Chapter 4

Quantum Entanglement

4.1 Nonseparability of EPR pairs

4.1.1 Hidden quantum information

The deep ways that quantum information differs from classical information involve the properties, implications, and uses of *quantum entanglement.* Recall from §2.4.1 that a bipartite pure state is *entangled* if its Schmidt number is greater than one. Entangled states are interesting because they exhibit correlations that have no classical analog. We will begin the study of these correlations in this chapter.

Recall, for example, the *maximally entangled* state of two qubits defined in §3.4.1:

$$|\phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB}). \tag{4.1}$$

"Maximally entangled" means that when we trace over qubit B to find the density operator $\boldsymbol{\rho}_A$ of qubit A, we obtain a multiple of the identity operator

$$\boldsymbol{\rho}_A = tr_B(|\phi^+\rangle_{AB}\ {}_{AB}\langle\phi^+) = \frac{1}{2}\mathbf{1}_A, \tag{4.2}$$

(and similarly $\boldsymbol{\rho}_B = \frac{1}{2}\mathbf{1}_B$). This means that if we measure spin A along *any* axis, the result is completely random – we find spin up with probability 1/2 and spin down with probability 1/2. Therefore, if we perform any local measurement of A or B, we acquire no information about the preparation of the state, instead we merely generate a random bit. This situation contrasts sharply with case of a single qubit in a pure state; there we can store a bit by preparing, say, either $|\uparrow_{\hat{n}}\rangle$ or $|\downarrow_{\hat{n}}\rangle$, and we can recover that bit reliably by measuring along the $\hat{n}$-axis. With two qubits, we ought to be able to store two bits, but in the state $|\phi^+\rangle_{AB}$ this information is *hidden*; at least, we can't acquire it by measuring A or B.

In fact, $|\phi^+\rangle$ is one member of a basis of four mutually orthogonal states for the two qubits, all of which are maximally entangled — the basis

$$\begin{aligned} |\phi^\pm\rangle &= \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \\ |\psi^\pm\rangle &= \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle), \end{aligned} \tag{4.3}$$

introduced in §3.4.1. We can choose to prepare one of these four states, thus encoding two bits in the state of the two-qubit system. One bit is the *parity* bit ($|\phi\rangle$ or $|\psi\rangle$) – are the two spins aligned or antialigned? The other is the *phase* bit ($+$ or $-$) – what superposition was chosen of the two states of like parity. Of course, we can recover the information by performing an orthogonal measurement that projects onto the $\{|\phi^+\rangle, |\phi^-\rangle, |\psi^+\rangle, |\psi^-\rangle\}$ basis. But if the two qubits are distantly separated, we cannot acquire this information locally; that is, by measuring A or measuring B.

What we *can* do locally is *manipulate* this information. Suppose that Alice has access to qubit A, but not qubit B. She may apply $\boldsymbol{\sigma}_3$ to her qubit, flipping the relative phase of $|0\rangle_A$ and $|1\rangle_A$. This action flips the phase bit stored in the entangled state:

$$\begin{aligned} |\phi^+\rangle &\leftrightarrow |\phi^-\rangle, \\ |\psi^+\rangle &\leftrightarrow |\psi^-\rangle. \end{aligned} \tag{4.4}$$

On the other hand, she can apply $\boldsymbol{\sigma}_1$, which flips her spin ($|0\rangle_A \leftrightarrow |1\rangle_A$), and also flips the parity bit of the entangled state:

$$\begin{aligned} |\phi^+\rangle &\leftrightarrow |\psi^+\rangle, \\ |\phi^-\rangle &\leftrightarrow -|\psi^-\rangle. \end{aligned} \tag{4.5}$$

Bob can manipulate the entangled state similarly. In fact, as we discussed in §2.4, either Alice or Bob can perform a local unitary transformation that changes one maximally entangled state to any other maximally entangled state.[1] What their local unitary transformations *cannot* do is alter $\boldsymbol{\rho}_A = \boldsymbol{\rho}_B = \frac{1}{2}\mathbf{1}$ – the information they are manipulating is information that neither one can read.

But now suppose that Alice and Bob are able to exchange (classical) messages about their measurement outcomes; together, then, they can learn about how their measurements are correlated. The entangled basis states are conveniently characterized as the simultaneous eigenstates of two commuting observables:

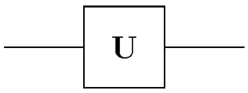
$$\begin{aligned} &\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)}, \\ &\boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)}; \end{aligned} \tag{4.6}$$

the eigenvalue of $\boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)}$ is the parity bit, and the eigenvalue of $\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)}$ is the phase bit. Since these operators commute, they can in principle be measured simultaneously. But they cannot be measured simultaneously if Alice and Bob perform localized measurements. Alice and Bob could both choose to measure their spins along the z-axis, preparing a simultaneous eigenstate of $\boldsymbol{\sigma}_3^{(A)}$ and $\boldsymbol{\sigma}_3^{(B)}$. Since $\boldsymbol{\sigma}_3^{(A)}$ and $\boldsymbol{\sigma}_3^{(B)}$ both commute with the parity operator $\boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)}$, their orthogonal measurements do not disturb the parity bit, and they can combine their results to infer the parity bit. But $\boldsymbol{\sigma}_3^{(A)}$ and $\boldsymbol{\sigma}_3^{(B)}$ do *not* commute with phase operator $\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)}$, so their measurement disturbs the phase bit. On the other hand, they could both choose to measure their spins along the x-axis; then they would learn the phase bit at the cost of disturbing the parity bit. But they can't have it both ways. To have hope of acquiring the parity bit without disturbing the phase bit, they would need to learn about the product $\boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)}$ without finding out anything about $\boldsymbol{\sigma}_3^{(A)}$ and $\boldsymbol{\sigma}_3^{(B)}$ separately. That cannot be done locally.

Now let us bring Alice and Bob together, so that they can operate on their qubits jointly. How might they acquire both the parity bit and the phase bit of their pair? By applying an appropriate unitary transformation, they can rotate the entangled basis $\{|\phi^\pm\rangle, |\psi^\pm\rangle\}$ to the unentangled basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Then they can measure qubits A and B separately to acquire the bits they seek. How is this transformation constructed?

[1]But of course, this does not suffice to perform an arbitrary unitary transformation on the four-dimensional space $\mathcal{H}_A \otimes \mathcal{H}_B$, which contains states that are not maximally entangled. The maximally entangled states are *not* a subspace – a superposition of maximally entangled states typically is *not* maximally entangled.

This is a good time to introduce notation that will be used heavily later in the course, the quantum circuit notation. Qubits are denoted by horizontal lines, and the single-qubit unitary transformation $\mathbf{U}$ is denoted:



A particular single-qubit unitary we will find useful is the *Hadamard transform*

$$\mathbf{H} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}}(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_3), \tag{4.7}$$

which has the properties

$$\mathbf{H}^2 = \mathbf{1}, \tag{4.8}$$

and

$$\begin{aligned} \mathbf{H}\boldsymbol{\sigma}_1\mathbf{H} &= \boldsymbol{\sigma}_3, \\ \mathbf{H}\boldsymbol{\sigma}_3\mathbf{H} &= \boldsymbol{\sigma}_1. \end{aligned} \tag{4.9}$$

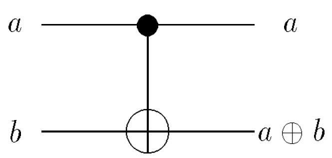
(We can envision $\mathbf{H}$ (up to an overall phase) as a $\theta = \pi$ rotation about the axis $\hat{n} = \frac{1}{\sqrt{2}}(\hat{n}_1 + \hat{n}_3)$ that rotates $\hat{x}$ to $\hat{z}$ and vice-versa; we have

$$R(\hat{n}, \theta) = \mathbf{1}\cos\frac{\theta}{2} + i\hat{n}\cdot\vec{\boldsymbol{\sigma}}\sin\frac{\theta}{2} = i\frac{1}{\sqrt{2}}(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_3) = i\mathbf{H}.) \tag{4.10}$$

Also useful is the two-qubit transformation known as the XOR or controlled-NOT transformation; it acts as

$$\mathbf{CNOT} : |a, b\rangle \rightarrow |a, a \oplus b\rangle, \tag{4.11}$$

on the basis states $a, b = 0, 1$, where $a \oplus b$ denotes addition modulo 2, and is denoted:

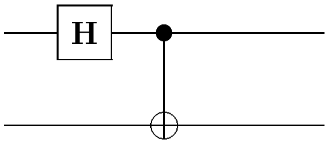


Thus this gate flips the second bit if the first is 1, and acts trivially if the first bit is 0; we see that

$$(\mathbf{CNOT})^2 = \mathbf{1}. \tag{4.12}$$

We call a the *control* (or source) bit of the **CNOT**, and b the *target* bit.

By composing these "primitive" transformations, or quantum *gates*, we can build other unitary transformations. For example, the "circuit"



(to be read from left to right) represents the product of **H** applied to the first qubit followed by **CNOT** with the first bit as the source and the second bit as the target. It is straightforward to see that this circuit transforms the standard basis to the entangled basis,

$$\begin{aligned}
|00\rangle &\to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle \to |\phi^+\rangle,\\
|01\rangle &\to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|1\rangle \to |\psi^+\rangle,\\
|10\rangle &\to \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)|0\rangle \to |\phi^-\rangle,\\
|11\rangle &\to \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)|1\rangle \to |\psi^-\rangle,
\end{aligned} \tag{4.13}$$

so that the first bit becomes the phase bit in the entangled basis, and the second bit becomes the parity bit.

Similarly, we can invert the transformation by running the circuit backwards (since both **CNOT** and **H** square to the identity); if we apply the inverted circuit to an entangled state, and then measure both bits, we can learn the value of both the phase bit and the parity bit.

Of course, **H** acts on only one of the qubits; the "nonlocal" part of our circuit is the controlled-NOT gate – this is the operation that establishes or removes entanglement. If we could only perform an "interstellar **CNOT**," we would be able to create entanglement among distantly separated pairs, or extract the information encoded in entanglement. But we can't. To do its job, the **CNOT** gate must act on its target without revealing the value of its source. Local operations and classical communication will not suffice.

4.1.2 Einstein locality and hidden variables

Einstein was disturbed by quantum entanglement. Eventually, he along with Podolsky and Rosen sharpened their discomfort into what they regarded as a paradox. As later reinterpreted by Bohm, the situation they described is really the same as that discussed in §2.5.3. Given a maximally entangled state of two qubits shared by Alice and Bob, Alice can choose one of several possible measurements to perform on her spin that will realize different possible ensemble interpretations of Bob's density matrix; for example, she can prepare either $\boldsymbol{\sigma}_1$ or $\boldsymbol{\sigma}_3$ eigenstates.

We have seen that Alice and Bob are unable to exploit this phenomenon for faster-than-light communication. Einstein knew this but he was still dissatisfied. He felt that in order to be considered a *complete* description of physical reality a theory should meet a stronger criterion, that might be called Einstein locality:

> Suppose that A and B are spacelike separated systems. Then in a *complete* description of physical reality an action performed on system A must not modify the description of system B.

But if A and B are entangled, a measurement of A is performed and a *particular* outcome is known to have been obtained, then the density matrix of B *does* change. Therefore, by Einstein's criterion, the description of a quantum system by a wavefunction cannot be considered complete.

Einstein seemed to envision a more complete description that would remove the indeterminacy of quantum mechanics. A class of theories with this feature are called *local hidden-variable theories*. In a hidden variable theory, measurement is actually fundamentally deterministic, but appears to be probabilistic because some degrees of freedom are not precisely known. For example, perhaps when a spin is prepared in what quantum theory would describe as the pure state $|\uparrow_{\hat{z}}\rangle$, there is actually a deeper theory in which the state prepared is parametrized as $(\hat{z}, \lambda)$ where λ $(0 \leq \lambda \leq 1)$ is the hidden variable. Suppose that with present-day experimental technique, we have no control over λ, so when we prepare the spin state, λ might take any value – the probability distribution governing its value is uniform on the unit interval.

Now suppose that when we measure the spin along an axis rotated by θ from the $\hat{z}$ axis, the outcome will be

$$\begin{aligned} &|\uparrow_\theta\rangle, \text{ for } 0 \leq \lambda \leq \cos^2\frac{\theta}{2} \\ &|\downarrow_\theta\rangle, \text{ for } \cos^2\frac{\theta}{2} < \lambda \leq 1. \end{aligned} \tag{4.14}$$

If we know λ, the outcome is deterministic, but if λ is completely unknown, then the probability distribution governing the measurement will agree with the predictions of quantum theory.

Now, what about entangled states? When we say that a hidden variable theory is *local*, we mean that it satisfies the Einstein locality constraint. A measurement of A does not modify the values of the variables that govern the measurements of B. This seems to be what Einstein had in mind when he envisioned a more complete description.

4.1.3 Bell Inequalities

John Bell's fruitful idea was to test Einstein locality by considering the quantitative properties of the correlations between measurement outcomes obtained by Bob and Alice.[2] Let's first examine the predictions of quantum mechanics regarding these correlations.

Note that the state $|\psi^-\rangle$ has the properties

$$(\vec{\boldsymbol{\sigma}}^{(A)} + \vec{\boldsymbol{\sigma}}^{(B)})|\psi^-\rangle = 0, \tag{4.15}$$

as we can see by explicit computation. Now consider the expectation value

$$\langle\phi^-|(\vec{\boldsymbol{\sigma}}^{(A)}\cdot\hat{n})(\vec{\boldsymbol{\sigma}}^{(B)}\cdot\hat{m})|\psi^-\rangle. \tag{4.16}$$

Since we can replace $\vec{\boldsymbol{\sigma}}^{(B)}$ by $-\vec{\boldsymbol{\sigma}}^{(A)}$ acting on $|\psi^-\rangle$, this can be expressed as

$$\begin{aligned} &-\langle(\vec{\boldsymbol{\sigma}}^{(A)}\cdot\hat{n})(\vec{\boldsymbol{\sigma}}^{(A)}\cdot\hat{m})\rangle = \\ &-n_i m_j tr(\boldsymbol{\rho}_A \boldsymbol{\sigma}_i^{(A)}\boldsymbol{\sigma}_j^{(A)}) = -n_i m_j \delta_{ij} = -\hat{n}\cdot\hat{m} = -\cos\theta, \end{aligned} \tag{4.17}$$

[2]A good reference on Bell inequalities is A. Peres, *Quantum Theory: Concepts and Methods*, chapter 6.

where θ is the angle between the axes $\hat{n}$ and $\hat{m}$. Thus we find that the measurement outcomes are always perfectly anticorrelated when we measure both spins along the same axis $\hat{n}$, and we have also obtained a more general result that applies when the two axes are different. Since the projection operator onto the spin up (spin down) states along $\hat{n}$ is $\mathbf{E}(\hat{n}, \pm) = \frac{1}{2}(\mathbf{1} \pm \hat{n} \cdot \vec{\boldsymbol{\sigma}})$, we also obtain

$$\begin{aligned}
&\langle \psi^- | \mathbf{E}^{(A)}(\hat{n}, +) \mathbf{E}^{(B)}(\hat{m}, +) | \psi^- \rangle \\
&= \langle \psi^- | \mathbf{E}^{(A)}(\hat{n}, -) \mathbf{E}^{(B)}(\hat{m}, -) | \psi^- \rangle = \frac{1}{4}(1 - \cos\theta), \\
&\langle \psi^- | \mathbf{E}^{(A)}(\hat{n}, +) \mathbf{E}^{(B)}(\hat{m}, -) | \psi^- \rangle \\
&= \langle \psi^- | \mathbf{E}^{(A)}(\hat{n}, -) \mathbf{E}^{(B)}(\hat{m}, +) | \psi - \rangle = \frac{1}{4}(1 + \cos\theta);
\end{aligned} \tag{4.18}$$

The probability that the outcomes are opposite is $\frac{1}{2}(1 + \cos\theta)$, and the probability that the outcomes are the same is $\frac{1}{2}(1 - \cos\theta)$.

Now suppose Alice will measure her spin along one of the three axes in the $x - z$ plane,

$$\begin{aligned}
\hat{n}_1 &= (0, 0, 1) \\
\hat{n}_2 &= \left(\frac{\sqrt{3}}{2}, 0, -\frac{1}{2}\right) \\
\hat{n}_3 &= \left(-\frac{\sqrt{3}}{2}, 0, -\frac{1}{2}\right).
\end{aligned} \tag{4.19}$$

Once she performs the measurement, she disturbs the state of the spin, so she won't have a chance to find out what would have happened if she had measured along a different axis. Or will she? If she shares the state $|\psi^-\rangle$ with Bob, then Bob can help her. If Bob measures along, say, $\hat{n}_2$, and sends the result to Alice, then Alice knows what would have happened *if* she had measured along $\hat{n}_2$, since the results are perfectly anticorrelated. Now she can go ahead and measure along $\hat{n}_1$ as well. According to quantum mechanics, the probability that measuring along $\hat{n}_1$, and $\hat{n}_2$ give the *same* result is

$$P_{same} = \frac{1}{2}(1 - \cos\theta) = \frac{1}{4}. \tag{4.20}$$

(We have $\cos\theta = 1/2$ because Bob measures along $-\hat{n}_2$ to obtain Alice's result for measuring along $\hat{n}_2$). In the same way, Alice and Bob can work together to determine outcomes for the measurement of Alice's spin along any two of the axes $\hat{n}_1, \hat{n}_2$, and $\hat{n}_3$.

It is as though three coins are resting on a table; each coin has either the heads (H) or tails (T) side facing up, but the coins are covered, at first, so we don't know which. It is possible to reveal two of the coins (measure the spin along two of the axes) to see if they are H or T, but then the third coin always disappears before we get a chance to uncover it (we can't measure the spin along the third axis).

Now suppose that there are actually local hidden variables that provide a *complete* description of this system, and the quantum correlations are to arise from a probability distribution governing the hidden variables. Then, in this context, the Bell inequality is the statement

$$P_{same}(1,2) + P_{same}(1,3) + P_{same}(2,3) \geq 1, \tag{4.21}$$

where $P_{same}(i,j)$ denotes the probability that coins i and j have the *same* value (HH or TT). This is satisfied by any probability distribution for the three coins because no matter what the values of the coins, there will always be two that are the same. But in quantum mechanics,

$$P_{same}(1,2) + P_{same}(1,3) + P_{same}(2,3) = 3 \cdot \frac{1}{4} = \frac{3}{4} < 1. \tag{4.22}$$

We have found that the correlations predicted by quantum theory are incompatible with the local hidden variable hypothesis.

What are the implications? To some people, the peculiar correlations unmasked by Bell's theorem call out for a deeper explanation than quantum mechanics seems to provide. They see the EPR phenomenon as a harbinger of new physics awaiting discovery. But they may be wrong. We have been waiting over 60 years since EPR, and so far no new physics.

Perhaps we have learned that it can be dangerous to reason about what might have happened, but didn't actually happen. (Of course, we do this all the time in our everyday lives, and we usually get away with it, but sometimes it gets us into trouble.) I claimed that Alice knew what would happen when she measured along $\hat{n}_2$, because Bob measured along $\hat{n}_2$, and every time we have ever checked, their measurement outcomes are always perfectly anticorrelated. But Alice did *not* measure along $\hat{n}_2$; she measured along $\hat{n}_1$ instead. We got into trouble by trying to assign probabilities to the outcomes of measurements along $\hat{n}_1, \hat{n}_2$, and $\hat{n}_3$, even though we can only perform one of those measurements. This turned out to lead to mathematical inconsistencies, so we had better not do it. From this viewpoint we have affirmed Bohr's principle of *complementary* — we are forbidden to consider simultaneously the possible outcomes of two mutually exclusive experiments.

Another common attitude is that the violations of the Bell inequalities (confirmed experimentally) have exposed an essential nonlocality built into the quantum description of Nature. One who espouses this view has implicitly rejected the complementarity principle. *If* we do insist on talking about outcomes of mutually exclusive experiments *then* we are forced to conclude that Alice's choice of measurement actually exerted a subtle *influence* on the outcome of Bob's measurement. This is what is meant by the "nonlocality" of quantum theory.

By ruling out local hidden variables, Bell demolished Einstein's dream that the indeterminacy of quantum theory could be eradicated by adopting a more complete, yet still local, description of Nature. If we accept locality as an inviolable principle, then we are forced to accept randomness as an unavoidable and intrinsic feature of quantum measurement, rather than a consequence of incomplete knowledge.

The human mind seems to be poorly equipped to grasp the correlations exhibited by entangled quantum states, and so we speak of the weirdness of quantum theory. But whatever your attitude, experiment forces you to accept the existence of the weird correlations among the measurement outcomes. There is no big mystery about how the correlations were established – we saw that it was necessary for Alice and Bob to get together at some point to create entanglement among their qubits. The novelty is that, even when A and B are distantly separated, we cannot accurately regard A and B as two separate qubits, and use classical information to characterize how they are correlated. They are more than just correlated, they are a single *inseparable* entity. They are *entangled*.

4.1.4 Photons

Experiments that test the Bell inequality are done with entangled photons, not with spin$-\frac{1}{2}$ objects. What are the quantum-mechanical predictions for photons?

Suppose, for example, that an excited atom emits two photons that come out back to back, with vanishing angular momentum and even parity. If $|x\rangle$ and $|y\rangle$ are horizontal and vertical linear polarization states of the photon, then we have seen that

$$
\begin{aligned}
|+\rangle &= \frac{1}{\sqrt{2}}(|x\rangle + i|y\rangle), \\
|-\rangle &= \frac{1}{\sqrt{2}}(i|x\rangle + |y\rangle),
\end{aligned} \tag{4.23}
$$

are the eigenstates of *helicity* (angular momentum along the axis of propagation $\hat{z}$. For two photons, one propagating in the $+\hat{z}$ direction, and the other in the $-\hat{z}$ direction, the states

$$
\begin{aligned}
&|+\rangle_A|-\rangle_B \\
&|-\rangle_A|+\rangle_B
\end{aligned} \tag{4.24}
$$

are invariant under rotations about $\hat{z}$. (The photons have opposite values of J_z, but the *same* helicity, since they are propagating in opposite directions.) Under a reflection in the $y-z$ plane, the polarization states are modified according to

$$
\begin{aligned}
|x\rangle \to -|x\rangle, &\quad |+\rangle \to +i|-\rangle, \\
|y\rangle \to |y\rangle, &\quad |-\rangle \to -i|+\rangle;
\end{aligned} \tag{4.25}
$$

therefore, the parity eigenstates are *entangled* states

$$
\frac{1}{\sqrt{2}}(|+\rangle_A|-\rangle_B \pm |-\rangle_A|+\rangle_B). \tag{4.26}
$$

The state with $J_z = 0$ and even parity, then, expressed in terms of the linear polarization states, is

$$
\begin{aligned}
&-\frac{i}{\sqrt{2}}(|+\rangle_A|-\rangle_B + |-\rangle_A|+\rangle_B) \\
&= \frac{1}{\sqrt{2}}(|xx\rangle_{AB} + |yy\rangle_{AB})_n = |\phi^+\rangle_{AB}.
\end{aligned} \tag{4.27}
$$

Because of invariance under rotations about $\hat{z}$, the state has this form irrespective of how we orient the x and y axes.

We can use a polarization analyzer to measure the linear polarization of either photon along any axis in the xy plane. Let $|x(\theta)\rangle$ and $|y(\theta)\rangle$ denote the linear polarization eigenstates along axes rotated by angle θ relative to the canonical x and y axes. We may define an operator (the analog of $\vec{\boldsymbol{\sigma}} \cdot \hat{n}$)

$$\boldsymbol{\tau}(\theta) = |x(\theta)\rangle\langle x(\theta)| - |y(\theta)\rangle\langle y(\theta)|, \tag{4.28}$$

which has these polarization states as eigenstates with respective eigenvalues ± 1. Since

$$|x(\theta)\rangle = \begin{pmatrix} \cos & \theta \\ \sin & \theta \end{pmatrix}, \quad |y(\theta)\rangle = \begin{pmatrix} -\sin & \theta \\ \cos & \theta \end{pmatrix}, \tag{4.29}$$

in the $|x\rangle, |y\rangle$ basis, we can easily compute the expectation value

$$_{AB}\langle\phi^+|\boldsymbol{\tau}^{(A)}(\theta_1)\boldsymbol{\tau}^{(B)}(\theta_2)|\phi^+\rangle_{AB}. \tag{4.30}$$

Using rotational invariance:

$$\begin{aligned}
&= \ _{AB}\langle\phi^+|\boldsymbol{\tau}^{(A)}(0)\boldsymbol{\tau}^{(B)}(\theta_2-\theta_1)|\phi^+\rangle_{AB} \\
&= \frac{1}{2}{}_B\langle x|\boldsymbol{\tau}^{(B)}(\theta_2-\theta_1)|x\rangle_B - \frac{1}{2}{}_B\langle y|\boldsymbol{\tau}^{(B)}(\theta_2-\theta_1)|y\rangle_B \\
&= \cos^2(\theta_2-\theta_1) - \sin^2(\theta_2-\theta_1) = \cos[2(\theta_2-\theta_1)].
\end{aligned} \tag{4.31}$$

(For spin-$\frac{1}{2}$ objects, we would obtain

$$_{AB}\langle\phi^+|(\vec{\boldsymbol{\sigma}}^{(A)}\cdot\hat{n}_1)(\vec{\boldsymbol{\sigma}}^{(B)}\cdot\hat{n}_2) = \hat{n}_1\cdot\hat{n}_2 = \cos(\theta_2-\theta_1); \tag{4.32}$$

the argument of the cosine is different than in the case of photons, because the half angle $\theta/2$ appears in the formula analogous to eq. (4.29).)

4.1.5 More Bell inequalities

So far, we have considered only one (particularly interesting) case of the Bell inequality. Here we will generalize the result.

Consider a correlated pair of photons, A and B. We may choose to measure the polarization of photon A along either one of two axes, α or α'. The corresponding observables are denoted

$$\begin{aligned}
\boldsymbol{a} &= \boldsymbol{\tau}^{(A)}(\alpha) \\
\boldsymbol{a}' &= \boldsymbol{\tau}^{(A)}(\alpha').
\end{aligned} \tag{4.33}$$

Similarly, we may choose to measure photon B along either axis β or axis β'; the corresponding observables are

$$\begin{aligned} \boldsymbol{b} &= \boldsymbol{\tau}^{(B)}(\beta) \\ \boldsymbol{b} &= \boldsymbol{\tau}^{(B)}(\beta'). \end{aligned} \tag{4.34}$$

We will, to begin with, consider the special case $\alpha' = \beta' \equiv \gamma$.

Now, if we make the local hidden variable hypothesis, what can be infer about the correlations among these observables? We'll assume that the prediction of quantum mechanics is satisfied if we measure $\boldsymbol{a}'$ and $\boldsymbol{b}'$, namely

$$\langle \boldsymbol{a}'\boldsymbol{b}' \rangle = \langle \boldsymbol{\tau}^{(B)}(\gamma)\boldsymbol{\tau}^{(B)}(\gamma) \rangle = 1; \tag{4.35}$$

when we measure both photons along the same axes, the outcomes *always* agree. Therefore, these two observables have exactly the *same* functional dependence on the hidden variables – they are really the *same* observable, with we will denote $\boldsymbol{c}$.

Now, let $\boldsymbol{a}, \boldsymbol{b}$, and $\boldsymbol{c}$ be *any* three observables with the properties

$$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} = \pm 1; \tag{4.36}$$

i.e., they are functions of the hidden variables that take only the two values ± 1. These functions satisfy the identity

$$\boldsymbol{a}(\boldsymbol{b} - \boldsymbol{c}) = \boldsymbol{ab}(1 - \boldsymbol{bc}). \tag{4.37}$$

(We can easily verify the identity by considering the cases $\boldsymbol{b} - \boldsymbol{c} = 0, 2, -2$.) Now we take expectation values by integrating over the hidden variables, weighted by a nonnegative probability distribution:

$$\langle \boldsymbol{ab} \rangle - \langle \boldsymbol{ac} \rangle = \langle \boldsymbol{ab}(1 - \boldsymbol{bc}) \rangle. \tag{4.38}$$

Furthermore, since $\boldsymbol{ab} = \pm 1$, and $1 - \boldsymbol{bc}$ is nonnegative, we have

$$\begin{aligned} &|\langle \boldsymbol{ab}(1 - \boldsymbol{bc}) \rangle| \\ &\leq |\langle 1 - \boldsymbol{bc} \rangle| = 1 - \langle \boldsymbol{bc} \rangle. \end{aligned} \tag{4.39}$$

We conclude that

$$|\langle \boldsymbol{ab} \rangle - \langle \boldsymbol{ac} \rangle| \leq 1 - \langle \boldsymbol{bc} \rangle. \tag{4.40}$$

This is the Bell inequality.

To make contact with our earlier discussion, consider a pair of spin-$\frac{1}{2}$ objects in the state $|\phi^+\rangle$, where α, β, γ are separated by successive 60^o angles. Then quantum mechanics predicts

$$\begin{aligned} \langle \boldsymbol{ab} \rangle &= \frac{1}{2} \\ \langle \boldsymbol{bc} \rangle &= \frac{1}{2} \\ \langle \boldsymbol{ac} \rangle &= -\frac{1}{2}, \end{aligned} \tag{4.41}$$

which violates the Bell inequality:

$$1 = \frac{1}{2} + \frac{1}{2} \not\leq 1 - \frac{1}{2} = \frac{1}{2}. \tag{4.42}$$

For photons, to obtain the same violation, we halve the angles, so α, β, γ are separated by 30^o angles.

Return now to the more general case $\alpha' \neq \beta'$. We readily see that $\boldsymbol{a}, \boldsymbol{a}', \boldsymbol{b}, \boldsymbol{b}' = \pm 1$ implies that

$$(\boldsymbol{a} + \boldsymbol{a}')\boldsymbol{b} - (\boldsymbol{a} - \boldsymbol{a}')\boldsymbol{b}' = \pm 2, \tag{4.43}$$

(by considering the two cases $\boldsymbol{a} + \boldsymbol{a}' = 0$ and $\boldsymbol{a} - \boldsymbol{a}' = 0$), or

$$\langle \boldsymbol{ab} \rangle + \langle \boldsymbol{a}'\boldsymbol{b} \rangle + \langle \boldsymbol{a}'\boldsymbol{b}' \rangle - \langle \boldsymbol{ab}' \rangle = \langle \boldsymbol{\theta} \rangle, \tag{4.44}$$

where $\boldsymbol{\theta} = \pm 2$. Evidently

$$|\langle \boldsymbol{\theta} \rangle| \leq 2, \tag{4.45}$$

so that

$$|\langle \boldsymbol{ab} \rangle + \langle \boldsymbol{a}'\boldsymbol{b} \rangle + \langle \boldsymbol{a}'\boldsymbol{b}' \rangle - \langle \boldsymbol{ab}' \rangle| \leq 2. \tag{4.46}$$

This result is called the CHSH (Clauser-Horne-Shimony-Holt) inequality. To see that quantum mechanics violates it, consider the case for photons where $\alpha, \beta, \alpha', \beta'$ are separated by successive $22.5°$ angles, so that the quantum-mechanical predictions are

$$\begin{aligned} \langle \boldsymbol{ab} \rangle = \langle \boldsymbol{a}'\boldsymbol{b} \rangle = \langle \boldsymbol{a}'\boldsymbol{b}' \rangle &= \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}}, \\ \langle \boldsymbol{ab}' \rangle &= \cos\frac{3\pi}{4} = -\frac{1}{\sqrt{2}}, \end{aligned} \tag{4.47}$$

while

$$2\sqrt{2} \not\leq 2. \tag{4.48}$$

4.1.6 Maximal violation

We can see that the case just considered ($\alpha, \beta, \alpha', \beta'$ separated by successive 22.5^o angles) provides the largest possible quantum mechanical violation of the CHSH inequality. In quantum theory, suppose that $\boldsymbol{a}, \boldsymbol{a}', \boldsymbol{b}, \boldsymbol{b}'$ are observables that satisfy

$$\boldsymbol{a}^2 = \boldsymbol{a}'^2 = \boldsymbol{b}^2 = \boldsymbol{b}'^2 = \boldsymbol{1}, \tag{4.49}$$

and

$$0 = [\boldsymbol{a}, \boldsymbol{b}] = [\boldsymbol{a}, \boldsymbol{b}'] = [\boldsymbol{a}', \boldsymbol{b}] = [\boldsymbol{a}', \boldsymbol{b}']. \tag{4.50}$$

Let

$$\boldsymbol{C} = \boldsymbol{ab} + \boldsymbol{a}'\boldsymbol{b} + \boldsymbol{a}'\boldsymbol{b}' - \boldsymbol{ab}'. \tag{4.51}$$

Then

$$\boldsymbol{C}^2 = 4 + \boldsymbol{aba}'\boldsymbol{b}' - \boldsymbol{a}'\boldsymbol{bab}' + \boldsymbol{a}'\boldsymbol{b}'\boldsymbol{ab} - \boldsymbol{ab}'\boldsymbol{a}'\boldsymbol{b}. \tag{4.52}$$

(You can check that the other terms cancel)

$$= 4 + [\boldsymbol{a}, \boldsymbol{a}'][\boldsymbol{b}, \boldsymbol{b}']. \tag{4.53}$$

The *sup norm* $\parallel \mathbf{M} \parallel$ of a bounded operator $\mathbf{M}$ is defined by

$$\parallel \mathbf{M} \parallel = \begin{matrix} sup \\ |\psi\rangle \end{matrix} \left(\frac{\parallel \mathbf{M}|\psi\rangle \parallel}{\parallel |\psi\rangle \parallel} \right); \tag{4.54}$$

it is easy to verify that the sup norm has the properties

$$\begin{aligned} \parallel \mathbf{MN} \parallel &\leq \parallel \mathbf{M} \parallel \parallel \mathbf{N} \parallel, \\ \parallel \mathbf{M} + \mathbf{N} \parallel &\leq \parallel \mathbf{M} \parallel + \parallel \mathbf{N} \parallel, \end{aligned} \tag{4.55}$$

and therefore

$$\parallel [\mathbf{M}, \mathbf{N}] \parallel \leq \parallel \mathbf{MN} \parallel + \parallel \mathbf{NM} \parallel \leq 2 \parallel \mathbf{M} \parallel \parallel \mathbf{N} \parallel . \tag{4.56}$$

We conclude that

$$\| \boldsymbol{C}^2 \| \leq 4 + 4 \| \boldsymbol{a} \| \cdot \| \boldsymbol{a}' \| \cdot \| \boldsymbol{b} \| \cdot \| \boldsymbol{b}' \| = 8, \tag{4.57}$$

or

$$\| \boldsymbol{C} \| \leq 2\sqrt{2} \tag{4.58}$$

(Cirel'son's inequality). Thus, the expectation value of $\boldsymbol{C}$ cannot exceed $2\sqrt{2}$, precisely the value that we found to be attained in the case where $\alpha, \beta, \alpha', \beta'$ are separated by successive 22.5^o angles. The violation of the CHSH inequality that we found is the largest violation allowed by quantum theory.

4.1.7 The Aspect experiment

The CHSH inequality was convincingly tested for the first time by Aspect and collaborators in 1982. Two entangled photons were produced in the decay of an excited calcium atom, and each photon was directed by a switch to one of two polarization analyzers, chosen pseudo-randomly. The photons were detected about 12m apart, corresponding to a light travel time of about 40 ns. This time was considerably longer than either the cycle time of the switch, or the difference in the times of arrival of the two photons. Therefore the "decision" about which observable to measure was made after the photons were already in flight, and the events that selected the axes for the measurement of photons A and B were spacelike separated. The results were consistent with the quantum predictions, and violated the CHSH inequality by five standard deviations. Since Aspect, many other experiments have confirmed this finding.

4.1.8 Nonmaximal entanglement

So far, we have considered the Bell inequality violations predicted by quantum theory for a maximally entangled state such as $|\phi^+\rangle$. But what about more general states such as

$$|\phi\rangle = \alpha|00\rangle + \beta|11\rangle? \tag{4.59}$$

(Any pure state of two qubits can be expressed this way in the Schmidt basis; by adopting suitable phase conventions, we may assume that α and β are real and nonnegative.)

Consider first the extreme case of separable pure states, for which

$$\langle \boldsymbol{ab} \rangle = \langle \boldsymbol{a} \rangle \langle \boldsymbol{b} \rangle. \tag{4.60}$$

In this case, it is clear that no Bell inequality violation can occur, because we have already seen that a (local) hidden variable theory *does* exist that correctly reproduces the predictions of quantum theory for a pure state of a single qubit. Returning to the spin-$\frac{1}{2}$ notation, suppose that we measure the spin of each particle along an axis $\hat{n} = (\sin\theta, 0, \cos\theta)$ in the xz plane. Then

$$\begin{aligned} \boldsymbol{a} &= (\boldsymbol{\sigma}^{(A)} \cdot \hat{n}_1) = \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ \sin\theta_1 & -\cos\theta_1 \end{pmatrix}^{(A)}, \\ \boldsymbol{b} &= (\boldsymbol{\sigma}^{(B)} \cdot \hat{n}_2) = \begin{pmatrix} \cos\theta_2 & \sin\theta_2 \\ \sin\theta_2 & -\cos\theta_2 \end{pmatrix}^{(B)}, \end{aligned} \tag{4.61}$$

so that quantum mechanics predicts

$$\begin{aligned} \langle \boldsymbol{ab} \rangle &= \langle \phi | \boldsymbol{ab} | \phi \rangle \\ &= \cos\theta_1 \cos\theta_2 + 2\alpha\beta \sin\theta_1 \sin\theta_2 \end{aligned} \tag{4.62}$$

(and we recover $\cos(\theta_1 - \theta_2)$ in the maximally entangled case $\alpha = \beta = 1/\sqrt{2}$). Now let us consider, for simplicity, the (nonoptimal!) special case

$$\theta_A = 0, \quad \theta_A' = \frac{\pi}{2}, \quad \theta_B' = -\theta_B, \tag{4.63}$$

so that the quantum predictions are:

$$\begin{aligned} \langle \boldsymbol{ab} \rangle &= \cos\theta_B = \langle \boldsymbol{ab}' \rangle \\ \langle \boldsymbol{a}'\boldsymbol{b} \rangle &= 2\alpha\beta \sin\theta_B = -\langle \boldsymbol{a}'\boldsymbol{b}' \rangle \end{aligned} \tag{4.64}$$

Plugging into the CHSH inequality, we obtain

$$|\cos\theta_B - 2\alpha\beta \sin\theta_B| \le 1, \tag{4.65}$$

and we easily see that violations occur for θ_B close to 0 or π. Expanding to linear order in θ_B, the left hand side is

$$\simeq 1 - 2\alpha\beta\theta_B, \tag{4.66}$$

which surely exceeds 1 for θ_B negative and small.

We have shown, then, that *any* entangled pure state of two qubits violates some Bell inequality. It is not hard to generalize the argument to an arbitrary bipartite pure state. For bipartite pure states, then, "entangled" is equivalent to "Bell-inequality violating." For bipartite mixed states, however, we will see shortly that the situation is more subtle.

4.2 Uses of Entanglement

After Bell's work, quantum entanglement became a subject of intensive study among those interested in the foundations of quantum theory. But more recently (starting less than ten years ago), entanglement has come to be viewed not just as a tool for exposing the weirdness of quantum mechanics, but as a potentially valuable *resource*. By exploiting entangled quantum states, we can perform tasks that are otherwise difficult or impossible.

4.2.1 Dense coding

Our first example is an application of entanglement to communication. Alice wants to send messages to Bob. She might send classical bits (like dots and dashes in Morse code), but let's suppose that Alice and Bob are linked by a *quantum* channel. For example, Alice can prepare qubits (like photons) in any polarization state she pleases, and send them to Bob, who measures the polarization along the axis of his choice. Is there any advantage to sending qubits instead of classical bits?

In principle, if their quantum channel has perfect fidelity, and Alice and Bob perform the preparation and measurement with perfect efficiency, then they are no *worse* off using qubits instead of classical bits. Alice can prepare, say, either $|\uparrow_z\rangle$ or $|\downarrow_z\rangle$, and Bob can measure along $\hat{z}$ to infer the choice she made. This way, Alice can send one classical bit with each qubit. But in fact, that is the best she can do. Sending one qubit at a time, no matter how she prepares it and no matter how Bob measures it, no more than one classical bit can be carried by each qubit. (This statement is a special case of a bound proved by Kholevo (1973) on the classical information capacity of a quantum channel.)

But now, let's change the rules a bit – let's suppose that Alice and Bob share an entangled pair of qubits in the state $|\phi^+\rangle_{AB}$. The pair was prepared last year; one qubit was shipped to Alice and the other to Bob, anticipating that the shared entanglement would come in handy someday. Now, use of the quantum channel is very expensive, so Alice can afford to send only one qubit to Bob. Yet it is of the utmost importance for Alice to send Bob *two* classical bits of information.

Fortunately, Alice remembers about the entangled state $|\phi^+\rangle_{AB}$ that she shares with Bob, and she carries out a protocol that she and Bob had arranged for just such an emergency. On her member of the entangled pair, she can perform one of four possible unitary transformations:

1) $\mathbf{1}$ (she does nothing),

2) $\boldsymbol{\sigma}_1$ (180^o rotation about $\hat{x}$-axis),

3) $\boldsymbol{\sigma}_2$ (180^o rotation about $\hat{y}$-axis),

4) $\boldsymbol{\sigma}_3$ (180^o rotation about $\hat{z}$-axis).

As we have seen, by doing so, she transforms $|\phi^+\rangle_{AB}$ to one of 4 mutually orthogonal states:

1) $|\phi^+\rangle_{AB}$,

2) $|\psi^+\rangle_{AB}$,

3) $|\psi^-\rangle_{AB}$,

4) $|\phi^-\rangle_{AB}$.

Now, she sends her qubit to Bob, who receives it and then performs an orthogonal collective measurement on the pair that projects onto the maximally entangled basis. The measurement outcome unambiguously distinguishes the four possible actions that Alice could have performed. Therefore the single qubit sent from Alice to Bob has successfully carried 2 bits of classical information! Hence this procedure is called "dense coding."

A nice feature of this protocol is that, if the message is highly confidential, Alice need not worry that an eavesdropper will intercept the transmitted qubit and decipher her message. The transmitted qubit has density matrix $\boldsymbol{\rho}_A = \frac{1}{2}\mathbf{1}_A$, and so carries no information at all. All the information is in the correlations between qubits A and B, and this information is inaccessible unless the adversary is able to obtain both members of the entangled pair. (Of course, the adversary *can* "jam" the channel, preventing the information but reaching Bob.)

From one point of view, Alice and Bob really *did* need to use the channel twice to exchange two bits of information – a qubit had to be transmitted for them to establish their entangled pair in the first place. (In effect, Alice has merely sent to Bob two qubits chosen to be in one of the four mutually orthogonal entangled states.) But the first transmission could have taken place a long time ago. The point is that when an emergency arose and two bits had to be sent immediately while only one use of the channel was possible, Alice and Bob could exploit the pre-existing entanglement to communicate more efficiently. They used entanglement as a resource.

4.2.2 EPR Quantum Key Distribution

Everyone has secrets, including Alice and Bob. Alice needs to send a highly private message to Bob, but Alice and Bob have a very nosy friend, Eve, who they know will try to listen in. Can they communicate with assurance that Eve is unable to eavesdrop?

Obviously, they should use some kind of code. Trouble is, aside from being very nosy, Eve is also very smart. Alice and Bob are not confident that they are clever enough to devise a code that Eve cannot break.

Except there is one coding scheme that is surely unbreakable. If Alice and Bob share a *private key*, a string of random bits known only to them, then Alice can convert her message to ASCII (a string of bits no longer than the key) *add* each bit of her message (module 2) to the corresponding bit of the key, and send the result to Bob. Receiving this string, Bob adds the key to it to extract Alice's message.

This scheme is secure because even if Eve should intercept the transmission, she will not learn anything because the transmitted string itself carries no information – the message is encoded in a correlation between the transmitted string and the *key* (which Eve doesn't know).

There is still a problem, though, because Alice and Bob need to establish a shared random key, and they must ensure that Eve can't know the key. They could meet to exchange the key, but that might be impractical. They could entrust a third party to transport the key, but what if the intermediary is secretly in cahoots with Eve? They could use "public key" distribution protocols, but these are not guaranteed to be secure.

Can Alice and Bob exploit *quantum* information (and specifically entanglement) to solve the key exchange problem? They can! This observation is the basis of what is sometimes called "quantum cryptography." But since quantum mechanics is really used for key exchange rather than for encoding, it is more properly called "quantum key distribution."

Let's suppose that Alice and Bob share a supply of entangled pairs, each prepared in the state $|\psi^-\rangle$. To establish a shared private key, they may carry out this protocol.

For each qubit in her/his possession, Alice and Bob decide to measure either $\boldsymbol{\sigma}_1$ or $\boldsymbol{\sigma}_3$. The decision is pseudo-random, each choice occuring with probability 1/2. Then, after the measurements are performed, both Alice and Bob publicly announce what observables they measured, but do not reveal the outcomes they obtained. For those cases (about half) in which they measured their qubits along different axes, their results are discarded (as Alice and Bob obtained uncorrelated outcomes). For those cases in which they measured along the same axis, their results, though random, are *perfectly (anti-)correlated.* Hence, they have established a shared random key.

But, is this protocol really invulnerable to a sneaky attack by Eve? In particular, Eve might have clandestinely tampered with the pairs at some time and in the past. Then the pairs that Alice and Bob possess might be (unbeknownst to Alice and Bob) not perfect $|\psi^-\rangle$'s, but rather pairs that are entangled with qubits in Eve's possession. Eve can then wait until Alice and Bob make their public announcements, and proceed to measure her qubits in a manner designed to acquire maximal information about the results that Alice and Bob obtained. Alice and Bob must protect themselves against this type of attack.

If Eve has indeed tampered with Alice's and Bob's pairs, then the most general possible state for an AB pair and a set of E qubits has the form

$$\begin{aligned}|\Upsilon\rangle_{ABE} = &|00\rangle_{AB}|e_{00}\rangle_E + |01\rangle_{AB}|e_{01}\rangle_E \\ &+ |10\rangle_{AB}|e_{10}\rangle_E + |11\rangle_{AB}|e_{11}\rangle_E. \end{aligned} \tag{4.67}$$

But now recall that the defining property or $|\psi^-\rangle$ is that it is an eigenstate with eigenvalue -1 of both $\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)}$ and $\boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)}$. Suppose that A and B are able to verify that the pairs in their possession have this property. To satisfy $\boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)} = -1$, we must have

$$|\Upsilon\rangle_{AB} = |01\rangle_{AB}|e_{01}\rangle_E + |10\rangle_{AB}|e_{10}\rangle_E, \tag{4.68}$$

and to also satisfy $\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)} = -1$, we must have

$$|\Upsilon\rangle_{ABE} = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)|e\rangle_E = |\psi^-\rangle|e\rangle. \tag{4.69}$$

We see that it is possible for the AB pairs to be eigenstates of $\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)}$ and $\boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)}$ only if they are completely unentangled with Eve's qubits.

Therefore, Eve will not be able to learn anything about Alice's and Bob's measurement results by measuring her qubits. The random key is secure.

To verify the properties $\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)} = -1 = \boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)}$, Alice and Bob can sacrifice a portion of their shared key, and publicly compare their measurement outcomes. They should find that their results are indeed perfectly correlated. If so they will have high statistical confidence that Eve is unable to intercept the key. If not, they have detected Eve's nefarious activity. They may then discard the key, and make a fresh attempt to establish a secure key.

As I have just presented it, the quantum key distribution protocol seems to require entangled pairs shared by Alice and Bob, but this is not really so. We might imagine that Alice prepares the $|\psi^-\rangle$ pairs herself, and then measures one qubit in each pair before sending the other to Bob. This is completely equivalent to a scheme in which Alice prepares one of the four states

$$|\uparrow_z\rangle, |\downarrow_z\rangle, |\uparrow_x\rangle, |\downarrow_x\rangle, \tag{4.70}$$

(chosen at random, each occuring with probability 1/4) and sends the qubit to Bob. Bob's measurement and the verification are then carried out as before. This scheme (known as BB84 in quantum key distribution jargon) is just as secure as the entanglement-based scheme.[3]

Another intriguing variation is called the "time-reversed EPR" scheme. Here both Alice and Bob prepare one of the four states in eq. (4.70), and they both send their qubits to Charlie. Then Charlie performs a Bell measurement on the pair, orthogonally projecting out one of $|\phi^\pm\rangle|\psi^\pm\rangle$, and he publicly announces the result. Since all four of these states are simultaneous eigenstates of $\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)}$ and $\boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)}$, when Alice and Bob both prepared their spins along the same axis (as they do about half the time) they share a single bit.[4] Of course, Charlie could be allied with Eve, but Alice and Bob can verify that Charlie has acquired no information as before, by comparing a portion of their key. This scheme has the advantage that Charlie could

[3]Except that in the EPR scheme, Alice and Bob can wait until just before they need to talk to generate the key, thus reducing the risk that Eve might at some point burglarize Alice's safe to learn what states Alice prepared (and so infer the key).

[4]Until Charlie does his measurement, the states prepared by Bob and Alice are totally uncorrelated. A definite correlation (or anti-correlation) is established after Charlie performs his measurement.

operate a central switching station by storing qubits received from many parties, and then perform his Bell measurement when two of the parties request a secure communication link. A secure key can be established even if the quantum communication line is down temporarily, as long as both parties had the foresight to send their qubits to Charlie on an earlier occasion (when the quantum channel was open.)

So far, we have made the unrealistic assumption that the quantum communication channel is perfect, but of course in the real world errors will occur. Therefore even if Eve has been up to no mischief, Alice and Bob will sometimes find that their verification test will fail. But how are they to distinguish errors due to imperfections of the channel from errors that occur because Eve has been eavesdropping?

To address this problem, Alice and Bob must enhance their protocol in two ways. First they must implement (classical) error correction to reduce the effective error rate. For example, to establish each bit of their shared key they could actually exchange a block of three random bits. If the three bits are not all the same, Alice can inform Bob which of the three is different than the other two; Bob can flip that bit in his block, and *then* use majority voting to determine a bit value for the block. This way, Alice and Bob share the same key bit even if an error occured for one bit in the block of three.

However, error correction alone does not suffice to ensure that Eve has acquired negligible information about the key – error correction must be supplemented by (classical) privacy amplification. For example, after performing error correction so that they are confident that they share the same key, Alice and Bob might extract a bit of "superkey" as the *parity* of n key bits. To know *anything* about the parity of n bits, Eve would need to know *something* about each of the bits. Therefore, the parity bit is considerably more secure, on the average, than each of the individual key bits.

If the error rate of the channel is low enough, one can hope to show that quantum key distribution, supplemented by error correction and privacy amplification, is invulnerable to any attack that Eve might muster (in the sense that the information acquired by Eve can be guaranteed to be arbitrarily small). Whether this has been established is, at the moment, a matter of controversy.

4.2.3 No cloning

The security of quantum key distribution is based on an essential difference between quantum information and classical information. It is not possible to acquire information that *distinguishes* between nonorthogonal quantum states without *disturbing* the states.

For example, in the BB84 protocol, Alice sends to Bob any one of the four states $|\uparrow_z\rangle|\downarrow_z\rangle|\uparrow_x\rangle|\downarrow_x\rangle$, and Alice and Bob are able to verify that none of their states are perturbed by Eve's attempt at eavesdropping. Suppose, more generally, that $|\varphi\rangle$ and $|\psi\rangle$ are two nonorthogonal states in $\mathcal{H}$ ($\langle\psi|\varphi\rangle \neq 0$) and that a unitary transformation U is applied to $\mathcal{H}\otimes\mathcal{H}_E$ (where $\mathcal{H}_E$ is a Hilbert space accessible to Eve) that leaves both $|\psi\rangle$ and $|\varphi\rangle$ undisturbed. Then

$$\begin{aligned} U: \quad & |\psi\rangle \otimes |0\rangle_E \rightarrow |\psi\rangle \otimes |e\rangle_E, \\ & |\varphi\rangle \otimes |0\rangle_E \rightarrow |\varphi\rangle \otimes |f\rangle_E, \end{aligned} \tag{4.71}$$

and unitarity implies that

$$\begin{aligned} \langle\psi|\phi\rangle &= ({}_E\langle 0| \otimes \langle\psi|)(|\varphi\rangle \otimes |0\rangle_E) \\ &= ({}_E\langle e| \otimes \langle\psi|)(|\varphi\rangle \otimes |f\rangle_E) \\ &= \langle\psi|\varphi\rangle_E\langle e|f\rangle_E. \end{aligned} \tag{4.72}$$

Hence, for $\langle\psi|\varphi\rangle \neq 0$, we have ${}_E\langle e|f\rangle_E = 1$, and therefore since the states are normalized, $|e\rangle = |f\rangle$. This means that no measurement in $\mathcal{H}_E$ can reveal any information that distinguishes $|\psi\rangle$ from $|\varphi\rangle$. In the BB84 case this argument shows that the state in $\mathcal{H}_E$ will be the same irrespective of which of the four states $|\uparrow_z\rangle, |\downarrow_z\rangle, |\uparrow_x\rangle, |\downarrow_x\rangle$ is sent by Alice, and therefore Eve learns nothing about the key shared by Alice and Bob. On the other hand, if Alice is sending to Bob one of the two orthogonal states $|\uparrow_z\rangle$ or $|\downarrow_z\rangle$, there is nothing to prevent Eve from acquiring a copy of the information (as with classical bits).

We have noted earlier that if we have many identical copies of a qubit, then it is possible to measure the mean value of noncommuting observables like $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2$, and $\boldsymbol{\sigma}_3$ to completely determine the density matrix of the qubit. Inherent in the conclusion that nonorthogonal state cannot be distinguished without disturbing them, then, is the implicit provision that it is not possible to make a perfect copy of a qubit. (If we could, we would make as many copies as we need to find $\langle\boldsymbol{\sigma}_1\rangle, \langle\boldsymbol{\sigma}_2\rangle$, and $\langle\boldsymbol{\sigma}_3\rangle$ to any specified accuracy.) Let's now make this point explicit: there is no such thing as a perfect quantum Xerox machine.

Orthogonal quantum states (like classical information) *can* be reliably copied. For example, the unitary transformation that acts as

$$\begin{aligned} U: \quad & |0\rangle_A|0\rangle_B \to |0\rangle_A|0\rangle_B \\ & |1\rangle_A|0\rangle_B \to |1\rangle_A|1\rangle_B, \end{aligned} \tag{4.73}$$

copies the first qubit onto the second if the first qubit is in one of the states $|0\rangle_A$ or $|1\rangle_A$. But if instead the first qubit is in the state $|\psi\rangle = a|0\rangle_A + b|1\rangle_A$, then

$$\begin{aligned} U: \quad & (a|0\rangle_A + b|1\rangle_A)|0\rangle_B \\ & \to a|0\rangle_A|0\rangle_B + b|1\rangle_A|1\rangle_B. \end{aligned} \tag{4.74}$$

Thus is *not* the state $|\psi\rangle \otimes |\psi\rangle$ (a tensor product of the original and the copy); rather it is something very different – an entangled state of the two qubits.

To consider the most general possible quantum Xerox machine, we allow the full Hilbert space to be larger than the tensor product of the space of the original and the space of the copy. Then the most general "copying" unitary transformation acts as

$$\begin{aligned} U: \quad & |\psi\rangle_A|0\rangle_B|0\rangle_E \to |\psi\rangle_A|\psi\rangle_B|e\rangle_E \\ & |\varphi\rangle_A|0\rangle_B|0\rangle_E \to |\varphi\rangle_A|\varphi\rangle_B|f\rangle_E. \end{aligned} \tag{4.75}$$

Unitarity then implies that

$${}_A\langle\psi|\varphi\rangle_A = {}_A\langle\psi|\varphi\rangle_A \; {}_B\langle\psi|\varphi\rangle_B \; {}_E\langle e|f\rangle_E; \tag{4.76}$$

therefore, if $\langle\psi|\varphi\rangle \neq 0$, then

$$1 = \langle\psi|\varphi\rangle \; {}_E\langle e|f\rangle_E. \tag{4.77}$$

Since the states are normalized, we conclude that

$$|\langle\psi|\varphi\rangle| = 1, \tag{4.78}$$

so that $|\psi\rangle$ and $|\varphi\rangle$ actually represent the same ray. No unitary machine can make a copy of both $|\varphi\rangle$ and $|\psi\rangle$ if $|\varphi\rangle$ and $|\psi\rangle$ are *distinct*, *nonorthogonal* states. This result is called the no-cloning theorem.

4.2.4 Quantum teleportation

In dense coding, we saw a case where quantum information could be exploited to enhance the transmission of classical information. Now let's address a closely related issue: Can we use classical information to realize transmission of quantum information?

Alice has a qubit, but she doesn't know it's state. Bob needs this qubit desperately. But that darn quantum channel is down again! Alice can send only *classical* information to Bob.

She could try measuring $\vec{\boldsymbol{\sigma}} \cdot \hat{n}$, projecting her qubit to either $|\uparrow_{\hat{n}}\rangle$ or $|\downarrow_{\hat{n}}\rangle$. She could send the measurement outcome to Bob who could then proceed to prepare the state Alice found. But you showed in a homework exercise that Bob's qubit will not be a perfect copy of Alice's; on the average we'll have

$$F = |{}_B\langle \cdot \, |\psi\rangle_A|^2 = \frac{2}{3}, \tag{4.79}$$

Thus is a better fidelity than could have been achieved ($F = \frac{1}{2}$) if Bob had merely chosen a state at random, but it is not nearly as good as the fidelity that Bob requires.

But then Alice and Bob recall that they share some entangled pairs; why not use the entanglement as a *resource*? They carry out this protocol: Alice unites the unknown qubit $|\psi\rangle_C$ she wants to send to Bob with her member of a $|\phi^+\rangle_{AB}$ pair that she shares with Bob. On these two qubits she performs Bell measurement, projecting onto one of the four states $|\phi^\pm\rangle_{CA}, |\psi^\pm\rangle_{CA}$. She sends her measurement outcome (two bits of classical information) to Bob over the classical channel. Receiving this information, Bob performs one of four operations on his qubit $|\cdot\rangle_B$:

$$\begin{aligned} |\phi^+\rangle_{CA} &\rightarrow \mathbf{1}_B \\ |\psi^+\rangle_{CA} &\rightarrow \boldsymbol{\sigma}_1^{(B)} \\ |\psi^-\rangle_{CA} &\rightarrow \boldsymbol{\sigma}_2^{(B)} \\ |\phi^-\rangle_{CA} &\rightarrow \boldsymbol{\sigma}_3^{(B)}. \end{aligned} \tag{4.80}$$

This action transforms his qubit (his member of the $|\phi^+\rangle_{AB}$ pair that he initially shared with Alice) into a perfect copy of $|\psi\rangle_C$! This magic trick is called *quantum teleportation.*

It is a curious procedure. Initially, Bob's qubit $|\cdot\rangle_B$ is completely unentangled with the unknown qubit $|\psi\rangle_C$, but Alice's Bell measurement establishes a correlation between A and C. The measurement outcome is in fact completely random, as you'll see in a moment, so Alice (and Bob) actually acquire no information at all about $|\psi\rangle$ by making this measurement.

How then does the quantum state manage to travel from Alice to Bob? It is a bit puzzling. On the one hand, we can hardly say that the two classical bits that were transmitted carried this information – the bits were random. So we are tempted to say that the shared entangled pair made the teleportation possible. But remember that the entangled pair was actually prepared last year, long before Alice ever dreamed that she would be sending the qubit to Bob ...

We should also note that the teleportation procedure is fully consistent with the no-cloning theorem. True, a copy of the state $|\psi\rangle$ appeared in Bob's hands. But the original $|\psi\rangle_C$ had to be destroyed by Alice's measurement before the copy could be created.

How does it work? We merely note that for $|\psi\rangle = a|0\rangle + b|1\rangle$, we may write

$$
\begin{aligned}
|\psi\rangle_C|\phi^+\rangle_{AB} &= (a|0\rangle_C + b|1\rangle_C)\frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB}) \\
&= \frac{1}{\sqrt{2}}(a|000\rangle_{CAB} + a|011\rangle_{CAB} + b|100\rangle_{CAB} + b|111\rangle_{CAB}) \\
&= \frac{1}{2}a(|\phi^+\rangle_{CA} + |\phi^-\rangle_{CA})|0\rangle_B + \frac{1}{2}a(|\psi^+\rangle_{CA} + |\psi^-\rangle_{CA})|1\rangle_B \\
&+ \frac{1}{2}b(|\psi^+\rangle_{CA} - |\psi^-\rangle_{CA})|0\rangle_B + \frac{1}{2}b(|\phi^+\rangle_{CA} - |\phi^-\rangle_{CA})|1\rangle_B \\
&= \frac{1}{2}|\phi^+\rangle_{CA}(a|0\rangle_B + b|1\rangle_B) \\
&+ \frac{1}{2}|\psi^+\rangle_{CA}(a|1\rangle_B + b|0\rangle_B) \\
&+ \frac{1}{2}|\psi^-\rangle_{CA}(a|1\rangle_B - b|0\rangle_B) \\
&+ \frac{1}{2}|\phi^-\rangle_{CA}(a|0\rangle_B - b|1\rangle_B) \\
&= \frac{1}{2}|\phi^+\rangle_{CA}|\psi\rangle_B + \frac{1}{2}|\psi^+\rangle_{CA}\boldsymbol{\sigma}_1|\psi\rangle_B \\
&+ \frac{1}{2}|\psi^-\rangle_{CA}(-i\boldsymbol{\sigma}_2)|\psi\rangle_B + \frac{1}{2}|\phi^-\rangle_{CA}\boldsymbol{\sigma}_3|\psi\rangle_B. \qquad (4.81)
\end{aligned}
$$

Thus we see that when we perform the Bell measurement on qubits C and

A, all four outcomes are equally likely, and that the actions prescribed in Eq. (4.80) will restore Bob's qubit to the initial state $|\psi\rangle$.