At this point we have (more or less) all the components of quantum information theory that we’ll need. To recap:

- **Pure states**: unit vectors $|\psi\rangle \in \mathbb{C}^d$
- **Mixed states**: Trace 1 operators $\rho \in \mathcal{M}(\mathbb{C}^d)$
- **Gates**: unitary matrices $U : \mathbb{C}^d \rightarrow \mathbb{C}^d$
- **Measurement**: $\sum \rho_i \leq 1$, $\sum \rho_i = I$

\[ p_i p_j = 0 \quad \forall i \neq j \]
\[ p_i^2 = p_i \quad \forall i \]

Before we move on to general quantum computation, let’s work through a few quantum communication protocols as a warm up. Today we’ll see how to send 2 bits of classical data by sending a single (entangled) qubit.
We've previously used the entangled state

\[ |14\rangle = \frac{1}{\sqrt{2}} (|100\rangle + |11\rangle) \]

in our protocols, but this is not the only 2-qubit entangled state. In fact there exists an orthonormal basis of \( \mathbb{C}^2 \otimes \mathbb{C}^2 \) which consists of entangled states. This is called the Bell basis, consisting of the Bell states

\[
|B_{00}\rangle = \frac{1}{\sqrt{2}} (|100\rangle + |11\rangle), \\
|B_{01}\rangle = \frac{1}{\sqrt{2}} (|101\rangle + |110\rangle), \\
|B_{10}\rangle = \frac{1}{\sqrt{2}} (|100\rangle - |11\rangle), \\
|B_{11}\rangle = \frac{1}{\sqrt{2}} (|101\rangle - |110\rangle).
\]

It can be readily verified that these 4 states are orthonormal and non-separable (i.e. entangled). As an entangled basis of \( \mathbb{C}^2 \otimes \mathbb{C}^2 \), these states arise frequently in quantum computation.

**Ex.**
Recall that the circuit

\[
\begin{array}{c}
\text{H} \\
\hline
\otimes
\end{array}
\]

(\( \text{CNOT} (H \otimes I) \))

generates the EPR pair or Bell state \( |B_{00}\rangle \) from the \( |100\rangle \) initial state. We can ask what states it maps the other 3 computational basis states to:

\[
\begin{align*}
\text{CNOT} (H \otimes I) |101\rangle &= \text{CNOT} (\frac{1}{\sqrt{2}} |101\rangle + \frac{1}{\sqrt{2}} |110\rangle) \\
&= \frac{1}{\sqrt{2}} |101\rangle + \frac{1}{\sqrt{2}} |110\rangle \\
&= |B_{01}\rangle \\
\text{CNOT} (H \otimes I) |110\rangle &= \text{CNOT} (\frac{1}{\sqrt{2}} |100\rangle - \frac{1}{\sqrt{2}} |110\rangle) \\
&= |B_{10}\rangle.
\end{align*}
\]
\[
\text{CNOT}(H \otimes I) 111 = \text{CNOT} \left( \frac{1}{\sqrt{2}} 101 - \frac{1}{\sqrt{2}} 111 \right) \\
= \frac{1}{\sqrt{2}} 110
\]

So this circuit implements a change of basis from the computational basis to the Bell basis. Since the circuit is unitary, it follows that its dagger

\[
(\text{CNOT}(H \otimes I))^* = (H \otimes I) \text{CNOT}
\]

maps the Bell basis back to the computational basis:

\[
(H \otimes I) \text{CNOT} \ket{100} = (H \otimes I) \left( \frac{1}{\sqrt{2}} 100 + \frac{1}{\sqrt{2}} 110 \right) \\
= \frac{1}{\sqrt{2}} (100 + 110 + 100 - 110) \\
= 100
\]

(Bell basis measurement)

We denote measurement in the Bell basis by

We observe that Bell basis measurement is equivalent to a change of basis followed by a computational basis measurement and a change back to the Bell basis.

We’ve used this trick before to measure in the \(\{+,-\}\) basis. So far it’s just a mathematical convenience, but soon we’ll see that it’s important for building a real quantum computer. In general, for any orthonormal basis \(\{e_i\}\), the matrix \(U = [e_1 \ldots e_d]\) is unitary, and

\[
\begin{array}{c}
\text{Bell} \\
\hline
\end{array} \\
= \\
\begin{array}{c}
U^T \\
\hline
\end{array} \\
\begin{array}{c}
\chi \\
\hline
\end{array} \\
\begin{array}{c}
\chi \\
\hline
\end{array} \\
\begin{array}{c}
U \\
\hline
\end{array}
\]
The ability to measure in the Bell basis leads to some surprising consequences. The first one we will discuss is superdense coding.

(Superdense coding)

Suppose Alice wants to send Bob two bits of information over a telephone line. Basic information theory states that Alice actually needs to send both bits (is this obvious?)

What if Alice can also send qubits over the telephone line? If Alice and Bob share an entangled pair, it turns out that Alice only needs to send a single qubit to send two (classical) bits of information. Let's see how this works.

Recall the definitions of the $X$ & $Z$ gates:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Alice and Bob share a pre-entangled $|\psi_{oo}\rangle$ state. To send the bits $a$ and $b$, Alice applies the following transformation to her qubit

<table>
<thead>
<tr>
<th>$a$</th>
<th>$b$</th>
<th>transformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>$I$</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>$X$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>$Z$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$ZX$</td>
</tr>
</tbody>
</table>

Alternatively, we can write the transformation as $Z^a X^b$. Alice then sends her qubit to Bob who measures both qubits in the Bell basis to get $|\psi_{ab}\rangle$. 
The protocol can be summarized as below:

To see why this works, observe that

\[
(\mathbf{X} \otimes \mathbf{I}) |B_{00}\rangle = \frac{1}{\sqrt{2}} \left( (\mathbf{X} \otimes \mathbf{I}) |00\rangle + (\mathbf{X} \otimes \mathbf{I}) |11\rangle \right)
= \frac{1}{\sqrt{2}} \left( |10\rangle + |01\rangle \right)
= |B_{01}\rangle
\]

Likewise,

\[
(\mathbf{Z} \otimes \mathbf{I}) |B_{00}\rangle = \frac{1}{\sqrt{2}} \left( (\mathbf{Z} \otimes \mathbf{I}) |00\rangle + (\mathbf{Z} \otimes \mathbf{I}) |11\rangle \right)
= \frac{1}{\sqrt{2}} \left( |10\rangle - |11\rangle \right)
= |B_{10}\rangle
\]

\[
(2 \otimes \mathbf{I}) |B_{00}\rangle = \frac{1}{\sqrt{2}} \left( (2 \otimes \mathbf{I}) |00\rangle + (2 \otimes \mathbf{I}) |11\rangle \right)
= \frac{1}{\sqrt{2}} \left( -|10\rangle + |01\rangle \right)
= |B_{11}\rangle
\]

In short,

\[
(\mathbf{Z}^a \otimes \mathbf{X}^b \otimes \mathbf{I}) |B_{00}\rangle = |B_{ab}\rangle
\]

So when Bob measures in the Bell basis, his result is $B_{ab}$ which tells him Alice’s bits $a$ & $b$. 
At first glance superdense coding seems to contradict the fact that local operations Alice performs can’t alter Bob’s reduced density matrix. The key here is that while local operations don’t impact Bob’s reduced density matrix, which accounts for all local measurements of Bob’s qubit, they can change the joint (entangled) state, which is observable through a joint measurement.

You may wonder whether we really need the two entangled qubits — that is, could Alice just send a state encoding her bits like

\[ (-1)^a |1b\rangle \]

and then have Bob “decode” the state? The answer is a resounding NO by Holevo’s Theorem which states, roughly, that you can’t get more than one bit of information out of a qubit.