be assumed to be different depending on the full details of the measurement apparatus. Alternately, one could posit that the kind of description we have sketched would be assumed to hold only if the PVMs with which functions $\left\{P_k^{\mathcal{P}}(\lambda)\right\}$ are introduced are maximal PVMs, such as $\mathcal{P}_I = \{P_a, P_b, P_c\}$ and $\mathcal{P}_{II} = \{P_{a'}, P_{b'}, P_c\}$, rather than nonmaximal, such as $\{P_c, I - P_c\}$. While functions $\left\{P_k^{\mathcal{P}}(\lambda)\right\}$ could be used for maximal PVMs, the results for nonmaximal PVMs would then be taken to follow from coarse-graining the results of the maximal PVMs, and it would simply be acknowledged that for a particular implementation of a nonmaximal PVM the sub-quantum behaviour would depend on the particular coarse-graining that was used to get to the nonmaximal PVM, reflecting the difference in the gadgets, as in the example above. While partner contextuality would still arise, we could at least maintain a gadget noncontextuality for *maximal PVMs*. We can summarize this kind of model in this way:

> A “maximal PVM faithful” (mPVM-F) hidden variable model for maximal PVMs involving *sets of projector variables* that are functions of principal variables λ is possible if, for every ray ψ, there is a probability distribution function $\mu_\psi(\lambda) \geq 0$ with
>
> $$\int \mu_\psi(\lambda) d\lambda = 1 \tag{5}$$
>
> and if, for every maximal PVM $\mathcal{P} = \{P_k\}$, there is a set of characteristic functions $\left\{P_k^{\mathcal{P}}(\lambda)\right\}$ with
>
> $$\begin{aligned} & P_k^{\mathcal{P}}(\lambda) \text{ either 0 or 1 for all } \lambda, \\ \sum_k P_k^{\mathcal{P}}(\lambda) &= 1 \text{ for all } \lambda \end{aligned} \tag{6}$$
>
> such that
>
> $$\langle\psi|P_k|\psi\rangle = \int \mu_\psi(\lambda) P_k^{\mathcal{P}}(\lambda) d\lambda \tag{7}$$

A mPVM-F model, or something like it, gives up the idea that the sub-quantum behavior completely respects the equivalence classes of PVMs at the operational level: While in operational quantum mechanics all expectation values of projectors are independent of which PVM characterizes the measurement that involves them, in a mPVM-F at the sub-quantum level different PVMs will lead to different results in particular experimental runs, since generally $P_c^{\mathcal{P}_I}(\lambda) \neq P_c^{\mathcal{P}_{II}}(\lambda)$, for example. While no contradiction arises from these two statements, it does lead to the suggestion of a Nature that is conspiratorial, always hiding in averages over experimental runs (4) the differences that appear in individual runs. Is there a way to avoid that, and somehow hold on to measurement noncontextuality? One suggestion would be to acknowledge that an PVM-F model is indeed too simple, but because there is either a fundamental indeterminism in what measurement outcome results, or that the result depends on the values of principal variables of the measurement gadget as well, and so cannot be expressed deterministically in terms of the values of the principal variables of the quantum system being measurement alone. In either case a natural extension the PVM-F would be to replace the set of characteristic functions $\{P_k(\lambda)\}$ by a set of *indicator functions* $\{\xi_k(\lambda)\}$. This is defined as a set of functions satisfying $\xi_k(\lambda) \geq 0$ and

$$\sum_k \xi_k(\lambda) = 1.$$

Here $\xi_k(\lambda)$ is understood as the probability that outcome k is the result if the values of the principal variables of the quantum system are λ. Whether the indeterminism is considered fundamental or not, this kind of model could then be stated as:

> An "indeterministic PVM" (iPVM) hidden variable model for PVMs involving *indicator functions* of principal variables λ is possible if, for every ray ψ, there is a probability distribution function $\mu_\psi(\lambda) \geq 0$ with
>
> $$\int \mu_\psi(\lambda) d\lambda = 1$$
>
> and if, for every PVM $\{P_k\}$ there is a set of indicator functions $\{\xi_k(\lambda)\}$ with
>
> $$\begin{aligned} \xi_k(\lambda) &\geq 0, \\ \sum_k \xi_k(\lambda) &= 1 \text{ for all } \lambda \end{aligned}$$
>
> such that
>
> $$\langle\psi|P_k|\psi\rangle = \int \mu_\psi(\lambda)\xi_k(\lambda)d\lambda \tag{8}$$

The model is measurement noncontextual because the probability $\xi_k(\lambda)$ depends only on the projector P_k and not on what PVM, of which P_k is an element, it is that characterizes the measurement; although probabilities enter here, they respect the equivalence class structure of the measurements in operational quantum mechanics. Of course, one could easily generalize this to define an outcome-indeterministic hidden variable model that would exhibit partner contextuality, in which they would not.

Are either of these two new hidden variable models viable?

3 The "de-Ockhamization" of quantum mechanics

mPVM-F models can be identified that are in agreement with the predictions of quantum mechanics. Here is one way to do it: For each maximal PVM $\mathcal{P} = \{P_k\}$ – where $\mathcal{P} = I, II, III....$, in principle an infinite number of maximal PVMS – introduce a principal variable $\lambda_{\mathcal{P}}$, a real number that ranges from 0 to 1. We take $P_k^{\mathcal{P}}(\lambda)$ to depend only on $\lambda_{\mathcal{P}}$, and if there are N elements in the PVM $\mathcal{P}$ we put

$$\begin{aligned}
P_1^{\mathcal{P}}(\lambda_{\mathcal{P}}) &= 1 \text{ for } 0 \leq \lambda_{\mathcal{P}} \leq \frac{1}{N},\ P_1^{\mathcal{P}}(\lambda_{\mathcal{P}}) = 0 \text{ otherwise,} \\
P_2^{\mathcal{P}}(\lambda_{\mathcal{P}}) &= 1 \text{ for } \frac{1}{N} < \lambda_{\mathcal{P}} \leq \frac{2}{N},\ P_1^{\mathcal{P}}(\lambda_{\mathcal{P}}) = 0 \text{ otherwise,} \\
&\vdots \\
P_N^{\mathcal{P}}(\lambda_{\mathcal{P}}) &= 1 \text{ for } \frac{N-1}{N} < \lambda_{\mathcal{P}} \leq 1,\ P_N^{\mathcal{P}}(\lambda_{\mathcal{P}}) = 0 \text{ otherwise.}
\end{aligned}$$

For the probability distribution function we take

$$\mu_\psi(\lambda) = f_\psi^I(\lambda_I) f_\psi^{II}(\lambda_{II}) \cdots,$$

and there is a different function $f_\psi^{\mathcal{P}}(\lambda_{\mathcal{P}})$ associated with each of them. For $\mathcal{P} = \{P_k\}$, we put

$$\begin{aligned}
f_\psi^{\mathcal{P}}(\lambda_{\mathcal{P}}) &= N \langle\psi|P_1|\psi\rangle \text{ for } 0 \leq \lambda_{\mathcal{P}} \leq \frac{1}{N},\ f_\psi^{\mathcal{P}}(\lambda_{\mathcal{P}}) = 0 \text{ otherwise,} \\
f_\psi^{\mathcal{P}}(\lambda_{\mathcal{P}}) &= N \langle\psi|P_2|\psi\rangle \text{ for } \frac{1}{N} < \lambda_{\mathcal{P}} \leq \frac{2}{N},\ f_\psi^{\mathcal{P}}(\lambda_{\mathcal{P}}) = 0 \text{ otherwise,} \\
&\vdots \\
f_\psi^{\mathcal{P}}(\lambda_{\mathcal{P}}) &= N \langle\psi|P_N|\psi\rangle \text{ for } \frac{N-1}{N} < \lambda_{\mathcal{P}} \leq 1,\ f_\psi^{\mathcal{P}}(\lambda_{\mathcal{P}}) = 0 \text{ otherwise.}
\end{aligned}$$

It is easy to confirm that (5,6,7) are satisfied, with $\mu_\psi(\lambda) \geq 0$. The possibility of this sort of hidden variable model was pointed out by van Fraassen. Note that in this model two non-maximal measurements that are in the same operational equivalence class need not be represented similarly.

What is one to think of such a model? It certainly demonstrates that a hidden variable model with a description of measurement based on (5,7) can be constructed. So it demonstrates that the probabilities in operational quantum mechanics can always be understood epistemically – *i.e.,* as a mark of our ignorance – if one so desires. Beyond that most find the model of little interest, and their initial reaction is one of dismissal. It seems a completely uneconomical description of an experiment, and indeed of Nature, as it assumes a new independent variable of each system for every rank 1 PVM. Since there are, in principle, an infinite number of such PVMs, and infinite number of independent variables λ are necessary! An obvious criticism of this approach is that it contradicts, in an emphatic way, "Ockham's Razor," the dictum of William of Ockham (*c.* 1280-1349) that descriptions of nature should be as simple and economical as possible. While generally in science one aims at describing a rich range of phenomena with a few variables, here one simply introduces a new variable for almost every new measurement.

Even deeper than this initial reaction is the sense that such a description of an experiment would never result from any realist theory of quantum mechanics. For a realist theory must do more than just meet the challenge of obtaining agreement with the predictions of operational quantum mechanics. As a theory of fundamental entities, it must provide explanations of what occurs. It is hard to believe than any theory leading to the surfeit of principal variables postulated here could meet any of the "burdens of explanation" identified in the previous chapter.

What of an iPVM model? One of these can be constructed as well, but in a way that might seem even more artificial. The strategy is to take for the principal variables λ the set of rays themselves, what is called the "projective Hilbert space" associated with the Hilbert space of interest for the system. Imagine an integration over these rays, described by $d\psi$, and define a Dirac delta function such that

$$\int \delta(\psi_o - \psi)d\psi = 1,$$

with $\delta(\psi_o - \psi)$ vanishing unless $\psi_o = \psi$. Then we can form on iPVM by taking

$$\begin{aligned} \mu_\psi(\psi') &= \delta(\psi - \psi'), \\ \xi_k(\psi') &= \langle\psi'|P_k|\psi'\rangle \end{aligned}$$

We then see that (8) is satisfied in the form

$$\langle\psi|P_k|\psi\rangle = \int \mu_\psi(\psi')\xi_k(\psi')d\psi'.$$

Here the principal variable is the ray itself. Unlike the Bell-Mermin model, where the ray is part of the ontology as one of the principal variables, in this model the ray constitutes the *entire* ontology of the quantum system. We have a return to "textbook quantum mechanics," and thus to the problems that arose there in trying to turn this into a general realist theory.

4 Local causality and contextuality

This is a good point at which to put aside the construction of various hidden variable models, and consider an interesting contradiction that seems to arise between contextuality and local causality. Consider a system of two spin$-\frac{1}{2}$

systems, and the PVM $\mathcal{P}_I = \{P_A, P_B, P_C\}$, where

$$\begin{aligned} P_A &= P^{(1)}_{-\hat{\mathbf{z}}} \otimes P^{(2)}_{+\hat{\mathbf{x}}}, \\ P_B &= P^{(1)}_{-\hat{\mathbf{z}}} \otimes P^{(2)}_{-\hat{\mathbf{x}}}, \\ P_C &= P^{(1)}_{+\hat{\mathbf{z}}} \otimes \mathcal{I}^{(2)}, \end{aligned}$$

where the superscripts refer to the separate spins, and where $\mathcal{I}^{(2)}$ indicates the identity operator for the second spin. Suppose the spins are separated in space, perhaps after some initial interaction. Then implementing this PVM would require a $z-$oriented Stern-Gerlach device at the location of the first spin, and an $x-$oriented Stern Gerlach device at the location of the second spin. Alternately, one could consider the PVM $\mathcal{P}_{II} = \{P_{A'}, P_{B'}, P_C\}$, where

$$\begin{aligned} P_{A'} &= P^{(1)}_{-\hat{\mathbf{z}}} \otimes P^{(2)}_{+\hat{\mathbf{y}}}, \\ P_{B'} &= P^{(1)}_{-\hat{\mathbf{z}}} \otimes P^{(2)}_{-\hat{\mathbf{y}}}, \\ P_C &= P^{(1)}_{+\hat{\mathbf{z}}} \otimes \mathcal{I}^{(2)}, \end{aligned}$$

which would require a different implementation: a $z-$oriented Stern-Gerlach device at the location of the first spin, and a $y-$oriented Stern-Gerlach device at the location of the second spin. While the two PVMs $\mathcal{P}_I$ and $\mathcal{P}_{II}$ are different than the similarly labeled PVMs describing a spin 1 particle and discussed above, if the same conditions for requiring measurement contextuality here as arose there, we would have in general $P_C^{\mathcal{P}_I}(\lambda_1) \neq P_C^{\mathcal{P}_{II}}(\lambda_1)$, at least for some λ_1. Here we have used the assumption that in a locally realist theory we would expect $P_C^{\mathcal{P}_I}(\lambda_1)$ and $P_C^{\mathcal{P}_{II}}(\lambda_1)$ to depend only on a set of principal variables relevant to the first spin, which we label λ_1.

Since for at least some values of the principal variables we would have $P_C^{\mathcal{P}_I}(\lambda_1) \neq P_C^{\mathcal{P}_{II}}(\lambda_1)$, we would have a violation of local causality: For if at the time the Stern-Gerlach devices were implemented the two spins were in spacelike separated regions, then even if the values of the principal variables λ_1 were known, perhaps through their initial values at a time before the spins separated and some law of state evolution, the result of the Stern-Gerlach device for the first spin would *also* depend on the setting of the Stern-Gerlach device for the second spin.

How could one avoid such a violation of local causality in a hidden variable model? Such a violation would not arise in a PVM-F model, but we have seen that the assumption of that kind of model is generally not tenable. One possibility would be that, although a PVM-F model can indeed not be constructed *in general,* perhaps it can be for specific situations of compound systems where each subsystem is in a spacelike separated region. Even were this not possible, perhaps a iPVM model could be constructed for such compound systems, and the indeterminism that arose there would avoid the violation of local causality indicated above, since the actual result that would be recorded for the measurement on the first spin would not be determined whether the orientation of the Stern-Gerlach device measuring the second spin were specified or not. Are either of these possible?

4.1 Local realism and principal variables

With this as motivation, let us look at the constraints that local realism would place on the nature of principal variables and how they can enter in a hidden variable model that can reproduce the predictions of operational quantum mechanics. We have seen that for measurements in spacelike separated regions a PVM can be constructed that describes the measurements in both regions. In our case if in region 1 (henceforth "Alice") the PVM $\left\{P^{(1)}_{+\hat{\mathbf{a}}}, P^{(1)}_{-\hat{\mathbf{a}}}\right\}$ were implemented and in region 2 (henceforth "Bob") the PVM $\left\{P^{(2)}_{+\hat{\mathbf{b}}}, P^{(2)}_{-\hat{\mathbf{b}}}\right\}$ were implemented, the full PVM would be $\left\{P^{(1)}_{+\hat{\mathbf{a}}} \otimes P^{(2)}_{+\hat{\mathbf{b}}}, P^{(1)}_{+\hat{\mathbf{a}}} \otimes P^{(2)}_{-\hat{\mathbf{b}}}, P^{(1)}_{-\hat{\mathbf{a}}} \otimes P^{(2)}_{+\hat{\mathbf{b}}}, P^{(1)}_{-\hat{\mathbf{a}}} \otimes P^{(2)}_{-\hat{\mathbf{b}}}\right\}$. We want to be able to consider the possibility that Alice and Bob can orient their Stern-Gerlach devices in different directions, and so we introduce $q_{ab}(k,l)$ as the probability in our hidden variable model that, for Alice's setting a and Bob's setting b Alice obtains result $k = (+ \text{ or } -)$ and Bob obtains result $l = (+ \text{ or } -)$. For the hidden variable model to agree with the predictions of operational quantum mechanics we must have, for all a and b and all ψ,

$$\begin{aligned}
\left\langle \psi | P^{(1)}_{+\hat{\mathbf{a}}} \otimes P^{(2)}_{+\hat{\mathbf{b}}} | \psi \right\rangle &= q_{ab}(+,+), \\
\left\langle \psi | P^{(1)}_{+\hat{\mathbf{a}}} \otimes P^{(2)}_{-\hat{\mathbf{b}}} | \psi \right\rangle &= q_{ab}(+,-), \\
\left\langle \psi | P^{(1)}_{-\hat{\mathbf{a}}} \otimes P^{(2)}_{+\hat{\mathbf{b}}} | \psi \right\rangle &= q_{ab}(-,+), \\
\left\langle \psi | P^{(1)}_{-\hat{\mathbf{a}}} \otimes P^{(2)}_{-\hat{\mathbf{b}}} | \psi \right\rangle &= q_{ab}(-,-).
\end{aligned} \tag{9}$$

We now turn to the form $q_{ab}(k,l)$ must take in a local realist theory.

Combining any initial preparation procedure, when the spins were close to each other, with a transformation describing the propagation of the spins to the spacelike regions in which they are to be measured, we consider that the ray appropriate just before the measurements is ψ. At this time, since we are assuming a separable realist theory some of the the principal variables λ (say a set labeled λ_1) should be associated with Alice's spacetime region, and the some of them (say a set labeled λ_2) should be associated with Bob's spacetime region. We take the probability distribution function just before the measurements to be $\mu_\psi(\lambda_1;\lambda_2)$, with

$$\begin{aligned}
\mu_\psi(\lambda_1;\lambda_2) &\geq 0 \quad \text{and} \\
\int \mu_\psi(\lambda_1;\lambda_2) d\lambda_1 d\lambda_2 &= 1.
\end{aligned} \tag{10}$$

For Alice's measurement the result can only depend on λ_1, and not on λ_2 nor on the setting of Bob's device; otherwise local causality would be violated. We take a general stance and assume that knowledge of the values of the principal variables λ_1 only allows a prediction of the probability that "spin up" is the result, and the probability that "spin down" is the result; these probabilities must add to unity. As before, this could arise either because a full description would necessarily involve principal variables of the measuring gadget, and those we average over in constructing the probabilities, or because there is a fundamental indeterminism in how the values of the principal variables lead to the result. We characterize those probabilities by the indicator functions $\xi^{(1)}_{a;+}(\lambda_1)$ and $\xi^{(1)}_{a;-}(\lambda_1)$, where $\xi^{(1)}_{a;k}(\lambda_1) > 0$ and $\xi^{(1)}_{a;+}(\lambda_1) + \xi^{(1)}_{a;-}(\lambda_1) = 1$ for all λ_1.

With a similar characterization of the situation in Bob's spacetime region, the average $q_{ab}(k,l)$ of a host of measurements following the sample preparation procedure characterized by ψ will then involve an average over the probability distribution function $\mu_\psi(\lambda_1, \lambda_2)$, and so in a local realist theory we must have

$$q_{ab}(k,l) = \int \mu_\psi(\lambda_1;\lambda_2)\xi^{(1)}_{a;k}(\lambda_1)\xi^{(2)}_{b;l}(\lambda_2)d\lambda_1 d\lambda_2. \tag{11}$$

Or so it seems. Later we will consider possible generalizations, but for the moment we focus on attempts to get agreement with the predictions (9) of operational quantum mechanics with this kind of hidden variable model, which certainly is consistent with local causality. We call a hidden variable model satisfying (9,10,11) a *local hidden variable model.*

> A *local hidden variable model* for measurements of two spins in spacelike separated regions introduces principal variables λ_1 and λ_2 associated with the two separate regions, a probability distribution functions $\mu_\psi(\lambda_1;\lambda_2) \geq 0$ such that
>
> $$\int \mu_\psi(\lambda_1;\lambda_2)d\lambda_1 d\lambda_2 = 1, \tag{12}$$
>
> and satisfies
>
> $$\left\langle \psi | P^{(1)}_{k\hat{\mathbf{a}}} \otimes P^{(2)}_{l\hat{\mathbf{b}}} | \psi \right\rangle = q_{ab}(k,l) \tag{13}$$
>
> for all rays ψ, and spin direction measurements $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ for the indicated spins, where
>
> $$q_{ab}(k,l) \equiv \int \mu_\psi(\lambda_1;\lambda_2)\xi^{(1)}_{a;k}(\lambda_1)\xi^{(2)}_{b;l}(\lambda_2)d\lambda_1 d\lambda_2, \tag{14}$$
>
> with sets of indicator functions $\left\{\xi^{(1)}_{a;k}(\lambda_1)\right\}$, $\left\{\xi^{(2)}_{b;l}(\lambda_2)\right\}$ for the first and second spin respectively, with $\xi^{(1)}_{a;k}(\lambda_1), \xi^{(2)}_{b;l}(\lambda_2) \geq 0$ and
>
> $$\begin{aligned} \xi^{(1)}_{a;+}(\lambda_1) + \xi^{(1)}_{a;-}(\lambda_1) &= 1, \\ \xi^{(2)}_{b;+}(\lambda_2) + \xi^{(2)}_{b;-}(\lambda_2) &= 1, \end{aligned}$$
>
> for all λ_1 and λ_2.

Combining (9,11), we see that for agreement with operational quantum mechanics we require

$$\left\langle\psi|P^{(1)}_{k\hat{\mathbf{a}}}\otimes P^{(2)}_{l\hat{\mathbf{b}}}|\psi\right\rangle=\int\mu_\psi(\lambda_1;\lambda_2)\xi^{(1)}_{a;k}(\lambda_1)\xi^{(2)}_{b;l}(\lambda_2)d\lambda_1 d\lambda_2. \tag{15}$$

It is interesting to compare (15), where the focus is on local realism rather than issues of noncontextuality, with (8), the requirement for a general iPVM, which is indeterministic but measurement noncontextual. In one sense (15) is less general than (8) because in the former the overall indicator function takes the role of a product, $\xi^{(1)}_{a;k}(\lambda_1)\xi^{(2)}_{b;l}(\lambda_2)$; this of course follows from the assumption of local causality. However, in another sense (15) is more general than (8), since the argument leading to (15) does not presuppose gadget noncontextuality. For example, we need *not* assume that the indicator functions $\left\{\xi^{(1)}_{a;+}(\lambda_1),\xi^{(1)}_{a;-}(\lambda_1)\right\}$ hold for any Stern-Gerlach device that would operationally be associated with the PVM $\left\{P^{(1)}_{+\hat{\mathbf{a}}},P^{(1)}_{-\hat{\mathbf{a}}}\right\}$; in the absence of gadget noncontextuality, they could just be assumed to be the indicator functions relevant for the actual, particular gadget being used. Of course, the issue of partner noncontextuality does not arise for the spins of Alice and Bob, since each is described by a two-dimensional Hilbert space. And $\mu_\psi(\lambda_1;\lambda_2)$ can be taken to describe the particular preparation gadget employed.

A special case of a local hidden variable model would be a deterministic local hidden variable model, where instead of indicator functions we have characteristic functions. Here the appropriate principal variables actually specify with certainty the value of the spin component that will be recorded:

> A *deterministic local hidden variable model* for measurements of two spins in spacelike separated regions introduces principal variables λ_1 and λ_2 associated with the two separate regions, a probability distribution functions $\mu_\psi(\lambda_1;\lambda_2)\geq 0$ such that
>
> $$\int\mu_\psi(\lambda_1;\lambda_2)d\lambda_1 d\lambda_2=1, \tag{16}$$
>
> and satisfies
>
> $$\left\langle\psi|P^{(1)}_{k\hat{\mathbf{a}}}\otimes P^{(2)}_{l\hat{\mathbf{b}}}|\psi\right\rangle=q_{ab}(k,l) \tag{17}$$
>
> for all rays ψ, and spin direction measurements $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ for the indicated spins, where
>
> $$q_{ab}(k,l)\equiv\int\mu_\psi(\lambda_1;\lambda_2)\chi^{(1)}_{a;k}(\lambda_1)\chi^{(2)}_{b;l}(\lambda_2)d\lambda_1 d\lambda_2, \tag{18}$$
>
> with sets characteristic functions $\left\{\chi^{(1)}_{a;k}(\lambda_1)\right\}$, $\left\{\chi^{(2)}_{b;l}(\lambda_2)\right\}$ for the first and second spin respectively, with $\chi^{(1)}_{a;k}(\lambda_1),\chi^{(2)}_{b;l}(\lambda_2)$ each equal

to zero or one, and

$$\begin{aligned}
\chi_{a;+}^{(1)}(\lambda_1)+\chi_{a;-}^{(1)}(\lambda_1) &= 1, \\
\chi_{b;+}^{(2)}(\lambda_2)+\chi_{b;-}^{(2)}(\lambda_2) &= 1,
\end{aligned}$$

for all λ_1 and λ_2.

4.2 A useful lemma

It turns out that if a local hidden variable model can be found that is in agreement with the predictions of operational quantum mechanics, then one can also find a local deterministic hidden variable model that is as well. The second model is somewhat formal, but here is how it can be constructed starting with a local hidden variable model.

We suppose we have a model satisfying (15). Introduce new formal variables $\tilde{\lambda}_{1a}$ and $\tilde{\lambda}_{2b}$ that range from zero to unity. One $\tilde{\lambda}_{1a}$ is introduced for each orientation $\hat{\mathbf{a}}$ of Alice's Stern-Gerlach device, and one $\tilde{\lambda}_{2b}$ for each orientation $\hat{\mathbf{b}}$ of Bob's Stern-Gerlach device. So, again, as in earlier examples, there is a plethora of new hidden variables. For each orientation of Alice's Stern-Gerlach device introduce characteristic functions $\chi_{a;+}^{(1)}(\lambda_1,\tilde{\lambda}_{1a})$ and $\chi_{a;-}^{(1)}(\lambda_1,\tilde{\lambda}_{1a})$, such that

$$\begin{aligned}
\chi_{a;+}^{(1)}(\lambda_1,\tilde{\lambda}_{1a}) &= 1 \\
\text{for } 0 &< \tilde{\lambda}_{1a} < \xi_{a;+}^{(1)}(\lambda_1) \\
\chi_{a;+}^{(1)}(\lambda_1,\tilde{\lambda}_{1a}) &= 0 \\
\text{for } \xi_{a;+}^{(1)}(\lambda_1) &< \tilde{\lambda}_{1a} < 1,
\end{aligned}$$

and $\chi_{a;-}^{(1)}(\lambda_1,\tilde{\lambda}_{1a}) = 1-\chi_{a;+}^{(1)}(\lambda_1,\tilde{\lambda}_{1a})$, and for each orientation of Bob's Stern-Gerlach device we introduce characteristic functions $\chi_{b;+}^{(2)}(\lambda_2,\tilde{\lambda}_{2b})$ and $\chi_{b;-}^{(2)}(\lambda_2,\tilde{\lambda}_{2b})$ in the same way. Now if we denote the original set of principal variables λ_1 for the first spin, together with the whole set $\left\{\tilde{\lambda}_{1a},\tilde{\lambda}_{1a'},\tilde{\lambda}_{1a''}...\right\}$ by $\bar{\lambda}_1$, and the original set of principal variables λ_2 for the second spin, together with the whole set $\left\{\tilde{\lambda}_{2b},\tilde{\lambda}_{2b'},\tilde{\lambda}_{2b''}...\right\}$ by $\bar{\lambda}_2$, and if we introduce a new probability distribution function

$$\bar{\mu}_\psi(\bar{\lambda}_1;\bar{\lambda}_2) \equiv \mu_\psi(\lambda_1;\lambda_2),$$

from (10) we have

$$\int \bar{\mu}_\psi(\bar{\lambda}_1;\bar{\lambda}_2) = 1,$$

and from (11) we have

$$q_{kl}^{ab} = \int \bar{\mu}_\psi(\bar{\lambda}_1;\bar{\lambda}_2)\chi_{a;k}^{(1)}(\bar{\lambda}_1)\chi_{b;l}^{(2)}(\bar{\lambda}_2)d\bar{\lambda}_1 d\bar{\lambda}_2,$$

although of course $\chi_{a;k}^{(1)}(\bar{\lambda}_1)$ really depends only on λ_1 and $\tilde{\lambda}_{1a}$, while $\chi_{b;l}^{(2)}(\bar{\lambda}_2)$ really depends only on λ_2 and $\tilde{\lambda}_{2b}$.

The bottom line is that if we have a local hidden variable model satisfying (12,13,14), and in agreement with operational quantum mechanics, it is possible to derive from it a local hidden variable model involving characteristic functions, and satisfying (16,17,14). This was first pointed out by Fine in 1982; before then it was generally thought that (indeterministic) local hidden variable models were more general than deterministic local hidden variable models. Schematically, we have

$$\begin{array}{c}\text{possibility of local}\\\text{hidden variable model}\\\text{in agreement with}\\\text{operational quantum mechanics}\end{array} \Rightarrow \begin{array}{c}\text{possibility of local}\\\text{deterministic}\\\text{hidden variable model}\\\text{in agreement with}\\\text{operational quantum mechanics}\end{array}.$$

From this it follows that

$$\begin{array}{c}\text{impossibility of local}\\\text{hidden variable model}\\\text{in agreement with}\\\text{operational quantum mechanics}\end{array} \Leftarrow \begin{array}{c}\text{impossibility of local}\\\text{deterministic}\\\text{hidden variable model}\\\text{in agreement with}\\\text{operational quantum mechanics}\end{array}.$$

Since we will see that such *impossibility* does arise, we focus on locally deterministic models, since the algebra of characteristic functions is simpler than that of indicator functions, and so some of the demonstrations can be done easier. Cleaning up the notation, it is useful to denote the full set of principal variables as $\lambda = (\bar{\lambda}_1, \bar{\lambda}_2)$. Although some quantities will depend only on the variables $\bar{\lambda}_1$ associated with the first spin and some only on the variables $\bar{\lambda}_2$, this will not be essential to the demonstrations that follow. So we define a model less restrictive than a local deterministic hidden variable model, which we call simply a *PC hidden variable model,* because it involves the *P*roduct of *C*haracteristic functions $\chi_{a;k}^{(1)}(\lambda)\chi_{b;l}^{(2)}(\lambda)$:

> A *PC hidden variable model* for measurements of two spins in space-like separated regions introduces principal variables λ and a probability distribution functions $\mu_\psi(\lambda) \geq 0$ such that
>
> $$\int \mu_\psi(\lambda) d\lambda = 1, \tag{19}$$
>
> and satisfies
>
> $$\left\langle \psi | P_{k\hat{\mathbf{a}}}^{(1)} \otimes P_{l\hat{\mathbf{b}}}^{(2)} | \psi \right\rangle = q_{ab}(k,l) \tag{20}$$
>
> for all rays ψ, and spin direction measurements $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ for the indicated spins, where
>
> $$q_{ab}(k,l) \equiv \int \mu_\psi(\lambda) \chi_{a;k}^{(1)}(\lambda) \chi_{b;l}^{(2)}(\lambda) d\lambda, \tag{21}$$

with sets characteristic functions $\left\{\chi^{(1)}_{a;k}(\lambda)\right\}$, $\left\{\chi^{(2)}_{b;l}(\lambda)\right\}$ for the first and second spin respectively, with $\chi^{(1)}_{a;k}(\lambda), \chi^{(2)}_{b;l}(\lambda)$ each equal to zero or one and

$$\begin{aligned}\chi^{(1)}_{a;+}(\lambda) + \chi^{(1)}_{a;-}(\lambda) &= 1, \\ \chi^{(2)}_{b;+}(\lambda) + \chi^{(2)}_{b;-}(\lambda) &= 1,\end{aligned}$$

for all λ.

If it impossible to construct a PC hidden variable model in agreement with quantum mechanics, it is impossible to construct a local deterministic hidden variable model in agreement with quantum mechanics, for the latter is a special case of the former. And as we have seen, if it is impossible to construct a local deterministic hidden variable model in agreement with quantum mechanics, it is impossible to construct a local hidden variable model in agreement with quantum mechanics. So the impossibility of a successful PC hidden variable model implies the impossibility of a successful local hidden variable model.

5 Bell's theorem

We now demonstrate that indeed (19,20,21) cannot be generally satisfied for any choice of probability distribution functions and characteristic functions, and thus quantum mechanics will not admit a local hidden variable model. The strategy is to assume that (19,20,21) hold, and then derive a contradiction with the predictions of operational quantum mechanics. Since Bell's original work in 1966 this issue has been looked at from a host of different perspectives, and a number of demonstrations have been presented; the subject would constitute a whole book in itself. Any such demonstration we refer to as a "Bell's theorem," and we just present three examples here.

5.1 Bell's example

The first is Bell's. He looked at $\left\langle\psi|\sigma^{(1)}_{\hat{\mathbf{a}}}\otimes\sigma^{(2)}_{\hat{\mathbf{b}}}|\psi\right\rangle$, where $\sigma^{(1)}_{\hat{\mathbf{a}}} = P^{(1)}_{+\hat{\mathbf{a}}} - P^{(1)}_{-\hat{\mathbf{a}}}$, etc., and $|\psi\rangle$ is the singlet state that can be written as

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|+\hat{\mathbf{n}}\rangle^{(1)}\otimes|-\hat{\mathbf{n}}\rangle^{(2)} - |-\hat{\mathbf{n}}\rangle^{(1)}\otimes|+\hat{\mathbf{n}}\rangle^{(2)}\right)$$

for any $\hat{\mathbf{n}}$. Writing $\left\langle\psi|\sigma^{(1)}_{\hat{\mathbf{a}}}\otimes\sigma^{(2)}_{\hat{\mathbf{b}}}|\psi\right\rangle$ simply as $\left\langle\sigma^{(1)}_{\hat{\mathbf{a}}}\sigma^{(2)}_{\hat{\mathbf{b}}}\right\rangle$, if we assume (20) we would have

$$\left\langle\sigma^{(1)}_{\hat{\mathbf{a}}}\sigma^{(2)}_{\hat{\mathbf{b}}}\right\rangle = \int \mu_\psi(\lambda) f^{(1)}_a(\lambda) f^{(2)}_b(\lambda) d\lambda, \tag{22}$$

where

$$\begin{aligned}f^{(1)}_a(\lambda) &= \chi^{(1)}_{a;+}(\lambda) - \chi^{(1)}_{a;-}(\lambda), \\ f^{(2)}_b(\lambda) &= \chi^{(2)}_{b;+}(\lambda) - \chi^{(2)}_{b;-}(\lambda),\end{aligned}$$

with each function equal to ± 1 in any experimental run. Now from elementary quantum mechanics we have $\left\langle \sigma^{(1)}_{\hat{\mathbf{a}}} \sigma^{(2)}_{\hat{\mathbf{b}}} \right\rangle = -\hat{\mathbf{a}} \cdot \hat{\mathbf{b}}$, and so from (22) we have

$$-1 = \int \mu_\psi(\lambda) f_b^{(1)}(\lambda) f_b^{(2)}(\lambda) d\lambda.$$

Since $\mu_\psi(\lambda)$ is positive and its integral is unity, we must have $f_b^{(1)}(\lambda) f_b^{(2)}(\lambda) = -1$, at least for all λ for which $\mu_\psi(\lambda) \neq 0$, and since the functions $f_b^{(1)}(\lambda)$ and $f_b^{(2)}(\lambda)$ are both either $+1$ or -1 we must have $f_b^{(2)}(\lambda) = -f_b^{(1)}(\lambda)$ for such λ; this allows us to write (22) as

$$\left\langle \sigma^{(1)}_{\hat{\mathbf{a}}} \sigma^{(2)}_{\hat{\mathbf{b}}} \right\rangle = -\int \mu_\psi(\lambda) f_a^{(1)}(\lambda) f_b^{(1)}(\lambda) d\lambda.$$

Thus

$$\begin{aligned} & \left\langle \sigma^{(1)}_{\hat{\mathbf{a}}} \sigma^{(2)}_{\hat{\mathbf{b}}} \right\rangle - \left\langle \sigma^{(1)}_{\hat{\mathbf{a}}} \sigma^{(2)}_{\hat{\mathbf{c}}} \right\rangle \\ = \; & -\int \mu_\psi(\lambda) \left[f_a^{(1)}(\lambda) f_b^{(1)}(\lambda) - f_a^{(1)}(\lambda) f_c^{(1)}(\lambda) \right] d\lambda \\ = \; & \int \mu_\psi(\lambda) \left[-f_a^{(1)}(\lambda) f_b^{(1)}(\lambda) \right] \left[1 - f_b^{(1)}(\lambda) f_c^{(1)}(\lambda) \right] d\lambda, \end{aligned} \tag{23}$$

where in the last line we have used the fact that $\left[f_b^{(1)}(\lambda) \right]^2 = 1$ for all λ. Now $\left[1 - f_b^{(1)}(\lambda) f_c^{(1)}(\lambda) \right]$ is either 0 or 2 for all λ, and therefore always nonnegative. The term $\left[-f_a^{(1)}(\lambda) f_b^{(1)}(\lambda) \right]$ is either ± 1, so an upper bound on the norm of (23) is obtained by replacing $\left[-f_a^{(1)}(\lambda) f_b^{(1)}(\lambda) \right]$ by unity,

$$\begin{aligned} & \left| \left\langle \sigma^{(1)}_{\hat{\mathbf{a}}} \sigma^{(2)}_{\hat{\mathbf{b}}} \right\rangle - \left\langle \sigma^{(1)}_{\hat{\mathbf{a}}} \sigma^{(2)}_{\hat{\mathbf{c}}} \right\rangle \right| \\ \leq \; & \int \mu_\psi(\lambda) \left[1 - f_b^{(1)}(\lambda) f_c^{(1)}(\lambda) \right] d\lambda \end{aligned}$$

Noting that $f_c^{(1)}(\lambda) = -f_c^{(2)}(\lambda)$ by an argument of the form given above, and using the fact that $\mu_\psi(\lambda)$ integrates over λ to unity, we can write this as

$$\left| \left\langle \sigma^{(1)}_{\hat{\mathbf{a}}} \sigma^{(2)}_{\hat{\mathbf{b}}} \right\rangle - \left\langle \sigma^{(1)}_{\hat{\mathbf{a}}} \sigma^{(2)}_{\hat{\mathbf{c}}} \right\rangle \right| \leq 1 + \left\langle \sigma^{(1)}_{\hat{\mathbf{b}}} \sigma^{(2)}_{\hat{\mathbf{c}}} \right\rangle,$$

an inequality that would hold between the expectation values were a local hidden variable model possible. But operational quantum mechanics gives

$$\left| -\hat{\mathbf{a}} \cdot \hat{\mathbf{b}} + \hat{\mathbf{a}} \cdot \hat{\mathbf{c}} \right| \leq 1 - \hat{\mathbf{b}} \cdot \hat{\mathbf{c}},$$

which does not in general hold. Taking $\hat{\mathbf{b}} = \hat{\mathbf{x}}$, $\hat{\mathbf{c}} = \hat{\mathbf{y}}$, and $\hat{\mathbf{a}} = (-\hat{\mathbf{x}} + \hat{\mathbf{y}})/\sqrt{2}$, for example, this would give

$$\sqrt{2} \leq 1,$$

a contradiction. Hence operational quantum mechanics will not admit a local hidden variable model.

5.2 An algebraic version

Bell's demonstration is sometimes called a "statistical" version of Bell's theorem, since it relies on expectation values. Other demonstrations, called "algebraic," rely on the underlying algebra of the operators and the characteristic functions; contradictions follow from the conditions on characteristic functions that arise when the ket is an eigenket of combinations of the operators involved. We present one that involves three spins, assumed to be in spacelike separated regions when the measurements occur. The obvious generalization of (19,20,21) is

$$\begin{aligned} \mu_\psi(\lambda) &\geq 0 \quad \text{and} \\ \int \mu_\psi(\lambda) d\lambda &= 1. \end{aligned} \tag{24}$$

and

$$\begin{aligned} \left\langle \psi | P^{(1)}_{k\hat{\mathbf{a}}} \otimes P^{(2)}_{l\hat{\mathbf{b}}} \otimes P^{(3)}_{m\hat{\mathbf{c}}} | \psi \right\rangle &= \int \mu_\psi(\lambda) \chi^{(1)}_{a;k}(\lambda) \chi^{(2)}_{b;l}(\lambda) \chi^{(3)}_{c;m}(\lambda) d\lambda \\ \chi^{(1)}_{a;k}(\lambda)\chi^{(1)}_{a;l}(\lambda) &= 0 \text{ or } 1 \text{ for all } \lambda, \qquad \chi^{(1)}_{a;+}(\lambda) + \chi^{(1)}_{a;-}(\lambda) = 1 \text{ for all } \lambda, \\ \chi^{(2)}_{b;k}(\lambda)\chi^{(2)}_{b;l}(\lambda) &= 0 \text{ or } 1 \text{ for all } \lambda, \qquad \chi^{(2)}_{b;+}(\lambda) + \chi^{(2)}_{b;-}(\lambda) = 1 \text{ for all } \lambda, \\ \chi^{(3)}_{c;k}(\lambda)\chi^{(3)}_{c;l}(\lambda) &= 0 \text{ or } 1 \text{ for all } \lambda, \qquad \chi^{(3)}_{c;+}(\lambda) + \chi^{(3)}_{c;-}(\lambda) = 1 \text{ for all } \lambda. \end{aligned} \tag{25}$$

The ket employed here is the Greenberger-Horne-Zeilinger ket,

$$\begin{aligned} & |\psi_{GHZ}\rangle \\ = \; & \frac{1}{2} \left(|{+\hat{\mathbf{x}}}\rangle |{+\hat{\mathbf{x}}}\rangle |{+\hat{\mathbf{x}}}\rangle - |{+\hat{\mathbf{x}}}\rangle |{+\hat{\mathbf{y}}}\rangle |{+\hat{\mathbf{y}}}\rangle - |{+\hat{\mathbf{y}}}\rangle |{+\hat{\mathbf{x}}}\rangle |{+\hat{\mathbf{y}}}\rangle - |{+\hat{\mathbf{y}}}\rangle |{+\hat{\mathbf{y}}}\rangle |{+\hat{\mathbf{x}}}\rangle \right). \end{aligned}$$

where in each term on the right-hand-side the first ket refers to the first spin, the second to the second spin, and the third to the third spin; we leave off the labels as well as the direct product sign $\otimes$. The latter we leave off in writing our spin operators as well; the operators of interest have $|\psi_{GHZ}\rangle$ as an eigenket,

$$\begin{aligned} \sigma^{(1)}_{\hat{\mathbf{x}}} \sigma^{(2)}_{\hat{\mathbf{x}}} \sigma^{(3)}_{\hat{\mathbf{x}}} |\psi_{GHZ}\rangle &= -|\psi_{GHZ}\rangle, \\ \sigma^{(1)}_{\hat{\mathbf{y}}} \sigma^{(2)}_{\hat{\mathbf{y}}} \sigma^{(3)}_{\hat{\mathbf{x}}} |\psi_{GHZ}\rangle &= |\psi_{GHZ}\rangle, \\ \sigma^{(1)}_{\hat{\mathbf{y}}} \sigma^{(2)}_{\hat{\mathbf{x}}} \sigma^{(3)}_{\hat{\mathbf{y}}} |\psi_{GHZ}\rangle &= |\psi_{GHZ}\rangle, \\ \sigma^{(1)}_{\hat{\mathbf{x}}} \sigma^{(2)}_{\hat{\mathbf{y}}} \sigma^{(3)}_{\hat{\mathbf{y}}} |\psi_{GHZ}\rangle &= |\psi_{GHZ}\rangle. \end{aligned} \tag{26}$$

From (25) we have

$$\left\langle \psi_{GHZ} | \sigma^{(1)}_{\hat{\mathbf{a}}} \sigma^{(2}_{\hat{\mathbf{b}}} \sigma^{(3)}_{\hat{\mathbf{c}}} | \psi_{GHZ} \right\rangle = \int \mu_{GHZ}(\lambda) f^{(1)}_a(\lambda) f^{(2)}_b(\lambda) f^{(3)}_c(\lambda) d\lambda,$$

as in the development of Bell's inequality, where all the $f's$ here are either ± 1. Now from the first of (26) we have

$$-1 = \int \mu_{GHZ}(\lambda) f^{(1)}_x(\lambda) f^{(2)}_x(\lambda) f^{(3)}_x(\lambda) d\lambda.$$

But since all the $f's$ are ± 1, $\mu_{GHZ}(\lambda)$ is nonnegative, and integrates to unity, it must be that $f_x^{(1)}(\lambda)f_x^{(2)}(\lambda)f_x^{(3)}(\lambda) = -1$ for *all* the λ for which $\mu_{GHZ}(\lambda) \neq 0$. This strong condition follows because $|\psi\rangle$ is an eigenket of the operator $\sigma_{\hat{\mathbf{x}}}^{(1)}\sigma_{\hat{\mathbf{x}}}^{(2}\sigma_{\hat{\mathbf{x}}}^{(3)}$; corresponding results follow for the other operators in (26) of which $|\psi_{GHZ}\rangle$ is an eigenket, and collecting all such results we find

$$\begin{aligned} f_x^{(1)}(\lambda)f_x^{(2)}(\lambda)f_x^{(3)}(\lambda) &= -1, \\ f_x^{(1)}(\lambda)f_y^{(2)}(\lambda)f_y^{(3)}(\lambda) &= 1, \\ f_y^{(1)}(\lambda)f_x^{(2)}(\lambda)f_y^{(3)}(\lambda) &= 1, \\ f_y^{(1)}(\lambda)f_y^{(2)}(\lambda)f_x^{(3)}(\lambda) &= 1. \end{aligned}$$

for *all* the λ for which $\mu_{GHZ}(\lambda) \neq 0$. Multiplying all of these expressions together we find

$$1 = -1,$$

since on the left hand side each f appears twice; thus a contradiction, demonstrating that a local hidden variable theory is not tenable.

5.3 The Clauser-Horne inequalities

As a third example we derive the Clauser-Horne inequalities. They involve two spins, so from (20,21) we want to examine the possibility of satisfying

$$\left\langle \psi | P_{k\hat{\mathbf{a}}}^{(1)} \otimes P_{l\hat{\mathbf{b}}}^{(2)} | \psi \right\rangle = q_{ab}(k,l) \equiv \int \mu_\psi(\lambda)\chi_{a;k}^{(1)}(\lambda)\chi_{b;l}^{(2)}(\lambda)d\lambda \tag{27}$$

where $\left\langle \psi | P_{k\hat{\mathbf{a}}}^{(1)} \otimes P_{l\hat{\mathbf{b}}}^{(2)} | \psi \right\rangle$ is evaluated from operational quantum mechanics. Summing over l, and using the fact that

$$\begin{aligned} \sum_l P_{l\hat{\mathbf{b}}}^{(2)} &= \mathcal{I}^{(2)}, \\ \sum_k \chi_{b;l}^{(2)}(\lambda) &= 1, \end{aligned}$$

this would lead to

$$\left\langle \psi | P_{k\hat{\mathbf{a}}}^{(1)} | \psi \right\rangle = \int \mu_\psi(\lambda)\chi_{a;k}^{(1)}(\lambda)d\lambda \equiv q_a^{(1)}(k) \tag{28}$$

We want to derive conditions that this structure places on the q_{ab}^{kl}, the $q_a^{(1)}(k)$, and the $q_b^{(2)}(l)$. Here is one way to do it. Construct the quantity

$$\begin{aligned} K(\lambda) \equiv{}& \chi_{a;k}^{(1)}(\lambda)\left[\chi_{b;l}^{(1)}(\lambda)(1-\chi_{c;m}^{(2)}(\lambda)) + (1-\chi_{b;l}^{(1)}(\lambda))(1-\chi_{d;n}^{(2)}(\lambda))\right] \\ &+(1-\chi_{a;k}^{(1)}(\lambda))\left[\chi_{b;l}^{(1)}(\lambda)\chi_{d;n}^{(2)}(\lambda) + (1-\chi_{b;l}^{(1)}(\lambda))\chi_{c;m}^{(2)}(\lambda)\right]. \end{aligned}$$

It is easy to see that for any λ, $K(\lambda)$ is either zero or unity. Thus

$$0 \leq \int \mu_\psi(\lambda) K(\lambda) d\lambda \leq 1. \tag{29}$$

With a bit of algebra, using the properties of characteristic functions we find that we can write

$$\begin{aligned}
K(\lambda) &= \chi^{(1)}_{a;k}(\lambda) + \chi^{(2)}_{c;m}(\lambda) \\
&\quad + \chi^{(1)}_{b;l}(\lambda)\chi^{(2)}_{d;n}(\lambda) - \chi^{(1)}_{a;k}(\lambda)\chi^{(2)}_{c;m}(\lambda) \\
&\quad - \chi^{(1)}_{b;l}(\lambda)\chi^{(2)}_{c;m}(\lambda) - \chi^{(1)}_{a;k}(\lambda)\chi^{(2)}_{d;n}(\lambda).
\end{aligned}$$

Using this in (29), together with the definitions (27,28), we find the inequalities

$$0 \leq q^{(1)}_a(k) + q^{(2)}_c(m) + q_{bd}(l,n) - q_{ac}(k,m) - q_{bc}(l,m) - q_{ad}(k,n) \leq 1. \tag{30}$$

If the hidden variable model were successful we would have this satisfied with

$$\begin{aligned}
q^{(1)}_a(k) &= \left\langle \psi | P^{(1)}_{k\hat{\mathbf{a}}} | \psi \right\rangle \\
q^{(2)}_c(m) &= \left\langle \psi | P^{(1)}_{m\hat{\mathbf{c}}} | \psi \right\rangle, \\
q_{ab}(k,l) &= \left\langle \psi | P^{(1)}_{k\hat{\mathbf{a}}} \otimes P^{(2)}_{l\hat{\mathbf{b}}} | \psi \right\rangle.
\end{aligned}$$

Evaluating the quantum mechanical expressions on the right hand sides, and using the results in (30), leads to contradictions (can you identify one?), and so again we see that operational quantum mechanics will not admit a local hidden variable model.

5.4 Joint probability distributions

The three proofs all attest to the cleverness of those who discovered them, but all seem quite different in detail. Is there a single principle that underlies them all, even though it is perhaps not apparent from the derivations given? Indeed there is.

Note first that we can understand

$$q^{(1)}_a(k) = \int \mu_\psi(\lambda)\chi^{(1)}_{a;k}(\lambda) d\lambda \tag{31}$$

as the probability that with setting a the measurement of particle 1 would yield result k ($+$ or $-$) were it performed, and

$$q_{ab}(k,l) = \int \mu_\psi(\lambda)\chi^{(1)}_{a;k}(\lambda)\chi^{(2)}_{b;l}(\lambda) d\lambda \tag{32}$$

as the probability that with setting a the measurement of particle 1 would yield result k were it performed, *and* that with setting b the measurement of particle

2 would yield result l were it performed. Now suppose we restrict ourselves to three possible settings on each measurement gadget, which we label a, b, c. Operationally, of course, only one of these settings can be adopted at each gadget, but we can nonetheless introduce the following quantity:

$$\begin{aligned} & p_{abc;abc}(k,m,n;l,s,u) \\ = & \int \mu_\psi(\lambda)\chi^{(1)}_{a;k}(\lambda)\chi^{(1)}_{b;m}(\lambda)\chi^{(1)}_{c;n}(\lambda)\chi^{(2)}_{a;l}(\lambda)\chi^{(2)}_{b;s}(\lambda)\chi^{(2)}_{c;u}(\lambda)d\lambda. \end{aligned} \tag{33}$$

We formally refer to this as:

> The probability that with setting a the measurement of particle 1 would yield result k were it performed, *and*
>
> that with setting b the measurement of particle 1 would yield result m were it performed, *and*
>
> that with setting c the measurement of particle 1 would yield result n were it performed, *and*
>
> that with setting a the measurement of particle 2 would yield result l were it performed, *and*
>
> that with setting b the measurement of particle 2 would yield result s were it performed, *and*
>
> that with setting c the measurement of particle 2 would yield result u were it performed.

Such a statement makes no sense in operational quantum mechanics, but note that from the form (31,32) the quantities like $q_{ab}(k,l)$ and $q^{(1)}_a(k)$ take in a (here deterministic) local hidden variable theory, the quantity $p_{abc;abc}(k,m,n;l,s,u)$ serves as a joint probability distribution from which $q_{ab}(k,l)$ and $q^{(1)}_a(k)$ can be derived as marginal probability distributions in the usual way. That is, we have

$$\begin{aligned} q^{(1)}_a(k) &= \sum_{m,n,l,s,u} p_{abc;abc}(k,m,n;l,s,u), \\ q_{ab}(k,l) &= \sum_{m,n,s,u} p_{abc;abc}(k,m,n;l,s,u), \end{aligned}$$

and so on. The reasons the constraints that constitute versions of Bell's theorem arise is because certain sets of probabilities and coincidence probabilities cannot arise as the marginals of a joint distribution.

Here is a trivial example. Suppose that instead of our nine possible experiments (three settings for each of two particles) we imagine a simple case of only three possible experiments ($\iota\tau\kappa$), all perhaps even in the same laboratory, of which two are done; we write the alleged joint probability distribution as $p_{\iota\tau\kappa}(i,j,k)$ where i, j, and k label the possible outcomes ($+$ or $-$), and an actual probability distribution function for two performed experiments as, e.g., $p_{\iota\tau}(i,j)$. Now suppose that three coincidence probabilities are $p_{\iota\tau}(+,+) = 1$,

$p_{\iota\kappa}(+,+) = 1$, and $p_{\tau\kappa}(+,+) = 0$. There is *no* joint probability distribution function $p_{\iota\tau\kappa}(i,j,k)$ that will generate these as marginals, because the first two coincidence probabilities require $p_{\iota\tau\kappa}(+,+,+) = 1$, but this would then lead to $p_{\tau\kappa}(+,+) = 1$, not $p_{\tau\kappa}(+,+) = 0$. In other words, the first two coincidence probabilities tell us that the outcomes of ι, τ, and κ must always all be $+$, while the last coincidence probability tells us that the outcomes of τ and κ cannot both be $+$.

Demonstrations of Bell's theorem are more complicated than this, of course, but the derivation of each can be written in such a way that the key element is is the demonstration that coincidence probabilities in local hidden variable theories must be marginals of higher-order distributions, and such marginals always satisfy a set of inequalities. Thus, to derive new Bell inequalities, one need only imagine an experiment with a greater number of regions, or measurements, or outcomes. This will correspond to a greater set of possible coincidences, and the probabilities for these coincidences, by virtue of being marginals of a higher-order distribution, will necessarily be required to satisfy a set of inequalities. It has been shown that the problem of determining all the inequalities that must be satisfied by a given set of marginals is a problem which is known to be computationally difficult. Nonetheless, if these inequalities can be derived, then one has another example of Bell's theorem.

In any event, one needs only a single counterexample to disprove a hypothesis. The proofs of Bell's theorem we have provided are therefore each sufficient to demonstrate that the predictions of operational quantum mechanics cannot be reproduced by a local hidden variable model.

6 Superdeterminism

So can we then say that the predictions of operational quantum mechanics are not consistent with a local realist theory? Not quite. Despite the fact that the two spins we have been considering are measured in spacelike separate regions, the backward light cones of the spacelike regions where the orientation of the Stern-Gerlach devices of the two measurements were set, *and* the backward light cone of the region of spacetime where the original spin preparation occurred, do overlap at an early enough time. And so one can imagine a number of variables – label them Λ_0, Λ_1, and Λ_2 – that could affect the very original preparation performed, as well as the orientation of the Stern-Gerlach devices. We suppose that Λ_0, Λ_1, and Λ_2 are specified at this early enough time and do not change with time, with Λ_0 later associated with the space time region of the original preparation, Λ_1 with the region of spacetime where the first spin's measurement device is set up, and Λ_2 with the region of spacetime where the second spin's measurement device is set up.

Consider first the preparation, which has so far been characterized by the ray ψ from operational quantum mechanics. Suppose that its details also depended on a variable Λ_0; then in a deterministic local hidden variable model the preparation could be denoted by $\mu_\psi(\lambda_1,\lambda_2;\Lambda_0)$. For each Λ_0 we still imagine

$$\int \mu_\psi(\lambda_1;\lambda_2;\Lambda_0)d\lambda_1 d\lambda_2 = 1,$$

but the form of the dependence on λ_1 and λ_2 will be different for different Λ_0.

Turning to the measurements, for the first spin we denote the appropriate characteristic functions by $\left\{\chi_k^{(1)}(\lambda_1,\Lambda_1)\right\}$. For some value of Λ_1 we would then have $\left\{\chi_k^{(1)}(\lambda_1,\Lambda_1)=\chi_{a;k}^{(1)}(\lambda_1)\right\}$, for others $\left\{\chi_k^{(1)}(\lambda_1,\Lambda_1)=\chi_{b;k}^{(1)}(\lambda_1)\right\}$, and so on. So while Alice might think that the choice of Stern-Gerlach orientation has been made by some nominally random procedure constructed in the lab, or simply by her own "free" choice, in fact it is determined by Λ_1. And so also for Bob. Now suppose $\Lambda_0=\Lambda_0(\Lambda_1,\Lambda_2)$; this is possible, because the values Λ_0, Λ_1, and Λ_2 are all specified in the same region of spacetime. Then one could imagine Λ_0 such that the distribution function $\mu_\psi(\lambda_1,\lambda_2;\Lambda_0(\Lambda_1,\Lambda_2))$ is adjusted to give the correct statistics for the measurements that *will* be done, as specified by Λ_1 and Λ_2. Since there is only one pair of measurements that need be considered, this can be easily done. And so for any $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ Stern-Gerlach orientations of Alice and Bob, if $\Lambda_1(\hat{\mathbf{a}})$ is the value of Λ_1 that leads to the orientation $\hat{\mathbf{a}}$ for Alice's Stern-Gerlach device, and $\Lambda_2(\hat{\mathbf{b}})$ is the value of Λ_2 that leads to the orientation $\hat{\mathbf{b}}$ for Bob's Stern-Gerlach device, we can easily arrange to have

$$\left\langle\psi|P_{k\hat{\mathbf{a}}}^{(1)}\otimes P_{l\hat{\mathbf{b}}}^{(2)}|\psi\right\rangle=\int \mu_\psi(\lambda_1;\lambda_2;\Lambda_0(\Lambda_1(\hat{\mathbf{a}}),\Lambda_2(\hat{\mathbf{b}}))\chi_k^{(1)}(\lambda_1;\Lambda_1(\hat{\mathbf{a}}))\chi_l^{(2)}(\lambda_2;\Lambda_2(\hat{\mathbf{b}}))d\lambda_1 d\lambda_2$$

for all $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, and thus identify a local realist model that is in agreement with the predictions of operational quantum mechanics.

What is one to make of this? An experimenter's claim of his or her "free will" in choosing the setting of a measurement device would of course obviate this kind of scenario, but the justification of claims of "free will" is in the best of cases problematic. Perhaps more significant is the fact that there is no precedent for any kind of scenario of the sort that would involve the variables Λ_0, Λ_1, and Λ_2 that we have identified; they would provide a correlation between the setting of measurement devices and the form of the probability distribution function characterizing the system being measured. Certainly we have no indication of any such kind of variables in our everyday expenience. Yet that everyday experience is of course classical, and so one might argue that it should not be given too much weight in exploring the quantum and sub-quantum regimes. Nonetheless, it does seem that such a scenario would involve Nature conspiring to confuse us, arranging in each and every instance for us to record statistics that seem to contradict what any local realist theory would produce by correlating preparation and measurement devices in a locally deterministic way. This kind of conspiratorial maintenance of local determinism is sometimes called *superdeterminism.*

There is apparently no way to rule it out. Thus, the best we can say is that in the absence of superdeterministic phenomena, Bell's theorem shows that the predictions of operational quantum mechanics are in contradiction with any local realist theory.

This result is a consequence of the equations of quantum mechanics, of course, and has nothing to do with experiment. However, one might imagine a local realist theory (without superdeterminism) with predictions close to those of operational quantum mechanics, but of course in some instances necessarily different. Experiment *can* distinguish between such a putative theory and quantum mechanics, and indeed recent experiments have done just that, confirming the predictions of operational quantum mechanics and thus ruling out an underlying local realist theory, subject only to the additional assumption that there is no superdeterminism.

7 More on contextuality

We conclude this chapter by returning to the issue contextuality. Recall that we had shown the possibility of constructing mPVM-F models, which are at least gadget noncontextual with respect to maximal PVMs, but involves partner contextuality. We also have the possibility of constructing iPVM models, which are still fully noncontextual. Despite the possibility of these models, it is hard to identify what would be an underlying theory that would lead to the predictions of operational quantum mechanics through the form of the probability distribution functions and characteristic (or indicator) functions that they employ.

Both mPVM-F and iPVM models involve pure preparations and PVMs. Such a framework would be reasonable for a realist who felt that the generalization to mixed preparations and POVMs in operational quantum mechanics obscure the deeper physics, since in all known cases those mixed preparations and POVMs can be associated with pure preparations and PVMs in larger Hilbert spaces that seem to have physical significance.

However, from a vantage point that is more sympathetic to modern operational quantum mechanics, recent work on contextuality has focused on mixed preparations and POVMs. Here striking results have been found. Inequalities similar to those of Bell have also been derived from the assumption of universal contextuality (ie., preparation and measurement noncontextuality), and these are of the form subject to experimental test. Thus one is in the position of being able to distinguish between the predictions of operational quantum mechanics, and those of a similar but necessarily different realist theory that would lead to a model satisfying universal noncontextuality.

Yet a new and different perspective on issues of contextuality arises from Bell's theorem and the results that favor the predictions of operational quantum mechanics rather than those of a local realist theory. The experiments demonstrate that, if one dismisses the possibility of superdeterminism, any realist theory in agreement with the predictions of operational quantum mechanics, and with what has been observed in the laboratory, must not be a local realist theory; we call such a putative theory a *nonlocal realist theory.* In such a theory the distinction between "preparation" and "measurement" seems artificial. We typically think of them as separate tasks, and identify different variables – hidden or otherwise – with them, because they implemented by separate gadgets sometimes separated some distance from each other. But in a nonlocal realist theory this carries no weight, and it is hard to see the justification for introducing principal variables with "just" the system, for example. Why would the statistics of the preparation not be allowed to depend on the principal variables that one might normally associate with the measurement device? This perspective would seem to argue for a new investigation treating preparation and measurement as indivisible components of simply an "experiment," with any hidden variables identified with the whole experiment and not just with particular components.

Finally, other realists might tire of only the kind of establishment of, or proof-of-impossibility-of-establishing, "models" of the type we have considered here. The real business of physics is presumably the construction of theories, and a realist might demand that this real business should be addressed directly. We turn next to where that can lead.

8 Appendix

The proof presented here is due to Friedberg. We consider the possibility of constructing a function $f(\hat{\mathbf{n}})$, where boldface indicates a unit vector, satisfying the *triad condition:* For any three orthogonal directions $(\hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2, \hat{\mathbf{n}}_3)$, we have $f(\hat{\mathbf{n}}_i) = 0$ for two of the $\hat{\mathbf{n}}_i$, while for the third we have $f(\hat{\mathbf{n}}) = 1$. Satisfying the triad condition is equivalent to being able to paint the surface of a sphere with two colours, say red and blue, such that for each three orthogonal directions $\hat{\mathbf{n}}_i$ two point toward blue, and the third toward red.

In fact it is impossible to construct a function that satisfies the triad condition; the proof is by contradiction.

- Assume that it is possible to find a function $f(\hat{\mathbf{n}})$ that satisfies the triad condition, and consider three sets of mutually orthogonal unit vectors:

$$\begin{aligned}
(\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \hat{\mathbf{a}}_3) &= (\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}) \\
(\hat{\mathbf{b}}_1, \hat{\mathbf{b}}_2, \hat{\mathbf{b}}_3) &= (\frac{\hat{\mathbf{x}} + \hat{\mathbf{y}}}{\sqrt{2}}, \frac{\hat{\mathbf{x}} - \hat{\mathbf{y}}}{\sqrt{2}}, \hat{\mathbf{z}}) \\
(\hat{\mathbf{c}}_1, \hat{\mathbf{c}}_2, \hat{\mathbf{c}}_3) &= (\frac{\hat{\mathbf{x}} + \hat{\mathbf{z}}}{\sqrt{2}}, \frac{\hat{\mathbf{x}} - \hat{\mathbf{z}}}{\sqrt{2}}, \hat{\mathbf{y}}).
\end{aligned}$$

- Now note $f(\hat{\mathbf{b}}_1), f(\hat{\mathbf{b}}_2), f(\hat{\mathbf{c}}_1), f(\hat{\mathbf{c}}_2)$ cannot all be zero. If they were, $f(\hat{\mathbf{b}}_3)$ and $f(\hat{\mathbf{c}}_3)$ would both be 1, and we would get a contradiction from the first triad.

- Next, note that each of $(\hat{\mathbf{b}}_1, \hat{\mathbf{b}}_2, \hat{\mathbf{c}}_1, \hat{\mathbf{c}}_2)$ is orthogonal to one of

$$\hat{\mathbf{h}}_1 = \frac{\hat{\mathbf{y}} + \hat{\mathbf{z}} + \hat{\mathbf{x}}}{\sqrt{3}},$$

$$\hat{\mathbf{h}}_2 = \frac{\hat{\mathbf{y}} + \hat{\mathbf{z}} - \hat{\mathbf{x}}}{\sqrt{3}}.$$

Thus, since at least one of $f(\hat{\mathbf{b}}_1), f(\hat{\mathbf{b}}_2), f(\hat{\mathbf{c}}_1), f(\hat{\mathbf{c}}_2)$ must be unity, $f(\hat{\mathbf{h}}_1)$ and $f(\hat{\mathbf{h}}_2)$ cannot both be unity. For if they were, then the vector from the set $(\hat{\mathbf{b}}_1, \hat{\mathbf{b}}_2, \hat{\mathbf{b}}_3, \hat{\mathbf{b}}_4)$ for which f was unity, together with vector $\hat{\mathbf{h}}_i$ to which it was orthogonal, would constitute two orthogonal vectors for which $f(\hat{\mathbf{n}})$ was unity, violating the triad condition. Now since

$$\hat{\mathbf{h}}_1 \cdot \hat{\mathbf{h}}_2 = \frac{1}{3},$$

if we put $\alpha = \cos^{-1}(\frac{1}{3}) \approx 70\frac{1}{2}^o$, we see by rotational symmetry that if the angle between any two unit vectors $\hat{\mathbf{h}}_a$ and $\hat{\mathbf{h}}_b$ is α, then $f(\hat{\mathbf{h}}_a)$ and $f(\hat{\mathbf{h}}_b)$ cannot both be unity. Call this the α *condition*.

- Next construct the vector $\hat{\mathbf{h}}_3$ in the (xy) plane:

  $$\hat{\mathbf{h}}_3 = \hat{\mathbf{y}} \cos\alpha + \hat{\mathbf{x}} \sin\alpha.$$

  Now assume $f(\hat{\mathbf{h}}_3) = 1$; then $f(\hat{\mathbf{y}}) = 0$ by the α condition, because by construction the angle between $\hat{\mathbf{h}}_3$ and $\hat{\mathbf{y}}$ is α. But then $f(\hat{\mathbf{z}}) = 0$ also, because $\hat{\mathbf{h}}_3$ and $\hat{\mathbf{z}}$ form two orthogonal vectors in a triad. And then since $f(\hat{\mathbf{y}}) = f(\mathbf{z}) = 0$, we have that $f(\hat{\mathbf{x}}) = 1$ by the triad condition. Note now that the angle β between $\hat{\mathbf{h}}_3$ and $\hat{\mathbf{x}}$ is $90^o - \alpha \approx 19\frac{1}{2}^o$. So by rotational symmetry we see that if the angle between and two vectors $\hat{\mathbf{h}}_a$ and $\hat{\mathbf{h}}_c$ is β, then $f(\hat{\mathbf{h}}_a)$ and $f(\hat{\mathbf{h}}_c)$ must both be unity. Call this the β *condition*.

- Finally, identify a vector $\hat{\mathbf{h}}_4$ for which $f(\hat{\mathbf{h}}_4) = 1$. Consider a vector $\hat{\mathbf{h}}_5$ that makes an angle α with $\hat{\mathbf{h}}_4$. Then, by the α condition, $f(\hat{\mathbf{h}}_5) = 0$. But it is also possible to get from $\hat{\mathbf{h}}_4$ to $\hat{\mathbf{h}}_5$ by an odd number of rotations of β; indeed, less than 5 will suffice. Hence, by repeated application of the β condition, we must have $f(\hat{\mathbf{h}}_5) = 1$.

- So we have derived a contradiction from assuming the triad condition holds for $f(\hat{\mathbf{n}})$; therefore no function satisfying the triad condition exists.