qmd411

Quantum Measurements

Robert B. Griffiths
Version of 24 October 2012

References:
CQT = *Consistent Quantum Theory* by Griffiths (Cambridge, 2002)

Contents

1 Introduction

1.1 Scope of these notes

★ In quantum foundations, the study of the conceptual basis of quantum mechanics, "measurements" have been an enormous conceptual headache. They have given rise to endless arguments.

• To resolve the arguments one needs to adopt a consistent approach in which the measuring apparatus itself is treated in quantum mechanical, not classical, terms, and probabilities are properly defined. This is discussed at some length in CQT Chs. 17 and 18. However, the reader who wants to delve into this would do well to begin with toy models in Sec. 7.4 of CQT.

★ The approaches found in most textbooks are inadequate, confusing, and out of date. Not always wrong, but also not very reliable.

• The purpose of these notes is not to discuss measurements in detail. For that see CQT Chs. 17, 18. Instead, it is to provide a brief introduction which gets to the essentials.

★ It is assumed that the reader knows something about quantum histories and consistency conditions.

• See CQT Ch. 8 for histories.

• We will use consistency conditions at the level of chain kets, CQT Sec. 11.6. Also in the notes *Probabilities for Quantum Histories II. Multitime Histories*.

★ In these notes we do *not* consider POVMs (positive operator valued measures). They are useful if one is interested in "noisy" measurements in which the outcome is not linked in a deterministic fashion with the property being measured. However, they involve additional mathematics. Instead we focus on *ideal* measurements, which are easier to understand.

1.2 Measurements and histories

★ A measuring apparatus determines properties of a quantum system by interacting with it in some way. Consequently, it is impossible to treat either the system or the apparatus as isolated, at least during the crucial time period when they interact. Instead, one should regard both together as constituting a single combined and isolated quantum system to which Schrödinger's equation or the corresponding unitary time development operators can be applied.

★ It is then necessary to introduce appropriate families of quantum histories in order to describe what is going on. We need families with the following features:

• At a time when the measurement is completed a suitable decomposition of the identity for the apparatus in which *measurement outcomes* are well defined.

◦ In the antique terminology employed in quantum foundations, where these issues were discussed well before the electronic age, measurement outcomes are traditionally referred to as *pointer positions*. However, any sort of macroscopically distinct states will do equally well. Symbols printed out on a sheet of paper. Bits stored in the computer memory, etc.

◦ Many physicists working in quantum foundations have failed to understand the importance of stochastic families of histories with projectors of this type, and this is the main reason for their lack of progress in resolving the infamous *measurement problem*.

• Next, in order to speak of the measurement as *having measured something* the family of histories must contain projectors corresponding to appropriate microscopic physical properties of the measured system *before* the measurement took place.

◦ Textbook quantum mechanics does not contain the concepts needed to deal with properties of the measured system before a measurement takes place, and this is the reason why the computational rules provided in textbooks appear so mysterious. The rules are adequate for calculating probabilities of measurement outcomes and certain types of correlations, but they do not provide an adequate physical understanding.

• Finally, the families must contain some sort of initial state(s) connected with the way in which the apparatus was set up, the way in which the quantum system of interest was prepared, etc.

◦ Since measuring apparatus is macroscopic one ought to specify initial states using "macroscopic" projectors or else density operators. But then the consistency of the relevant families can no longer be discussed using only chain kets. Consequently the treatment in these notes, which uses a pure initial state of the apparatus, is to some extent oversimplified, but it is not misleading.

★ Thus an adequate discussion of measurements using basic quantum principles rather than arm waving must be capable of considering histories containing at least three times: the initial time of preparation, the time just before the measurement takes place when the system possesses the measured property, and the time after the measurement has taken place, so that one can understand how the (macroscopic) "pointer position" is correlated with the earlier microscopic property one is interested in.

◦ Assigning probabilities to histories involving three (or more) times using the appropriate extension of the Born rule requires that one pay attention to *consistency conditions*.

2 Stern-Gerlach

★ A Stern-Gerlach measurement setup is shown schematically in Fig. 1. A sliver atom, which is a spin-half particle in its electronic ground state, passes through a magnetic field with a strong gradient in the z direction in such a way that if it is initially in the $S_z = +1/2$ spin state when it enters the apparatus (left side in the figure) it emerges from the field gradient region with a small upwards component of velocity, whereas if the initial spin state is $S_z = -1/2$ it has a small downwards component of velocity. Hence the particles come out of the apparatus in two separate beams, corresponding to the two possible values of S_z.

◦ The particles that emerge then have to be detected. In the original experiment this was done by letting the silver atoms strike and stick to a glass plate. It required a long run in order to get enough atoms so that the deposits on the glass plate could be seen. Experimental techniques have improved since then so that nowadays under appropriate circumstances it is possible to detect single atoms in such a beam.

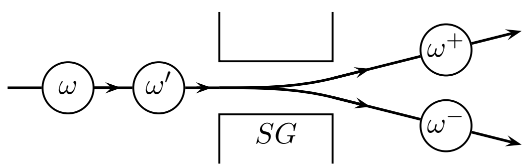

Figure 1: Schematic diagram of spin-half particle passing through Stern Gerlach apparatus. The circles labeled ω, ω', etc., stand for position wave packets.

★ This example illustrates the essence of a quantum measurement The system to be measured possesses one of several possible microscopic ("quantum") properties when it begins to interact with the measuring apparatus designed to measure these properties. Interaction with the apparatus results in the microscopic property producing a macroscopic difference after a certain time interval; this macroscopic difference is the "pointer position" of the traditional measurement discussion.

• In the Stern-Gerlach case the "pointer" can be thought of as the particle position as it exits the apparatus: is it in the upper or the lower beam? In this case the center of mass motion of the particle, a degree of freedom which is independent of the spin angular momentum, can be thought of as part of the measuring apparatus. When the atom emerges from the field gradient the spatial separation of the beams is macroscopic: small but still very large compared to the dimension of an atom.

• One could imagine a further stage of detection by, for example, ionizing the atom as it emerges in one beam or another, and then using an electron multiplier to turn it into a pulse of electric current.

◦ This sort of thing was not available in 1920 when the original apparatus was built, so Stern and Gerlach had to send a very large number of atoms through the apparatus in order to see a signal.

3 Ideal Measurements

3.1 Some definitions

★ An *ideal* measurement is one in which there is a one-to-one correspondence between the measured property at the time just before the measurement takes place and the measurement outcome as indicated by the apparatus "pointer position" when the measurement is over.

• The measurement may or may not preserve the property that is being measured, i.e., the measured system may or may not have the same property after the measurement is over. If it does, we say the measurement is *nondestructive* for this property; if the property is not preserved the measurement is *destructive*. A *nondemolition* measurement is a nondestructive measurement in which the property remains unchanged at still later times, as long as the measured system is no longer interacting with the measuring apparatus; i.e., the property is an energy eigenstate of the measured system.

• The Stern-Gerlach apparatus is an example of a nondestructive measurement in the sense that S_z for the particle emerging from the magnetic field gradient is the same as when it entered the field gradient. However, if the particle is itself detected either by letting it strike a glass place or by ionizing it, then the overall measurement is destructive; S_z will not be preserved.

• It is also a nondemolition measurement in the sense that as long as the emerging particle continues to travel through space the spin direction will be unchanged as long as there is no component of magnetic field perpendicular to the spin direction.

★ The Stern-Gerlach device also illustrates the concept of a *preparation*: if, for example, one places a small hole at the end of the field gradient that only allows the upper beam to emerge, then one has a source of spin-half particles with a definite value of the z component of spin, say $S_z = +1/2$. See the comments in Sec. 5.1 below.

• A particle prepared in this way can then in principle be subject to a further measurement, and if that is a measurement of S_z one expects that it will always give the same answer, in agreement with the state that has been prepared.

★ Measurements and preparations are concepts not limited to single particles. However, it will be convenient in the discussions which follow to use the term "particle" for the measured (or prepared) system and "apparatus" for the device that carries out the measurement (or the preparation). The apparatus will then include a "pointer" that indicates the state of the particle before the measurement took place (or after the preparation took place in the case of a preparation).

3.2 A simple measurement model

★ In this model we assume that there is a particle or system to be measured, Hilbert space $\mathcal{H}_a$, and a measuring device, Hilbert space $\mathcal{H}_m$, that will carry out the measurement. The Hilbert space of the two together is $\mathcal{H} = \mathcal{H}_a \otimes \mathcal{H}_m$. Let $\{|a^j\rangle\}$, $j = 1, 2, \ldots f$, be an orthonormal basis for the particle. Let $|M_0\rangle$ be the initial state—assume it is normalized—of the measuring apparatus, and assume that the interaction of the particle with the apparatus, which takes place during the time interval from t_1 to t_2, can be represented by the unitary time transformation

$$|a^j\rangle \otimes |M_0\rangle \rightarrow |s^j\rangle \otimes |M^j\rangle \text{ for } 1 \leq j \leq m. \tag{1}$$

We assume the $|M^j\rangle$ states are mutually orthogonal, $\langle M^j|M^k\rangle = \delta_{jk}$, but not that they form a basis for the apparatus Hilbert space $\mathcal{H}_m$, which may be very large.

• One possibility for the particle states on the right side of (1) is that $|s^j\rangle = |a^j\rangle$. This is appropriate if one assumes the measurement is nondestructive for the $\{[a^j]\}$, but otherwise it is not essential. One could just as well assume that all the $|s^j\rangle$ are identical to each other, say $|s^j\rangle = |s^0\rangle$ for all j. In many measurement situations one does not care what happens to the particle after the measurement is complete, and then the choice of the $|s^j\rangle$ is unimportant.

□ Exercise. Verify that in the case $|s^j\rangle = |s^0\rangle$ for all j the transformation (1) maps an orthonormal collection of states to an orthonormal collection of states, so it is consistent with unitary time development.

★ Next assume there is some decomposition of the identity $\{P^j\}$ for the apparatus such that

$$P^j|M^j\rangle = |M^j\rangle \text{ for } 1 \leq j \leq f, \tag{2}$$

along with $P^0 = I + m - \sum_j P^j$.

○ Why not just set $P^j = [M^j]$, the projector onto the space containing $|M^j\rangle$. This would yield the results we are interested in, but would be somewhat less realistic, in that distinct macroscopic states, distinct pointer positions, are best thought of as associated with very large subspaces of the apparatus Hilbert space $\mathcal{H}_m$, which in any realistic situation has an extremely large dimension.

○ Of course, to be realistic we should not have limited ourselves to a single pure initial state $|M_0\rangle$. See the discussion in CQT Sec. 17.5 for how to remove this restriction.

★ Next we assume that at an initial time t_0 the particle is in a normalized state

$$|\psi_0\rangle = \sum_j c_j|a^j\rangle, \tag{3}$$

where the c_j are complex coefficients. (One possibility is that $c_j = 0$ for all j with the exception $j = l$, which means the particle starts off in the basis state $|a^l\rangle$, but we don't assume this in general.) We also assume that at this time the apparatus is in the state $|M_0\rangle$, so the combined system plus apparatus is in the initial state

$$|\Psi_0\rangle = |\psi_0\rangle \otimes |M_0\rangle = \sum_j c_j|a^j\rangle \otimes |M_0\rangle. \tag{4}$$

• Assume that between this time and a slightly later time t_1 both the particle and the apparatus are isolated and undergo a trivial time dependence, so that unitary time development results in

$$|\Psi_1\rangle = T(t_1, t_0)|\Psi_0\rangle = |\Psi_0\rangle. \tag{5}$$

• During the time interval from from t_1 to t_2, when the particle interacts with the apparatus, the unitary time development is determined by (1), resulting in

$$|\Psi_2\rangle = T(t_2, t_1)|\Psi_1\rangle = T(t_2, t_0)|\Psi_0\rangle = \sum_j c_j|s^j\rangle \otimes |M^j\rangle. \tag{6}$$

★ To describe the measurement process it is convenient to employ a family of histories at $t_0 < t_1 < t_2$ of the form

$$[\Psi_0] \odot \{[a^j]\} \odot \{P^j\}, \tag{7}$$

which is to say at time t_1 we focus on particle states—of course, $[a^j]$ means $[a^j] \otimes I_m$—and at time t_2 on macroscopic pointer states of the apparatus—with, of course P^j understood as $I_a \otimes P^j$.

• It is then straightforward to show that the only nonzero chain kets for the family (7) are those in which the same j value occurs at t_2 as at t_1. And these are mutually orthogonal for different j, as follows from the fact that $P^j P^k = \delta_{jk} P^j$. Thus we arrive at probabilities (where the subscript 1 or 2 indicates the time)

$$\Pr([a^j]_1, P_2^k) = \delta_{jk}|c_j|^2, \quad \Pr([a^j]_1) = \Pr(P_2^j) = |c_j|^2. \tag{8}$$

As a consequence, for those cases in which $c_j \neq 0$, and therefore $\Pr([a^j]_1) > 0$ and $\Pr(P_2^j) > 0$, one has conditional probabilities

$$\Pr(P_2^k \mid [a^j]_1) = \delta_{jk} = \Pr([a^k]_1 \mid P_2^j) \tag{9}$$

□ Exercise. Work out the chain kets, and check (8) and (9).

★ This simple model exemplifies what might be called the *Competent Experimentalist Principle* or CEP. Apparatus built by a competent experimentalist and intended to measure some microscopic property will do what it was designed to do.

• This means in particular that if we analyze the apparatus and system to be measured in quantum mechanical terms using an appropriate *pointer basis* (it would be more accurate to say pointer decomposition of the identity) at the time when the apparatus has finished interacting with the particle, and an appropriate *property basis* for the particle before the measurement takes place, there should be a perfect statistical correlation between the two, as exhibited in (9)

• Note that this perfect correlation arises both if the particle is initially at time t_0 in one of the states $|a^j\rangle$ which the measurement is designed to measure, but also if it is in any superposition of such states, as assumed in (3).

★ The identity of the probabilities on the right side of (8) tells us that in the situation under discussion the probability of a particular measurement outcome at time t_2, $\Pr(P_2^j)$, is the same as the probability that the particle being measured was in the state $[a^j]$ at the earlier time t_1. As they are the same, one can refer either to the probability of the measurement outcome or to the probability that the measured system had this probability at the earlier time.

★ So why all the talk about measurements when all we need to calculate is the *microscopic* probability $\Pr([a^j]_1)$? It has to do with the historical development of quantum mechanics: "measurement" was invoked to get rid of certain conceptual problems and paradoxes back in the days before there was a fully consistent probabilistic formulation of quantum theory.

• In particular, the scheme proposed by von Neumann in his book [1] can be viewed from the modern perspective as an attempt to describe measurements using the consistent family

$$[\Psi_0] \odot \{[\Psi_1], I - [\Psi_1]\} \odot \{P^j\}. \tag{10}$$

That is, at time t_1 one uses the unitarily evolved state $[\Psi_1]$ in place of the properties $\{[a^j]\}$ the apparatus was designed to measure. In general, unless $|\psi_0\rangle$ happens to be one of the $|a^j\rangle$, $[\Psi_1]$ is incompatible with the properties of interest, and therefore the use of the family in (10) prevents any discussion of the latter.

• This is the fundamental difficulty with von Neumann's scheme, and the quantum textbooks that follow it: one can talk about the measurement outcome, but not about what it was that was (supposedly) measured!

◦ The student asked on an examination to "calculate the probability that a measurement of this particular sort has the following outcome" should assume (in the absence of any indication to the contrary) an implicit reference to the CEP, and then go ahead and use the Born rule (or whatever) to find the probability of the property (properties) of interest just before the measurement took place.

★ The CEP can be extended to *preparations* in an obvious way: Apparatus built by a competent experimentalist and intended to produce a particle (or other system) with a particular microscopic (quantum) property will do what it was designed to do.

• See the discussion of preparations in Sec. 5.1 below.

3.3 Example: Stern-Gerlach

★ Consider a destructive measurement of the z component of angular momentum of a spin-half particle, e.g., by a Stern-Gerlach magnet as in Fig. 1, followed by some detector which determines whether the emerging particle is in the upper or lower beam. Let the corresponding apparatus pointer positions be denoted by the projectors P^+ and P^- Assume the initial state at t_0 is

$$|\psi_0\rangle = \alpha|z^+\rangle + \beta|z^-\rangle, \tag{11}$$

with $|\alpha|^2 + |\beta|^2 = 1$, so $|\psi_0\rangle$ is normalized. What is the probability that the measurement will yield P^+ indicating $S_z = +1/2$?

• The device has been designed to determine the properties in the decomposition $[z^+]+[z^-] = I$, so all we need to do is to calculate these probabilities at a time t_1 just before the particle enters the measuring device, using the Born rule. Since the dynamics is trivial, $|\psi_1\rangle = |\psi_0\rangle$:

$$\Pr(S_z = +1/2) = \langle\psi_1|[z^+]|\psi_1\rangle = |\alpha|^2, \quad \Pr(S_z = -1/2) = \langle\psi_1|[z^-]|\psi_1\rangle = |\beta|^2. \tag{12}$$

★ Of course a competent experimentalist who can build a device to measure whether a qubit is in $|z^+\rangle$ or $|z^-\rangle$ can also build one to measure whether a system is in $|w^+\rangle$ or $|w^-\rangle$ for any pair of states that form an orthonormal basis for spin half. However, there is an alternative to building all manner of different devices, which is to place just ahead of the Stern-Gerlach magnet a region of uniform magnetic field which will make the state $S_w = +1/2$ precess into the state $S_z = +1/2$, and likewise $S_w = -1/2$ into $S_z = -1/2$. Its action on the spin is given by the unitary operator

$$T(t_1', t_1) = W = |z^+\rangle\langle w^+| + |z^-\rangle\langle w^-|. \tag{13}$$

• The device including this "pre-processor" is shown schematically in Fig. 2.

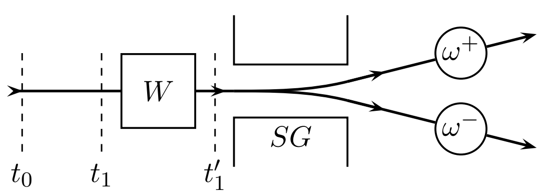

Figure 2: Device to measure S_z preceded by region of uniform magnetic field W which rotates $S_w = \pm 1/2$ into $S_z = \pm 1/2$.

• Now consider the family of histories at times $t_0 < t_1 < t_1'$, where the times are indicated in the figure:

$$[\psi_0] \odot \{[w^+], [w^-]\} \odot \{[z^+], [z^-]\}. \tag{14}$$

It is easy to show that there are only two nonzero chain kets, and they are orthogonal to each other.

☐ Exercise. Work out the chain kets. Note that the time development from t_0 to t_1 is given by $T = I$, and from t_1 to t_1' by W.

• The corresponding probabilities are

$$\Pr([\psi_0] \odot [w^+] \odot [z^+]) = |\langle w^+|\psi_0\rangle|^2, \quad \Pr([\psi_0] \odot [w^-] \odot [z^-]) = |\langle w^-|\psi_0\rangle|^2. \tag{15}$$

☐ Exercise. Check the math.

• Consequently, it follows that the final measurement outcome is P^+ if and only if the particle had the property $[w^+]$ at time t_1, and P^- if and only if it had the property $[w^-]$ at t_1. So we can think of the unitary W followed by the device which measures S_z as together comprising a slightly more complicated measuring device which measures whether the qubit is in the state $|w^+\rangle$ or $|w^-\rangle$ at the time t_1 just before it enters the compound ($W + SG$) device.

3.4 Using the right basis

★ CharliE is a *Competent Experimentalist* who knows how to build apparatus in such a way that it can reliably measure some microscopic property. He has just constructed a device that will measure S_x for a spin-half particle; $S_x = +1/2$ and $-1/2$ correspond to the the qubit states $|x^+\rangle$ and $|x^-\rangle$. Naturally, he also knows how to build an apparatus that will measure S_z, the states $|z^+\rangle$ and $|z^-\rangle$.

★ Charlie collaborates with a *Helpful Theorist* HerberT, who uses the Born rule (or sometimes its extensions, see CQT Ch. 10) to calculate the probability that if the qubit is in the state $|\psi_0\rangle$ at the beginning of the experiment, and then undergoes various interactions on the way to the final measurement, it will be in one of the two states $|z^+\rangle$ or $|z^-\rangle$ just before it reaches Charlie's detector.

• The two get together after a long night of hard work.

Charlie: What a tiresome business! I worked all night and only managed to make a thousand measurements of S_x. Nonetheless, that's enough so I can be confident that $\Pr(S_x = +1/2)$ is 0.33, and $\Pr(S_x = -1/2)$ is 0.67, given that I always used the starting state $|\psi_0\rangle$.

Herbert: What a nasty calculation. It took me all night because that system of yours is so complicated, but finally I figured out that the proper quantum calculation yields $\Pr(S_z = +1/2)$ is 0.82, and $\Pr(S_z = -1/2)$ is 0.18. Doesn't look like that agrees with what you measured.

Charlie: But you calculated the probabilities for S_z, not S_x. That wasn't very helpful!

Herbert: Oh come on. Why didn't you measure S_z rather than S_x?

• It is clear that they've had a long night, and Herbert is not being too helpful, so let's send them out for a few cups of coffee. After the break the discussion is more civil:

Charlie: Tell you what. During the next run I'll sometimes measure S_x and sometimes measure S_z, and I'll keep the data separate. And why don't you go ahead and calculate the S_x probabilities as well, so we can make a comparison in both cases.

Herbert: Agreed! See you tomorrow.

★ The moral to draw from this is that the theorist, if he wants to be helpful to the competent experimentalist who has built an apparatus to measure something, needs to adopt for his calculations a quantum sample space that includes, at the time just before the measurement takes place, the properties that the apparatus is designed to measure. There is no rule that says the theorist must do this, and of course it does no harm to calculate things in all sorts of incompatible frameworks. But the properties most directly correlated with measurement outcomes are the ones the competent experimentalist designed his apparatus to measure, the $\{|a^j\rangle\}$ employed in the discussion in Sec. 3.2.

3.5 The counterfactual mistake

• The following type of reasoning has a long history in quantum foundations, and has led to all sorts of confusion.

★ Suppose Charlie's apparatus is set up to measure S_z and, to be specific, yields the outcome (pointer position) corresponding to $S_z = +1/2$. But suppose that in this particular case, involving this particular particle, Charlie had instead used the S_x rather than the S_z apparatus.

◦ Sometimes the apparatus can be designed so that it has a switch with two positions, with the switch in the "z" position it measures S_z, and with the switch in the "x" position it measures S_x. E.g., switching on or off the current that produces the magnetic field in the region W in Fig. 2. Charlie might choose one or the other type of measurement just before the arrival of the particle.

• Then surely *if* Charlie had used the S_x apparatus it *would* have yielded a pointer position corresponding to one of the two values, $S_x = +1/2$ or $S_x = -1/2$, just before the measurement took place.

• Consequently, in this particular case when we know (using the Competent Experimentalist Principle) that S_z was $+1/2$ just before the measurement, we may be tempted to conclude, via a suitable exercise of imagination, that if Charlie had instead measured S_x it too would have had a definite value, $+1/2$ or $-1/2$. Thus we conclude that a particle must always have not only a definite S_z value, the one the measurement revealed, but also a definite S_x value which the counterfactual measurement would have revealed.

◦ The general idea of a *counterfactual* argument is to imagine a world similar to the actual world but different in some specific respect(s). Then one argues that *if* the world had been different in this way, something *would* have happened. The appearance of "if. . . would" in an English sentence is often an indication that a counterfactual argument is underway.

★ To be sure, the Hilbert space is not big enough to hold both an S_z and an S_x value for a spin-half particle. So what should we do? There are various approaches:

• One can try and enlarge the Hilbert space by adding "hidden variables." Approaches of this sort have been tried. The best known is due to de Broglie and Bohm, and leads to various odd results, such as an instantaneous action-at-a-distance that is difficult to reconcile with relativity.

• One can also publish lengthy papers and even books about quantum paradoxes, thus keeping up interest in the field while never solving the outstanding problems.

• Or one can look for problems in the logical structure of counterfactual reasoning as applied to quantum systems. When this is done, see Ch. 19 of CQT, one finds that in cases where it gives rise to paradoxes, counterfactual reasoning always involves a violation of the single framework rule of quantum reasoning: incompatible sample spaces are being combined, or consistency conditions ignored.

◦ In the case at hand it is pretty obvious that the reasoning involves combining incompatible sample spaces, since this is the only way to reach the nonsensical conclusion that a given spin-half particle has both an S_x and an S_z value at the same time.

• Conclusion: Be careful when using counterfactuals in quantum reasoning. There are correct as well as incorrect approaches. See Ch. 19 of CQT for more details.

4 Partial measurements

4.1 Introduction

★ One is often interested in situations in which only *part* of a quantum system is measured, and one wants to know what this can tell us about the other parts. Measurements of this sort were made famous by the 1935 paper of Einstein, Podolsky, and Rosen (EPR) [2] and they arise quite frequently in discussions of quantum information.

• In 1951 David Bohm wrote a book [3] about quantum mechanics in which he discussed the issues raised by EPR using a model system of two spin-half particles in a singlet state, (18) below. Most discussions of EPR use Bohm's approach.

★ In 1964 John Bell [4] published the first version of the famous "Bell inequality," in which he showed that the statistical correlations which are predicted by quantum mechanics for the singlet state for various sorts of measurements cannot be understood in terms of classical "hidden variables" added to the quantum mechanical description.

◦ For a more detailed discussion of these points, see Chs. 23 and 24 of CQT.

4.2 Alice, Bob, and Charlie

★ The following material is taken (with a few alterations) from Sec. 3 of [5]

Charlie in Chicago takes two slips of paper, one red and one green, places them in two opaque envelopes, and after shuffling them so that he himself does not know which is which, addresses one to Alice in Atlanta and the other to Bob in Boston, both of whom know the protocol Charlie is following. Upon receipt of the envelope addressed to her Alice opens it and sees a red slip of paper. From this she can immediately conclude that the slip in Bob's envelope is green, whether or not Bob has already opened his envelope, or will ever do so. Her conclusion is not based on a belief that opening her envelope to "measure" the color of the slip of paper has some magical long-range influence on the color of Bob's slip. Instead it employs statistical reasoning in the following way.

Before Alice opens the envelope she (or Bob or Charlie) can assign probabilities to the various situations as follows:

$$\begin{aligned} &\Pr(A = R, B = R) = \Pr(A = G, B = G) = 0, \\ &\Pr(A = R, B = G) = \Pr(A = G, B = R) = 1/2, \end{aligned} \tag{16}$$

where $A = R$ means a red slip in Alice's envelope, $B = G$ a green slip in Bob's, etc., and Pr() refers to the joint probability distribution. From the usual rule for conditional probabilities, $\Pr(C \mid D) = \Pr(C, D)/\Pr(D)$ it follows that

$$\Pr(B = R \mid A = R) = 0, \quad \Pr(B = G \mid A = R) = 1, \tag{17}$$

and this is the conditional probability distribution that Alice uses to infer the color of Bob's slip knowing that the one in her envelope is red. One could say that she uses the outcome of her observation to "collapse" the initial probability distribution (16) onto the conditional probability distribution (17). The colors of the two slips of paper, the ontological situation in the physical world, is not at all affected by Alice's "measurement." It is her *knowledge* of the world that changes, in a way which we do not find at all surprising. The "collapse," if that is what one wishes to call it, refers to a method of reasoning, not a physical effect.

Next consider a situation in which Charlie at the center of the laboratory pushes a button, and one member of a pair of spin-half particles initially in the entangled state (18) is sent towards

Alice's apparatus at one end of the building, while the other is simultaneously sent towards Bob's apparatus at the other end. If Alice measures the z component of spin of her particle and the apparatus indicates $A = 1$ corresponding to $S_z = 1/2$ before the measurement took place—let us assume that Alice is a competent experimentalist who has designed and built a piece of apparatus which can do a measurement of this sort[1]—what can she say about S_z for Bob's particle? The chain of probabilistic inference is identical to that discussed earlier for colored slips of paper, though, as we shall see, certain details must be discussed with greater care.

4.3 Two spin-half particles in a singlet state

★ Imagine two spin-half particles a and b located some distance away from each other and prepared in an entangled state

$$|\psi_0\rangle = \big(|z_a^+, z_b^-\rangle - |z_a^-, z_b^+\rangle\big)/\sqrt{2} \tag{18}$$

at time t_0. Here $|z_a^+, z_b^-\rangle$ is the tensor product, which could also be written $|z^+\rangle_a \otimes |z^-\rangle_b$, whose significance is that $S_{az} = +1/2$ for particle a and $S_{bz} = -1/2$ for particle b. Of course, $|\psi_0\rangle$ is an entangled superposition state, so it does not correspond to a definite property of either a or b.

∘ The state (18) is referred to as the *singlet* state because it is the unique (up to an overall state) state in the Hilbert space of two spin-half particles with total angular momentum quantum number $j = 0$. There are in addition three *triplet* states, $j = 1$, which together with the singlet form a basis of the four-dimensional tensor-product Hilbert space.

• Now suppose that at time t_1 particle a begins to pass through a Stern-Gerlach apparatus at time t_1, one designed to detect single particles emerging from the region with a field gradient, so that at time t_2 a macroscopic measurement outcome has been registered: either P_a^+ corresponding to $[z_a^+]$ or P_a^- corresponding to $[z_a^-]$. What can we say about the spin of particle b at time t_1 or at time t_2? Assume that it is traveling in a region free of magnetic fields, so there is no tendency of the spin to precess.

∘ While measurements of this sort are not inconceivable with modern techniques, they remain very difficult. However, it is relatively easy to carry out analogous measurements of the polarizations of two photons prepared with their polarizations in an entangled state analogous to (18), and such measurements are nowadays routine. There is ample evidence from these experiments to support the existence of correlations of the photon polarizations which are the analogs of those discussed here for spin-half particles.

★ In particular, assume that the measurement on particle a results in P_a^- corresponding to $[z_a^-]$. What can we say about particle b?

• The first and obvious answer is that particle b is in the state $|z_b^+\rangle$. One can come to this conclusion by "looking at" the right side of (18) and noticing that $|z^-\rangle$ for particle a is clearly paired with $|z^+\rangle$ for particle b in the second term.

★ Whereas experts may be able to discern the right answer by simply looking at the ket representing the entangled state, a safer procedure for the less experienced is to employ an orthonormal basis

$$|z_a^+, z_b^+\rangle, \quad |z_a^+, z_b^-\rangle, \quad |z_a^-, z_b^+\rangle, \quad |z_a^-, z_b^-\rangle \tag{19}$$

[1] In talks given by experimental particle physicists one hears all sorts of references to trajectories and spins of elementary particles before they are measured. That this talk is perfectly justified from the perspective of quantum mechanics has been known for a long time; see Chs. 17 and 18 of CQT. Competent experimentalists ignore the nonsense they were taught in their introductory quantum mechanics courses about measurements producing results out of nowhere. The quantum foundations community should do the same.

for the two particles and use the Born rule to calculate at time t_1, before any measurement takes place, the joint probabilities

$$\Pr(z_a^+, z_b^-) = 1/2 = \Pr(z_a^-, z_b^+), \quad \Pr(z_a^+, z_b^+) = 0 = \Pr(z_a^-, z_b^-). \tag{20}$$

From this it follows that

$$\Pr([z_b^+] \mid [z_a^-]) = 1, \tag{21}$$

consistent with the previous conclusion obtained by looking at (18).

• Then we invoke CEP: The outcome P_a^- at t_2, after the measurement is complete, means the a particle was earlier in the state $[z_a^-]$, and if it was in the state $[z_a^-]$, then that particle b was in the state $[z_b^+]$ is a consequence of (21).

☐ Exercise. Work through the analogy with the colored slips of paper in the narrative in Sec. 4.2. What corresponds to $[z_a^-]$? to $[z_b^+]$?

• This conclusion about the state of particle b applies at other times as well, as long as we assume that particle b is isolated in a region of zero magnetic field, so it is not precessing, and is also not interacting with anything else. To see this one can set up a family of histories in which the decomposition $\{[z_b^+], [z_b^-]\}$ is inserted at various different times. We leave that as an exercise.

☐ Exercise. Choose a time $t_{0.5}$ between t_0 and t_1 and at this time use $\{[z_b^+], [z_b^-]\}$ as a decomposition of the identity. At t_1 use the projectors corresponding to the basis (19), and show that (21) also works for $\Pr([z_b^+]_{0.5} \mid [z_a^-]_1)$, where we have used subscripts to refer to the different times.

★ One can rewrite (18) using the S_{ax} and S_{bx} bases in the form

$$|\psi_0\rangle = \big(|x_a^+, x_b^-\rangle - |x_a^-, x_b^+\rangle\big)/\sqrt{2} \tag{22}$$

up to an overall phase, which depends upon the precise way in which the x and z bases are related, but does not affect the physics.

• Given the similarity between (22) and (18) it is obvious that the above discussion extends to the case in which the z-basis kets in (20) are replaced by their x-basis counterparts; in particular (20) and (21) can be replaced with

$$\Pr(x_a^+, x_b^-) = 1/2 = \Pr(x_a^-, x_b^+), \quad \Pr(x_a^+, x_b^+) = 0 = \Pr(x_a^-, x_b^-), \quad \Pr([x_b^+] \mid [x_a^-]) = 1. \tag{23}$$

• And, as a consequence, a later measurement by Alice of S_{ax} which yields a value corresponding to $[x_a^-]$ at a time just before the measurement took place allows her to infer the state $[x_b^+]$ for particle b.

★ It is important to notice that what Alice, a competent experimentalist, learns from her measurement is a property possessed by the particle prior to the measurement. Using her knowledge of this property and of the initial (singlet) state of the two particles she can, by setting up an appropriate probabilistic framework, make inferences about the state of particle b.

• These inferences do not depend upon measurements which may or may not be carried out on particle b, though of course Alice's ability to infer something depends on assuming that particle b is moving freely and has not interacted with something else.

• However, Alice can use her knowledge of particle b gained from her measurement of particle a to assign probabilities to the outcomes of possible measurements that might be carried out on particle b. And in some cases these probabilities are 1 or 0, allowing Alice to say something quite definite.

• That she can carry out different types of measurement, say S_{ax} or S_{az}, means that Alice can acquire different sorts of information about particle a. And her ability to make (probabilistic) statements about particle b depends on the type of information she acquires. But her choice of measurement has absolutely no effect upon particle b, contrary to the misleading statements found in much of the literature on the subject and also in some textbooks.

4.4 Einstein, Podolsky, Rosen, and Bell

★ We have now reached the heart of the EPR paradox as formulated by Bohm. Einstein, Podolsky, and Rosen argued, in effect, that because Alice can use a measurement, along with her knowledge of the initial state (18), to infer the value of S_{bz} for particle b, or a different measurement to infer the value of S_{bx}, therefore particle b must simultaneously possess both a z and an x component of spin angular momentum.

• From this they concluded that there must be something incomplete in the formulation of quantum mechanics that existed at that time (1935).

★ The key point from the modern perspective is that Alice cannot measure both S_{az} and S_{ax}, for the very good reason that there is nothing there to be measured!

• What was missing from standard quantum mechanics in 1935, and is still (unfortunately) missing from standard textbooks, was an adequate understanding of the quantum measuring process. Because that process had been formulated in a "wave function collapse" manner (see Sec. 4.6 below) which did not allow one to work out the correlation between the measurement outcome and the prior state the apparatus was designed to measure, the real difficulty, the structure of the quantum Hilbert space that prevents simultaneous discussion of incompatible properties, was not obvious. One can think of EPR as using an indirect approach—measuring properties of a distant particle b by measurements carried out on a—to unearth the problem associated with trying to interpret quantum theory using "measurement" as a sort of axiom, rather than using a consistent probabilistic formulation of quantum theory capable of analyzing measurements in terms of fundamental quantum principles.

★ As for Bell's inequality, see the discussion in Ch. 24 of CQT. The fact that it is violated by quantum mechanics, and this violation is extensively supported by experimental measurements, demonstrates the inadequacy of hidden variables approaches to quantum mechanics, in which one attempts to supplement the quantum Hilbert space with some mathematical structure which possesses a more "classical" behavior.

◦ Claims continue to be made that the violation of Bell's inequality by quantum mechanics means that the quantum world is in some sense "nonlocal." For a recent discussion of quantum (non)locality, see [5, 6].

4.5 General case of two entangled particles

★ Let us extend the discussion of Sec. 4.3 to the case of two particles a and b, with Hilbert spaces $\mathcal{H}_a$ and $\mathcal{H}_b$ of arbitrary (finite) dimension, not necessarily equal to each other, with an initial entangled state $|\psi_0\rangle$. Given an orthonormal basis $\{|a^j\rangle\}$ of $\mathcal{H}_a$, such a state can always be written in the form

$$|\psi_0\rangle = \sum_j |a^j\rangle \otimes |\beta^j\rangle. \tag{24}$$

Here the expansion coefficients, the $|\beta_j\rangle$, are determined by $|\psi_0\rangle$. In general they are not normalized, and nor are they orthogonal to each other. (The case in (18) is a bit exceptional in this latter respect.)

☐ Exercise. Convince yourself that the expansion (24) will work for any $|\psi_0\rangle$ and any choice of orthonormal basis $\{|a^j\rangle\}$. One way to do this is to expand $|\psi_0\rangle$ in a product of bases $|a^j\rangle \otimes |b^k\rangle$.

☐ Exercise. Show using the Born rule that if $|\psi_0\rangle$ in (24) is regarded as a (normalized) pre-probability, then $\Pr([a^j]) = \langle\beta^j|\beta^j\rangle$.

☐ Exercise. Show that the reduced density operator for particle b, given that $|\psi_0\rangle$ in (24) is normalized, is

$$\rho_b = \sum_j |\beta^j\rangle\langle\beta^j| = \sum_j p_j[\bar{\beta}^j], \text{ where } p_j := \langle\beta^j|\beta^j\rangle. \tag{25}$$

★ Next introduce a decomposition of the identity at time t_1 of the form

$$\{[a^j] \otimes [\bar{\beta}^j], Q^0\}, \text{ where } Q^0 = I_a \otimes I_b - \sum_j [a^j] \otimes [\bar{\beta}^j], \text{ and } [\bar{\beta}^j] = |\beta^j\rangle\langle\beta^j|/\langle\beta^j|\beta^j\rangle. \tag{26}$$

• Now let a measurement using the basis $\{|a^j\rangle\}$, of the sort described in Sec. 3.2, be carried out between t_1 and t_2 on particle a. Using the CEP we can infer from the outcome of the measurement P^j at t_2 that at t_1 particle a was in the state $[a^j]$, and if we employ the decomposition (26), we can infer the state $[\bar{\beta}^j]$ for particle b.

★ But there is a worry. How can we discuss which state particle b is in when the $[\bar{\beta}^j]$ for different j are in general not orthogonal to each other? Will this not violate the single framework rule?

• The proper way to think about this is that while the different $[\bar{\beta}^j]$ are not orthogonal, the $[a^j] \otimes [\bar{\beta}^j]$ *are* orthogonal. Think of the property $[\bar{\beta}^j]$ as being tagged by $[a^j]$ or by the measurement outcome P^j. Given a particular measurement outcome—we know the pointer position j—we can then assign a property to particle b, or use $[\bar{\beta}^j]$ as a pre-probability to calculate the probability of some property of particle b associated with a decomposition of I_b.

◦ We are dealing here with an example of a *dependent* or *contextual* property in the language of CQT Ch. 14, which see for a more detailed discussion.

• A straightforward analysis of the family

$$[\Psi_0] \odot \{[a^j] \otimes [\bar{\beta}^j], Q^0\} \odot \{P^k\}, \text{ with } |\Psi_0\rangle = |\psi_0\rangle \otimes |M_0\rangle, \tag{27}$$

for times $t_0 < t_1 < t_2$, where we assume that the time evolution of the measuring device interacting with particle a in the time interval from t_1 to t_2 is given by the appropriate analog of (6), shows that

$$\Pr([\bar{\beta}^j]_1 \mid P_2^j) = 1. \tag{28}$$

This confirms our earlier conclusion, that from outcome P^j at time t_2 of the measurement on particle a we can infer the state $[\bar{\beta}^j]$ of particle b at time t_1.

★ It is worth noting that as long as particle b remains undisturbed during the time intervals of interest, (28) remains valid even when $[\bar{\beta}^j]$ refers to particle b at a time later than the measurement outcome P^j.

☐ Exercise. Work this out by extending the family in (27) to a time $t_3 > t_2$, with a decomposition of the identity $\{I_a \otimes [\bar{\beta}^j] \otimes P^j\}$ at t_3, assuming that the unitary time transformation on particle b is trivial (I_b) at all times after t_0.

4.6 Wave function collapse

★ The analysis given in the previous section, where a state $[\bar{\beta}^j]$ is assigned to particle b on the basis of the outcome of a measurement on particle a corresponds closely to what is often called "wave function collapse".

• The idea of wave function collapse goes back to von Neumann's attempt [1] to understand quantum dynamics in terms of measurements. A measurement is a process that somehow "collapses" a wave function so that after the measurement takes place the measured particle is assigned a wave function corresponding to the measurement outcome. This proposal has caused an enormous amount of confusion in quantum foundations.

★ The reader is advised to think of wave function collapse as a *calculational procedure* for computing a conditional probability, such as the one in (28). Confusion comes about when the calculational procedure is thought of as some sort of physical process by which a measurement on particle a changes the physical state of particle b.

★ The idea that such a physical influence exists has given rise to all sorts of nonsense. In particular it has been claimed that quantum mechanics violates the theory of (special) relativity in cases in which particles a and b are far quite far apart when the measurement is carried out, because the measurement on a leads to an instantaneous change in the state of particle b.

○ See the discussion in [5, 6].

5 Preparations and Nondestructive Measurements

5.1 Preparations

★ The idea of *preparing* a quantum state is closely related to that of measurement, but there are some differences. The basic idea is that if a particle has been prepared in some state there is some sort of macroscopic indication that this is the case.

• The Stern-Gerlach apparatus described in Sec. 2 can be used as a preparation device if one imagines that there is a small hole placed at the output such that, for example, only the particles with $S_z = +1/2$, only the upper beam in Fig. 1, emerges from the apparatus. This produces a beam of polarized spin-half particles.

★ If two particles are known to be in an entangled state $|\psi_0\rangle$ of the sort discussed in Sec. 4 and one of them is measured, the measurement outcome will tell one the state of the other particle, and this can be considered a "preparation."

• But suppose the measurement outcome does not give one the desired result? Throw away the b particle in that case and repeat the experiment until the result on particle a gives what one desires.

★ But how does one prepare the state $|\psi_0\rangle$ in the first place, and how does one *know* that it is in that state?

• One strategy is illustrated by the process of photon down-conversion. An intense laser beam of light with frequency 2ω is sent through a suitable nonlinear crystal. Under the right circumstances this can produce a pair of photons of frequency ω emerging from the crystal with their polarizations in a correlated state that is the (or at least a) photon analog of the singlet state (18). Measurements are then carried out to determine how the polarizations of the photons are correlated. Due to the prodigious number of pairs which can be produced in a short time, a large number of measurements are possible, and if these are what one expects for photon polarizations in the desired state, one can be reasonably confident that, as long as the parameters governing the experimental setup are not changed, the next photon pair that emerges will be in this same quantum state.

5.2 Nondestructive Measurements

★ A *nondestructive measurement* as that term is used here means one in which the particle possess the same property, a property the apparatus was designed to (and does) measure, after the measurement as it did before the measurement.

• A Stern-Gerlach setup in which one imagines detecting whether the particle emerges in the upper or lower beam in Fig. 1, but without disturbing the spin of the particle, is an example of a nondestructive measurement.

○ This may be difficult to actually carry out in the laboratory, but it does not violate the laws of quantum mechanics, so is potentially possible.

• As this example indicates, a measurement which is nondestructive for some property, in this case S_z, can be very destructive for a different property, say S_x. Thus "nondestructive" has to be used relative to a specific set of (compatible) properties, i.e., some (projective) decomposition of the identity.

★ Nondestructive measurements have played a large role in discussions of the foundations of quantum mechanics because von Neumann uses them in his famous book [1]. Indeed, it is the only type of measurement which he considers, and it is connected with his idea of wave function collapse as discussed in Sec. 4.6 above.

• The measurement model discussed in Sec. 3.2 becomes nondestructive for the collection of properties $\{[a^j]\}$ if one assumes that $|s^j\rangle = |a^j\rangle$ for every j. If fact this is what is assumed in von Neumann's measurement model.

★ Another way to imagine a nondestructive measurement is that one first carries out an ordinary measurement (of the ideal sort discussed in Sec. 3), and then uses the resulting (macroscopic) outcome as the input to a preparation device which prepares a particle identical to the measured particle in a state which is the same as the measured state.

References

[1] Johann von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Springer-Verlag, Berlin, 1932. English translation: *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton (1955).

[2] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, 1935.

[3] David Bohm. *Quantum Theory*. Prentice Hall, Englewood Cliffs, N.J., 1951.

[4] John S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1:195–200, 1964. Reprinted in John S. Bell,*Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, 1987), p. 14.

[5] Robert B. Griffiths. Quantum locality. *Found. Phys.*, 41:705–733, 2011. arXiv:0908.2914.

[6] Robert B. Griffiths. EPR, Bell, and Quantum Locality. *Am. J. Phys.*, 79:954–965, 2011. arXiv:1007.4281.