

Chapter 9

Three No-Go Theorems and the Measurement Problem

If one is convinced by some EPR-like argument, then one might become a (possibly local) strong property realist. Indeed, even if one is not convinced by EPR, one might still believe that, in spite of the orthodox interpretation, there is no interesting link between strong property realism and the formalism of quantum mechanics. For example, one might hold that it is possible to assign definite eigenvalues to all observables in all states, in other words, to be a *strong* property realist in our sense, so that these values satisfy quantum mechanical predictions. However, in this chapter we consider three results that prove it is not so. None of them affect weak property realism. Finally, we introduce one of the central issues in discussions about quantum mechanics, namely, the measurement problem.

9.1 Some unsatisfactory no-go arguments

There are some quantum mechanical arguments that are occasionally presented as clear evidence that property realism is false. We look at two of them before we consider three significant no-go theorems.

Let us cast our minds back the distinction between superpositions and mixed states. If we assume EE (actually, all is needed is half EE, the eigenvalue to eigenstate link), we cannot say that in a system in state (6.5.2) definitely $O = \lambda_1$ or definitely $O = \lambda_2$, otherwise we would have to interpret (6.5.2) as saying that the system is definitely represented by $|\Psi_1\rangle$ or definitely by $|\Psi_2\rangle$, which would lead to the wrong statistics, as we saw. So, if one adopts EE plus completeness, the essential difference between

superpositions and mixed states shows that property realism is false. The problem, of course, is that one may reject EE or the claim that quantum mechanics is complete.¹

There is a most famous alleged no-go theorem produced in 1932 by von Neumann “proving” that no hidden variable theory allowing simultaneous values of incompatible observables could be constructed that would be compatible with quantum mechanical predictions or, for us, that strong property realism is incompatible with the machinery of quantum mechanics. The argument is complex, and we need not follow it but by noting that it rests on the assumption that, independently of whether \hat{O} and \hat{P} commute or not, if

$$\hat{O} + \hat{P} = \hat{Q}, \tag{9.1.1}$$

then

$$v(O) + v(P) = v(Q), \tag{9.1.2}$$

where $v(O)$ is the pre-measured value of O equal to the relevant eigenvalue of \hat{O} , and the same with the other two observables. Von Neumann then showed that a hidden variable theory cannot satisfy (9.1.2).

The rationale for the transition from (9.1.1) to (9.1.2) was provided by the theorem that in general if (9.1.1) is true, then

$$\langle O \rangle + \langle P \rangle = \langle Q \rangle \tag{9.1.3}$$

for all states, and independently of whether the operators commute or not. The problem here is that (9.1.3) does not yield (9.1.2). Obviously, the latter entails the former, but the converse is not true: the relations holding among averages need not hold among the items

¹ Arguments appealing to interference effects, such as those arising in the two slits experiments or in Mach-Zender interferometers share this pattern.

out of which the averages are constructed. In other words, a hidden variable theory could break (9.1.2) while satisfying (9.1.3). As it happens, von Neumann's claim can be obtained from quantum mechanical formalism *only* for commuting operators. As an aside, we should note that (9.1.2) makes no sense for standard quantum mechanics or for weak property realism if $[\hat{O}, \hat{P}] \neq 0$, for in those views no incompatible observables have simultaneous values.

9.2 Bell's Inequality

In 1964, J.S. Bell discovered a remarkable relation among strong property realism, the principle of locality (PL), and quantum mechanical predictions. There are various ways of proving the inequality, or its equivalent, that constitutes the core of Bell's original theorem; some are due to Bell himself and some to others, but often they are all referred to as "Bell's Theorem". Here we look at a particularly simple and general version that does not employ the machinery of hidden variables, and therefore makes no assumptions about the probability distribution of observables (Clauser, J. F., Horne, M. A., Shimony, A., Holt, R. A., (1969)). Bell's Theorem proves that the combination of strong property realism and the principle of locality (PL) is statistically incompatible with the predictions of quantum mechanics. In other words, any local hidden variable theory capturing strong property realism is statistically incompatible with quantum mechanics.²

As we saw in the EPRB example, when the system is in the singlet state

$\frac{1}{\sqrt{2}} \left(\left| \uparrow_z^a \downarrow_z^b \right\rangle - \left| \downarrow_z^a \uparrow_z^b \right\rangle \right)$, if we measure $S_z^a = +1$, then certainly $S_z^b = -1$. Suppose now that the

² Actually, the theorem applies to strong property contextualism as well, as we shall see.

Stern-Gerlach device A we use on a can be rotated along two unit vectors \mathbf{m} and \mathbf{m}' , and the Stern-Gerlach device B we use on b can be rotated along two unit vectors \mathbf{n} and \mathbf{n}' .

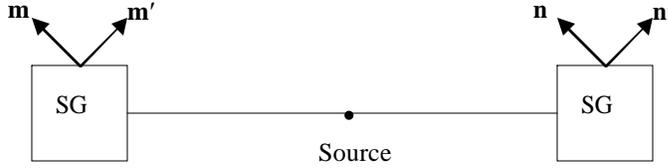


Figure 1

Let us symbolize the component of spin of particle a along vector \mathbf{q} as S_q^a , and similarly for other particles and vectors. Let us also assume that strong property realism obtains, which entails that the values of the spin components of a spin-half particle are always ± 1 ; in addition, PL guarantees that measurement returns on a particle are independent of any measurement return on the other particle (remote outcome independence) or experimental setup on the other particle (remote context independence or parameter independence).³

Consider the quantity

$$\mu = S_m^a S_n^b + S_m^a S_{n'}^b + S_{m'}^a S_n^b - S_{m'}^a S_{n'}^b, \quad (9.2.1)$$

³ The combination of remote outcome independence and remote context independence is equivalent to what is at times referred to as ‘Bell locality’, namely, that upon the joint measurement of two observables, one on a and one on b , the probability for a pair of measurement returns is equal to the product of the probabilities for each measurement return separately (Jarrett, J., P., 1984). Bell’s locality is what is needed for Bell’s proof to take off. Hence, the denial of either remote outcome independence or remote context independence blocks the proof, as we shall soon see.

where each factor in the summands represents a property with a definite value. As PL entails Bell's locality by guaranteeing that S_m^a and $S_{m'}^a$ are independent of any measurement return or experimental apparatus on b , we obtain

$$\mu = S_m^a(S_n^b + S_{n'}^b) + S_{m'}^a(S_n^b - S_{n'}^b). \quad (9.2.2)$$

If we assume that both S_n^b and $S_{n'}^b$ have simultaneous values, they must have either the same sign (that is, both must be equal to ± 1) or opposite signs, one and only one of the two addenda in the previous equation must be equal to zero, and consequently

$$\mu = \pm 2. \quad (9.2.3)$$

Applying (9.2.1) to an ensemble of pairs of particles, and taking the averages, we obtain

$$\langle \mu \rangle = \langle S_m^a S_n^b \rangle + \langle S_m^a S_{n'}^b \rangle + \langle S_{m'}^a S_n^b \rangle - \langle S_{m'}^a S_{n'}^b \rangle. \quad (9.2.4)$$

Now, since each μ is equal to ± 2 , it must be the case that

$$|\langle \mu \rangle| \leq 2. \quad (9.2.5)$$

Consequently,

$$\left| \langle S_m^a S_n^b \rangle + \langle S_m^a S_{n'}^b \rangle + \langle S_{m'}^a S_n^b \rangle - \langle S_{m'}^a S_{n'}^b \rangle \right| \leq 2, \quad (9.2.6)$$

which is Bell's inequality. Before we get to the issue of whether quantum mechanical predictions satisfy Bell's inequality, three things should be noticed.

First, the proof is compatible with property contextualism in the sense that the returns of a measurement may, although it need not, be a function not only of the measured system but also of the measuring apparatus. Second, $S_m^a S_n^b$, $S_m^a S_{n'}^b$, $S_{m'}^a S_n^b$, $S_{m'}^a S_{n'}^b$ are mutually compatible, and therefore (9.2.1) is uncontroversial. However, the transition from (9.2.1) to (9.2.2) appeals to strong property realism (or strong property contextualism) by assuming that both S_n^b and $S_{n'}^b$ have values even if they are

incompatible observables. Third, the transition from (9.2.1) to (9.2.6) does *not* involve any quantum mechanical machinery. This means that Bell's inequality can be proved with respect to any run of the mill dichotomic quality (one that only takes values ± 1) satisfying Bell's locality. Since any physical quantity can be made dichotomic by a bit of tinkering (for example, if length l is more than $3m$, then $l=1$, otherwise $l=0$), Bell's inequality can be obtained in any domain in which Bell's locality holds.

The truly amazing thing is that Bell's inequality is statistically incompatible with quantum mechanical predictions. The two next sections are devoted to proving this.

9.3 Stern-Gerlach Devices Oriented at an Angle

Up to now, we have studied spin by using the standard orthogonal components S_x , S_y , and S_z . To that effect, we have determined their operators and the eigenvalues and eigenvectors of those operators. However, nothing prevents us from finding operators representing any arbitrary component of spin. For example, consider the unit vector \mathbf{n} lying in the xz -plane and making an angle ϕ with the z -axis. Suppose now that we want to determine \hat{S}_n , the operator corresponding to S_n , the spin component along the n -axis in order to predict what might happen to particles traveling along the y -axis and entering a Stern-Gerlach device tilted at an angle ϕ with the z -axis. The normalized eigenvector associated with eigenvalue $-\hbar/2$ is

$$|\downarrow_n\rangle = \begin{pmatrix} -\sin\frac{\phi}{2} \\ \cos\frac{\phi}{2} \end{pmatrix}; \quad (9.3.1)$$

the eigenvector associated with eigenvalue $\hbar/2$ is

$$|\uparrow_n\rangle = \begin{pmatrix} \cos \frac{\phi}{2} \\ \sin \frac{\phi}{2} \end{pmatrix}.^4 \quad (9.3.2)$$

Suppose now that a stream of spin-half particles in state $|\uparrow_z\rangle$ enters a SGN device, a Stern-Gerlach apparatus tilted by an angle ϕ with the z -axis. Let us expand the vector state of the entering particles in terms of the eigenvectors associated with \hat{S}_n , thus obtaining

$$|\uparrow_z\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = c_1 \begin{pmatrix} -\sin \frac{\phi}{2} \\ \cos \frac{\phi}{2} \end{pmatrix} + c_2 \begin{pmatrix} \cos \frac{\phi}{2} \\ \sin \frac{\phi}{2} \end{pmatrix}.^5 \quad (9.3.3)$$

By simple inspection we note that $c_1 = -\sin \frac{\phi}{2}$ and $c_2 = \cos \frac{\phi}{2}$.⁶ Consequently, two particle streams will come out of the SGN device. The first stream is composed of particles in state $|\downarrow_n\rangle$ and with spin $S_n = -\hbar/2$. The probability that a particle belongs to this stream is $|c_1|^2 = \left| -\sin \frac{\phi}{2} \right|^2 = \sin^2 \frac{\phi}{2}$. The second stream is composed of particles in state $|\uparrow_n\rangle$ and with spin $S_n = \hbar/2$. The probability that a particle belongs to this stream is $|c_2|^2 = \cos^2 \frac{\phi}{2}$.

⁴ The proof is left as an exercise.

⁵ Of course, we know that we can expand the state of the entering particles in terms of the eigenvectors associated with \hat{S}_n because \hat{S}_n is Hermitian, and therefore its eigenvectors span the space.

⁶ This is because $\sin^2 x + \cos^2 x = 1$.

9.4 Bell's Inequality and Quantum Mechanics

We can now show that Bell's inequality is statistically incompatible with quantum mechanics. We choose our coordinates so that the z -axis is along the unit vector \mathbf{m} .

Suppose now that after having obtained $S_m^a = +1$ we measure a component S_n^b diverging from \mathbf{m} by an angle ϕ_{mn} and lying in the xz -plane. Since a and b are in the singlet state,

when $S_m^a = +1$ and a 's state is reduced to $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, b 's is $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. As usual, we need to expand

b 's state vector into the eigenvectors of S_m^b , thus obtaining

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \sin \frac{\phi_{mn}}{2} \begin{pmatrix} \cos \frac{\phi_{mn}}{2} \\ \sin \frac{\phi_{mn}}{2} \end{pmatrix} + \cos \frac{\phi_{mn}}{2} \begin{pmatrix} -\sin \frac{\phi_{mn}}{2} \\ \cos \frac{\phi_{mn}}{2} \end{pmatrix}. \quad (9.4.1)$$

Consequently, given that $S_m^a = +1$, we shall obtain $S_n^b = +1$ with probability $\sin^2 \frac{\phi_{mn}}{2}$ and

$S_n^b = -1$ with probability $\cos^2 \frac{\phi_{mn}}{2}$. Similarly, given that $S_m^a = -1$, we shall obtain

$S_n^b = +1$ with probability $\cos^2 \frac{\phi_{mn}}{2}$, and $S_n^b = -1$ with probability $\sin^2 \frac{\phi_{mn}}{2}$.

Now, the probability that $S_m^a = +1$ and $S_n^b = +1$ is equal to

$$\Pr(S_m^a = +1 \cap S_n^b = +1) = \frac{1}{2} \sin^2 \frac{\phi_{mn}}{2}, \quad (9.4.2)$$

since the probability that $S_m^a = +1$ is $1/2$ and the probability that $S_n^b = +1$ given that

$S_m^a = +1$ is

$$\Pr(S_n^b = +1 | S_m^a = +1) = \sin^2 \frac{\phi_{mn}}{2}. \quad (9.4.3)$$

The probabilities of the remaining three combinations can be calculated similarly. We can tabulate the results as follows:

$$\begin{aligned} P(S_m^a = +1 \cap S_n^b = +1) &= \frac{1}{2} \sin^2 \frac{\phi_{mn}}{2} \\ P(S_m^a = +1 \cap S_n^b = -1) &= \frac{1}{2} \cos^2 \frac{\phi_{mn}}{2} \\ P(S_m^a = -1 \cap S_n^b = -1) &= \frac{1}{2} \sin^2 \frac{\phi_{mn}}{2} \\ P(S_m^a = -1 \cap S_n^b = +1) &= \frac{1}{2} \cos^2 \frac{\phi_{mn}}{2}. \end{aligned} \quad (9.4.4)$$

Let $S_m^a S_n^b$ be the product of the returns of the measurements of S_m^a and S_n^b . Then, (9.4.4)

tells us that the probability for such a return is $\frac{1}{2} \cos^2 \frac{\phi_{mn}}{2}$. Consequently, remembering

that $\langle j \rangle = \sum_{j=0}^{\infty} jP(j)$, the expectation value of $S_m^a S_n^b$ is

$$\begin{aligned} \langle S_m^a S_n^b \rangle &= 1 \left(\frac{1}{2} \sin^2 \frac{\phi_{mn}}{2} \right) - 1 \left(\frac{1}{2} \cos^2 \frac{\phi_{mn}}{2} \right) + 1 \left(\frac{1}{2} \sin^2 \frac{\phi_{mn}}{2} \right) - 1 \left(\frac{1}{2} \cos^2 \frac{\phi_{mn}}{2} \right) = \\ &= \sin^2 \frac{\phi_{mn}}{2} - \cos^2 \frac{\phi_{mn}}{2}. \end{aligned} \quad (9.4.5)$$

A quick look at trigonometric tables tells us that

$$\langle S_m^a S_n^b \rangle = -\cos \phi_{mn}. \quad (9.4.6)$$

⁷ Here we use the notation where “ $p \cap q$ ” stands for “ p and q ” and “ $p | q$ ” for “ p given q .” Suppose we are dealing with a well-shuffled deck of cards, p =drawing an ace, q =drawing a queen, and r =drawing the ace of spades. Then, $P(p) = P(q) = 4/52$ and $P(r) = 1/52$. However, $P(r | p) = 1/4$; that is, *given that* we got an ace, the probability that it is the ace of spades is $1/4$. In general, $P(m \cap n) = P(m) \cdot \Pr(n | m)$.

Similarly, we can obtain

$$\begin{aligned}\langle S_m^a S_{n'}^b \rangle &= -\cos\phi_{mn'} \\ \langle S_m^a S_n^b \rangle &= -\cos\phi_{m'n} \\ \langle S_{m'}^a S_{n'}^b \rangle &= -\cos\phi_{m'n'}.\end{aligned}\tag{9.4.7}$$

Now let us take \mathbf{m} and \mathbf{n} to be parallel and $\phi_{m'n} = \phi_{mn'} = \theta$, so that $\phi_{m'n'} = 2\theta$.

Then,

$$\gamma = \left| \langle S_m^a S_n^b \rangle - \langle S_m^a S_{n'}^b \rangle + \langle S_{m'}^a S_n^b \rangle + \langle S_{m'}^a S_{n'}^b \rangle \right| = |1 + 2\cos\theta - \cos 2\theta|.\tag{9.4.8}$$

Hence, Bell's inequality is satisfied just in case

$$|1 + 2\cos\theta - \cos 2\theta| \leq 2.\tag{9.4.9}$$

However, when $0 < \theta < \pi/2$, the inequality is *not* satisfied by quantum mechanics, as one can easily verify. Consequently, local strong property realism is statistically incompatible with quantum mechanics.

Several experiments, largely using photons, have been performed to determine whether Bell's inequality is always satisfied in the quantum world, and the consensus is that they have come down on the side of quantum mechanics. Of course, it is always possible that something is wrong with the experiments. However, something like a conspiracy on nature's side would be required to dismiss all of them. In short, it looks very much as if Bell's inequality is false and quantum mechanical predictions are correct.

⁸ $\sin^2 \alpha = \frac{1}{2} - \frac{1}{2}\cos 2\alpha$ and $\cos^2 \alpha = \frac{1}{2} + \frac{1}{2}\cos 2\alpha$.

If so, strong property realism must relinquish PL, that is, must be non-local, or, which is the same, PL and strong property realism, contextual or not, cannot be both true.⁹

The failure of Bell's inequality at the quantum level has occasionally been taken to entail that PL is false, if quantum mechanics is correct. For, one might argue, the EPRB paper suggests that the only hope to avert non-locality is by adopting strong property realism and yet Bell, assuming the experiments are correct, has effectively shown that strong property realism must be non-local in order to agree with experience. This belief is then reinforced by the fact that Bohm's mechanics, the most developed version of a quantum theory interpreted in terms of strong property realism, is non-local. However, such a conclusion is unwarranted since, for example, it may be possible to develop an interpretation of quantum mechanics that adopts weak property realism, as we shall see.

9.5 The Greenberger-Horne-Zeilinger Theorem

Bell's theorem deals with averages, and therefore it only shows the statistical incompatibility of quantum mechanics and local strong property realism. By contrast, the Greenberger-Horne-Zeilinger theorem (GHZ for short) addresses the non-statistical incompatibility of local strong property realism and quantum mechanics (Home, D., (1997): 235-37).¹⁰

Consider three spin-half particles a , b , and c that are widely separated but in the entangled state

⁹ For a description of the relevant experiments, including possible loopholes, see Baggott, J., (2004).

¹⁰ As for Bell's theorem, contextuality does not affect the proof.

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow_z^a \uparrow_z^b \uparrow_z^c\rangle - |\downarrow_z^a \downarrow_z^b \downarrow_z^c\rangle \right). \quad (9.5.1)$$

By using $\hbar/2$ as a unit of measure, all spin component measurements will return ± 1 , and the spin operators will be Pauli matrices. Remembering that spin components of different particles commute, direct calculation shows that the following four eigenvalue equations obtain:

$$\sigma_x^a \sigma_x^b \sigma_x^c |\Psi\rangle = -|\Psi\rangle, \quad (9.5.2)$$

$$\sigma_x^a \sigma_y^b \sigma_y^c |\Psi\rangle = |\Psi\rangle, \quad (9.5.3)$$

$$\sigma_y^a \sigma_x^b \sigma_y^c |\Psi\rangle = |\Psi\rangle, \quad (9.5.4)$$

$$\sigma_y^a \sigma_y^b \sigma_x^c |\Psi\rangle = |\Psi\rangle. \quad (9.5.5)$$

Hence, the value of the x -component or y -component of the spin of any of the three particles can be determined with probability 1 just by measuring the x -components or y -components of the other two. For example, $\sigma_x^a \sigma_y^b \sigma_y^c |\Psi\rangle = |\Psi\rangle$ tells us that to know S_x^a all one needs are S_y^b and S_y^c . PL plus strong property realism tell us that $S_x^a, S_y^a, S_x^b, S_y^b, S_x^c, S_y^c$ have had definite values all along.

Now, if $\hat{A}, \hat{B}, \hat{C}$ are *mutually commuting* operators, then the functional relations among them are preserved among the observed values of their respective observables. In other words, if $f(\hat{A}, \hat{B}, \hat{C}, \dots) = 0$, then $f(A, B, C) = 0$, where A, B, C , are (the values of) the measured observables, that is, the relevant eigenvalues. In other words, the algebraic

relations among commuting operators are mirrored by those among the values of their observables (their eigenvalues). Let us call this the “functional rule.”¹¹

The functional rule plus the four eigenvalue equations above tell us that the following equalities obtain:

$$S_x^a S_x^b S_x^c = -1, \quad (9.5.6)$$

$$S_x^a S_y^b S_y^c = 1, \quad (9.5.7)$$

$$S_y^a S_x^b S_y^c = 1, \quad (9.5.8)$$

$$S_y^a S_y^b S_x^c = 1. \quad (9.5.9)$$

However, multiplying the last three equalities, we get

$$(S_x^a S_y^b S_y^c)(S_y^a S_x^b S_y^c)(S_y^a S_y^b S_x^c) = 1, \quad (9.5.10)$$

that is,

$$S_x^a S_x^b S_x^c = 1, \quad (9.5.11)$$

which is incompatible with (9.5.2).¹²

¹¹ Note that the classical mechanical counterpart of the functional rule is trivially true.

For example, given that the formula $E = \frac{p^2}{2m^2} + V$ holds for conservative systems

(systems in which no energy is introduced or lost), we expect that at any time for such systems the value of energy to be equal to that of the square of momentum divided by twice the square of the mass, plus that of potential energy.

¹² Tests involving three GHZ photons have, by and large, agreed with quantum mechanical predictions: Pan, J., Bouwmeester, D., Daniell, M., Weinfurter, H., Zeilinger, A., (2000).

9.6 The Kochen-Specker Theorem

Bell's theorem deals with strong property realism (strong property contextuality) and locality; by contrast, the Kochen-Specker theorem (KS for short) deals with strong property realism, local or not, but not with property contextuality. There is a long tradition of counterexamples to strong property realism going back at least to Schrödinger. For example, consider a quantum harmonic oscillator with energy E_n . Then, $E_n = T + V$, where T is the kinetic energy and V is the potential energy. Assume strong property realism. Then, at all times the particle itself has definite values for x , T , and V . Since both T and V are positive quantities and E_n is fixed, when $T = 0$, V should have its maximum value, that is $E_n = V$. Hence, x too should have a definite maximum value. However, if we assume that measurement returns provide the actual values of the measured observables, quantum mechanics and experiment tell us that this is not so: there is no limit to how large x can be, although the probability of finding the particle far from the equilibrium position is very low. So, something has to give in the premises of the argument, and the most likely candidate seems the assumption of non-contextual strong property realism (Schrödinger, E., (1935) in Wheeler, J. A., and Zureck, W. H., (eds.) (1983): 152-67).

While Schrödinger's counterexample, whatever its merits, is quite specific, KS, which we are going to consider in the simplified formulation provided by Home, establishes that in general strong property realism is incompatible with the machinery of quantum mechanical operators (Home, D., (1997): 32-37). To avoid confusion, in the following, if \hat{O} is an operator, then O is its observable and $v(O)$ the value of O .

Let us consider a system of two spin-half particles, with or without entanglement (PL is irrelevant here), and spin measurements using $\hbar/2$ as the unit of measure, so that the usual spin operators reduce to Pauli operators. By exploiting the commutativity of operators for observables belonging to different particles, the following two equalities can be shown to hold:

$$\hat{S}_x^a \hat{S}_y^b \hat{S}_y^a \hat{S}_x^b \hat{S}_z^a \hat{S}_z^b = 1 \quad (9.6.1)$$

and

$$\hat{S}_x^a \hat{S}_x^b \hat{S}_y^a \hat{S}_y^b \hat{S}_z^a \hat{S}_z^b = -1. \quad (9.6.2)$$

By the functional rule (section 9.5), we obtain

$$v[(S_x^a S_y^b)(S_y^a S_x^b)(S_z^a S_z^b)] = 1 \quad (9.6.3)$$

and

$$v[(S_x^a S_x^b)(S_y^a S_y^b)(S_z^a S_z^b)] = -1. \quad (9.6.4)$$

Now, a special case of the functional rule is the product rule: if \hat{A} and \hat{B} are commuting operators, then $v(AB) = v(A) \cdot v(B)$. By applying the product rule, we get

$$v(S_x^a S_y^b) v(S_y^a S_x^b) v(S_z^a S_z^b) = 1 \quad (9.6.5)$$

and

$$v(S_x^a S_x^b) v(S_y^a S_y^b) v(S_z^a S_z^b) = -1. \quad (9.6.6)$$

By applying the product rule again, we finally obtain

$$v(S_x^a) v(S_y^b) v(S_y^a) v(S_x^b) v(S_z^a) v(S_z^b) = 1 \quad (9.6.7)$$

¹³ It may be worth noticing that the following equalities hold: $(\sigma_x)^2 = (\sigma_y)^2 = (\sigma_z)^2 = 1$;

$\sigma_x \sigma_y = -\sigma_y \sigma_x = i\sigma_z$; $\sigma_y \sigma_z = -\sigma_z \sigma_y = i\sigma_x$; $\sigma_z \sigma_x = -\sigma_x \sigma_z = i\sigma_y$.

and

$$v(S_x^a)v(S_x^b)v(S_y^a)v(S_y^b)v(S_z^a)v(S_z^b) = -1. \quad (9.6.8)$$

However, the left sides of (9.6.7) and (9.6.8) are identical, while their right sides are not, which is impossible. The proof we just considered establishes the KS theorem for 4-dimensional space; however, it is possible to construct a much more complicated proof for 3-dimensional space. Lower dimensional spaces are exempt from KS. This limitation, however, is not very significant in that very little quantum mechanics can be carried on in such spaces.

Note that following the application of the product rule, we assumed that, for example, $v(S_x^a)$, the first factor in both (9.6.7) and (9.6.8), is the same in both equations, even if it enters into $S_x^a S_y^b$ and $S_x^a S_x^b$, which are not only different but also involve incompatible observables for b . But there is no a priori reason why this should be so. KS, then, can be escaped if one adopts property contextuality.

However, one can look at the lesson of KS differently. Algebraic contextuality disappears at the statistical level. If we measure O with P and Q in an ensemble of systems in state $|\Psi\rangle$, we get exactly the same statistics for O that we would have gotten if we had measured O with R and S , even if P and Q are incompatible with R and S . This seems to point in the direction of weak property realism, which would reject (9.6.7) and (9.6.8) as attributing simultaneous values to incompatible observables.

9.7 Measurement

Much of what is bizarre in standard quantum mechanics seems tied up with measurement. The wave function develops deterministically until measurement occurs, and then we have the collapse of the state vector, an unexpected and apparently

unreasonable suspension of TDSE. So, it might be a good idea to look at measurement more closely.

Typically, a measurement involves the establishment of a correlation between a quantity Q of a system S under investigation and a directly observable quantity R of a measuring device D . More precisely, if Q 's value is q , then there is a biunivocal function $f(q) = r$, where r is a value of R , such that to every value q of Q there corresponds one and only value of R , and vice versa. By reading off R 's value, let us say the position of a pointer on a graduated line, one can infer Q 's value. For example, if one dips the tip of a thermometer into a liquid, and the thermometer "reads" $30C^\circ$, then one infers that the liquid's temperature is $30C^\circ$. Since $f(q) = r$ is biunivocal, if we assume that S is in a specific state at the time of measurement, then the outcome of the measurement is determinate; in other words, every time we measure, we get a single determinate result. In addition, thermodynamics, and ultimately classical mechanics, explains why we got a reading of $30C^\circ$. It turns out that when we measure quantum mechanical observables, this expectation is satisfied: every time we measure, we do get a single determinate result. However, the way in which quantum theory explains this apparently obvious fact is rather problematic, *if* we assume that in the measurement both S and D can be described by quantum mechanics. Let us see how this comes about.

Suppose that system S is in the eigenstate $|\psi_n\rangle$ of \hat{Q} with eigenvalue q_n and that the measuring device D is in the eigenstate $|\chi_0\rangle$ of \hat{R} with eigenvalue r_0 . Suppose now that as the result of the interaction between S and D $|\chi_0\rangle$ changes into $|\chi_n\rangle$ with eigenvalue r_n , while $|\psi_n\rangle$ remains unchanged, so that the compound system goes from

state $|\psi_n\rangle \otimes |\chi_0\rangle$ to state $|\psi_n\rangle \otimes |\chi_n\rangle$.¹⁴ Then the eigenvector to eigenvalue link guarantees that the measuring device will record r_n , that Q had value q_n both before and after the measurement and that R had value r_0 before the measurement and value r_n after the measurement.

However, things change dramatically if the original state of S is a superposition of eigenstates $|\psi_i\rangle$ of Q so that the pre-measurement state of the compound $S+D$ is

$$|\Psi\rangle = \left(\sum_i c_i |\psi_i\rangle \right) \otimes |\chi_0\rangle. \quad (9.7.1)$$

For, since the time evolution of $|\Psi\rangle$ is linear, after measurement $|\Psi\rangle$ evolves into the pure state

$$|\Psi'\rangle = \sum_i c_i (|\psi_i\rangle \otimes |\chi_i\rangle) \quad (9.7.2)$$

a superposition of the tensor products of the various $|\psi_i\rangle$'s and their corresponding $|\chi_i\rangle$'s.¹⁵ Note that crucially (9.7.2) is a pure, not a mixed state. But then, after measurement D 's pointer is delocalized in the sense that its position, such as it is, is given

¹⁴ Measurements in which the measured property is preserved are variously called “non-destructive,” “non-demolition,” or “of the first kind.” Such measurements are often difficult to carry out in the quantum world. By contrast, measurements in which the measured property is seriously altered, often uncontrollably, are called “destructive,” “demolition,” or “of the second kind.”

¹⁵ Remember that tensor product is distributive under addition:

$$(A + B) \otimes C = A \otimes C + B \otimes C.$$

Remember also that whether a state is a superposition depends on the observable being measured.

by a superposition of the states $|\chi_i\rangle$ associated with each specific position. In fact, according to EE, D's pointer is not at any position. In other words, the superposition at work with S has infected D, as it were. The problem, of course, is that we know from experience (or so we think) that this is not what happens, since after a measurement the pointer is in some definite position. Note that the complexity of the superposition seems irrelevant to the argument since no limit need be set for the index i . So, it is certainly true that the final state vector will have about 10^{23} components, but by hypothesis a good measuring device insures that the system develops from $|\psi_n\rangle \otimes |\chi_0\rangle$ to $|\psi_n\rangle \otimes |\chi_n\rangle$ when $|\psi_n\rangle$ is an eigenstate of the observable's operator, and the linearity of the dynamical equations guarantees that the complex system ends up in state (9.7.2).

As D can be any macroscopic system working as a measuring device, Schrödinger's cat will do the job.

A cat is penned up in a steel chamber, along with the following diabolical device (which must be secured against the direct interference by the cat). In a Geiger counter there is a tiny bit of radioactive substance, *so* small that *perhaps* in the course of one hour one of the atoms decays, but also, with equal probability, perhaps none; if it happens, the counter tube discharges and through a relay releases a hammer which shatters a small flask of hydrocyanic acid. If one has left the entire system to itself for an hour, one would say that the cat still lives *if* meanwhile no atom has decayed. The first atomic decay would have poisoned it. The ψ -function of the entire system would express this by having in it the living and the dead cat (pardon the expression) mixed or smeared out in equal parts

(Schrödinger, E., (1935) in Wheeler, J. A., and Zureck, W. H., (eds.)
(1983): 157).

In other words, the state of the wretched beast will end up entangled with that of the quantum system and so neither live nor dead until checked upon, which, to use Schrödinger's word, seems quite "ridiculous". In sum, as long as both S and D are described by TDSE, it very much looks as if no measuring result at all can be obtained as long as S's pre-measurement state is a superposition of the eigenstates of the observable Q measured by the device D. As this is because (9.7.2) is not a (proper) mixture but a pure state, let us call this *the pure state problem*. Note that a solution to this problem need not involve a solution of *the determinate result problem*, to wit, of why we got *the determinate result* we did in fact get.

Let us consider what has gone into the production of the pure state problem. First, we have assumed that TDSE has universal validity. Let us call this *the universal validity principle*. Second, we have assumed that our observations have a certain degree of reliability: when it seems to us that the pointer is at a certain position and nowhere else, we are not deluded. Let us call this *the observer's reliability principle*. Third, we have assumed EE. More precisely, we have assumed half of the principle: an observable *O* possesses a certain eigenvalue only if the system is in an appropriate eigenstate of *O*. Let us call this *the eigenvalue only if eigenstate principle*. Fourth, we have assumed that state vectors are unqualifiedly about physical systems in that we need not specify from which point of view (in relation to what other physical system) we determine the state vector. Let us call this *the absolute state principle*. These four principles are necessary

ingredients for the production of the pure state problem; not surprisingly, as we shall see, many of the attempted solutions reject one of the four.

Since many solutions to the measurement problem have been put forth, it may be a good idea to set some criteria for judging success. There seem to be two such criteria. Obviously, a solution must explain the existence and uniqueness of measurement returns. In other words, while it is not necessary to explain why one got *the* return one actually got (why, say, one got +1 instead of -1 when measuring S_z on a spin-half particle), it is certainly necessary to explain why one got *one* measurement return. In addition, any solution must explain such existence and uniqueness in a non-traumatic way, as it were, without introducing excessively *ad hoc* abrupt breaks in the linear evolution of quantum states guaranteed by TDSE.

Exercises

Exercise 9.1

Determine the eigenvalues and the normalized eigenvectors of \hat{S}_n . [Hint: Since classically $\mathbf{S}_n = \mathbf{S}_z \cos\phi + \mathbf{S}_x \sin\phi$, it must be the case that $\hat{S}_n = \hat{S}_z \cos\phi + \hat{S}_x \sin\phi$].

Exercise 9.2

1. True or false: Any value determinist theory must be non-local to agree with experience.
2. True or false: The assumptions of the EPR paper are in conflict with experimental evidence.

Exercise 9.3

1. Does weak value determinism allow the transition from (9.5.10) to (9.5.11)?
2. True or false: The transition from (9.5.2-5) to (9.5.6-9) involves the product rule.

Exercise 9.4

1. Prove that $\sigma_x \sigma_y \sigma_z = i$.
2. Prove that $\hat{S}_x^a \hat{S}_y^b \hat{S}_y^a \hat{S}_x^b \hat{S}_z^a \hat{S}_z^b = 1$.
3. Prove that the product of commuting Hermitian operators is Hermitian.

Exercise 9.5

Suppose that the system S is a spin-half particle is in state $|\uparrow_x\rangle$ and the measuring apparatus is a SGZ in initial state $|\chi_0\rangle$. Determine the state of the compound system S+SGZ just before and after the measurement according to TDSE.

Answers to the Exercises

Exercise 9.1

Let us start by noting that classically $\mathbf{S}_n = \mathbf{S}_z \cos \phi + \mathbf{S}_x \sin \phi$. Since the functional relations among the operators representing observables are identical to those among the observables, it follows that $\hat{\mathbf{S}}_n = \hat{\mathbf{S}}_z \cos \phi + \hat{\mathbf{S}}_x \sin \phi$, that is,

$$\hat{\mathbf{S}}_n = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cos \phi + \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \sin \phi = \frac{\hbar}{2} \begin{pmatrix} \cos \phi & \sin \phi \\ \sin \phi & -\cos \phi \end{pmatrix}. \text{ So,}$$

$$\frac{\hbar}{2} \begin{pmatrix} \cos \phi & \sin \phi \\ \sin \phi & -\cos \phi \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \text{ Hence, we get two simultaneous equations:}$$

$$\begin{cases} \left(\frac{\hbar}{2} \cos \phi - \lambda \right) x_1 + \left(\frac{\hbar}{2} \sin \phi \right) x_2 = 0 \\ \left(\frac{\hbar}{2} \sin \phi \right) x_1 - \left(\lambda + \frac{\hbar}{2} \cos \phi \right) x_2 = 0 \end{cases}$$

$$\text{So, } 0 = \begin{vmatrix} \frac{\hbar}{2} \cos \phi - \lambda & \frac{\hbar}{2} \sin \phi \\ \frac{\hbar}{2} \sin \phi & -\frac{\hbar}{2} \cos \phi - \lambda \end{vmatrix} = \left(\frac{\hbar}{2} \cos \phi - \lambda \right) \left(-\frac{\hbar}{2} \cos \phi - \lambda \right) - \left(\frac{\hbar}{2} \right)^2 \sin^2 \phi.$$

$$\text{Developing, and after a bit of algebra, we obtain } 0 = \lambda^2 - \left(\frac{\hbar}{2} \right)^2 (\cos^2 \phi + \sin^2 \phi) = \lambda^2 - \left(\frac{\hbar}{2} \right)^2.$$

The solutions are $\lambda_1 = -\hbar/2$ and $\lambda_2 = \hbar/2$. Now, let us determine the eigenvector for

$\lambda_1 = -\hbar/2$. Plugging this eigenvalue into the first of the two original simultaneous

$$\text{equations, we obtain } x_1 \cos \phi + x_2 \sin \phi = -x_1, \text{ so that } x_2 = -\frac{1 + \cos \phi}{\sin \phi} x_1 = -\frac{\cos \left(\frac{\phi}{2} \right)}{\sin \left(\frac{\phi}{2} \right)} x_1.$$

Hence, the eigenvector is $X_{\lambda_1} = \begin{pmatrix} 1 \\ \cos\left(\frac{\phi}{2}\right) \\ -\sin\left(\frac{\phi}{2}\right) \end{pmatrix}$. Since $\langle X_{\lambda_1} | X_{\lambda_1} \rangle = 1 + \frac{\cos^2\left(\frac{\phi}{2}\right)}{\sin^2\left(\frac{\phi}{2}\right)} = \frac{1}{\sin^2\left(\frac{\phi}{2}\right)}$,

the normalized vector is $\|X_{\lambda_1}\| = \begin{pmatrix} \sin\left(\frac{\phi}{2}\right) \\ -\cos\left(\frac{\phi}{2}\right) \end{pmatrix}$. As any multiple of an eigenvector is also an

eigenvector, by multiplying by -1 we may rewrite the normalized eigenvector as

$\|X_{\lambda_1}\| = \begin{pmatrix} -\sin\left(\frac{\phi}{2}\right) \\ \cos\left(\frac{\phi}{2}\right) \end{pmatrix}$. Now, let us determine the eigenvector for $\lambda_1 = \hbar/2$. Plugging this

eigenvalue into the first of the two original simultaneous equations, we obtain

$(\cos\phi - 1)x_1 + x_2 \sin\phi = 0$, so that $x_2 = \frac{1 - \cos\phi}{\sin\phi} x_1 = \frac{\sin\left(\frac{\phi}{2}\right)}{\cos\left(\frac{\phi}{2}\right)} x_1$. Hence, the eigenvector is

$X_{\lambda_2} = \begin{pmatrix} 1 \\ \sin\left(\frac{\phi}{2}\right) \\ \cos\left(\frac{\phi}{2}\right) \end{pmatrix}$, and since $\langle X_{\lambda_2} | X_{\lambda_2} \rangle = 1 + \frac{\sin^2\left(\frac{\phi}{2}\right)}{\cos^2\left(\frac{\phi}{2}\right)} = \frac{1}{\cos^2\left(\frac{\phi}{2}\right)}$. Hence, the normalized

vector is $\|X_{\lambda_2}\| = \begin{pmatrix} \cos\left(\frac{\phi}{2}\right) \\ \sin\left(\frac{\phi}{2}\right) \end{pmatrix}$.

Exercise 9.2

1. False, as weak value determinism escapes the strictures of Bell's Theorem.
2. True, as the EPR paper endorses both strong value determinism and PL.

Exercise 9.3

1. No.
2. True.

Exercise 9.4

1. $\sigma_x \sigma_y \sigma_z = i \sigma_z \sigma_z = i$.
2. Since operators of spin components form different particles commute, and we are dealing with Pauli operators, $\hat{S}_x^a \hat{S}_y^b \hat{S}_y^a \hat{S}_x^b \hat{S}_z^a \hat{S}_z^b = \hat{S}_x^a \hat{S}_y^a \hat{S}_z^a \hat{S}_y^b \hat{S}_x^b \hat{S}_z^b = -ii = 1$
3. Let \hat{A} and \hat{B} be Hermitian and commute. Then, $(AB)^+ = B^+ A^+$ because the Hermitian conjugate of a product is the product of the Hermitian conjugates in reverse order. But since we are dealing with Hermitian operators, $B^+ A^+ = BA$ and they commute, $BA = AB$. Hence, $(AB)^+ = AB$.

Exercise 9.5

Since $|\uparrow_x\rangle = \frac{1}{\sqrt{2}} (|\uparrow_z\rangle + |\downarrow_z\rangle)$, the pre-measurement state of the compound system is

$$|\Xi\rangle = |\uparrow_x\rangle \otimes |\chi_0\rangle = \frac{1}{\sqrt{2}} (|\uparrow_z\rangle + |\downarrow_z\rangle) \otimes |\chi_0\rangle, \text{ from which we obtain}$$

$$|\Xi\rangle = \frac{1}{\sqrt{2}} (|\uparrow_z\rangle \otimes |\chi_0\rangle + |\downarrow_z\rangle \otimes |\chi_0\rangle). \text{ Since the time development of } |\Xi\rangle \text{ is linear, and}$$

$|\uparrow_z\rangle \otimes |\chi_0\rangle$ becomes $|\uparrow_z\rangle \otimes |\chi_+\rangle$ and $|\downarrow_z\rangle \otimes |\chi_0\rangle$ becomes $|\downarrow_z\rangle \otimes |\chi_-\rangle$, after measurement,

$$|\Xi\rangle \text{ becomes } |\Xi'\rangle = \frac{1}{\sqrt{2}} (|\uparrow_z\rangle \otimes |\chi_+\rangle + |\downarrow_z\rangle \otimes |\chi_-\rangle).$$