CS 294-2 Introduction, Axioms, Bell Inequalities 1/17/07
Spring 2007 Lecture 1

1 Why Quantum Computation?

- Quantum computers are the only model of computation that escape the limitations on computation imposed by the extended Church-Turing thesis. These theoretical limitations will soon have practical consequences as Moore's Law (the doubling every eighteen months in transistor density on chips and the resulting increase in computer speed) is expected in a decade or so to run into inherent (classical) physical limits such as transistors scaling down to the size of elementary particles.
- Quantum computers would have a profound effect on complexity based cryptography, via the polynomial time quantum algorithms for factoring and discrete logs. Understanding how to meet these threats by designing quantum resistant cryptosystems or by resorting to quantum cryptography is a major challenge.
- Quantum computation has brought about a renaissance in the examination of the foundations of quantum mechanics: highlighting phenomena such as entanglement and the resulting exponential power inherent in quantum physics, and showing that quantum error-correcting codes may be used to bring stability into the seemingly inherently unstable quantum world.
- Lastly, the study of quantum computation has helped broaden and deepen our insights into complexity theory (and tools such as Fourier analysis and information theory), and in a few cases helped resolve open questions in classical complexity theory.

2 Basic Quantum Mechanics

The basic formalism of quantum mechanics is very simple, though understanding and interpreting the results is much more challenging. There are three basic principles, enshrined in the three basic postulates of quantum mechanics:

- The superpostion principle: this axiom tells us what the state of a quantum system looks like.

  * An addendum to this axiom tells us given two subsystems, what the allowable states of the composte system are.
- The measurement principle: this axiom governs how much information about the state we can access.
- Unitary evolution: this axiom governs how the state of the quantum system evolves in time.

2.1 The superposition principle

Consider a system with k distinguishable states. For example, the electron in an atom might be either in its ground state or one of $k-1$ excited states, each of progressively higher energy. As a classical system, we might use the state of this system to store a number between 0 and $k-1$. The superposition principle says that if a quantum system is allowed to be any one of number of different states then it can also be placed in a linear superposition of these states with complex coefficients. Thus the quantum state of the k-state system above is described by a sequence of k complex numbers $\alpha_0,\dots,\alpha_{k-1} \in \mathscr{C}$. α_j is said to be the (complex) amplitude with which the system is in state j. We will require that the amplitudes are normalized so that $\sum_j |\alpha_j|^2 = 1$. It is natural to write the state of the system as a k dimensional vector:

$$\begin{pmatrix} \alpha_0 \\ \alpha_1 \\ . \\ . \\ \alpha_{k-1} \end{pmatrix}$$

The normalization on the complex amplitudes means that the state of the system is a unit vector in a k dimensional complex vector space — called a Hilbert space.

In quantum mechanics it is customary to use the Dirac's ket notation to write vectors. As we shall see later, this is a particularly useful notation in the context of quantum computation. In the ket notation, the above state is written as:

$$\left|\psi\right\rangle = \sum_{j=0}^{k-1} \alpha_j \left|j\right\rangle$$

Here $\left|0\right\rangle = \begin{pmatrix} 1 \\ 0 \\ . \\ . \\ 0 \end{pmatrix}$ and $\left|k-1\right\rangle = \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ 1 \end{pmatrix}$. The Dirac notation has the advantage that the it labels the basis vectors explicitly. This is very convenient because the notation expresses both that the state of the quantum system is a vector, while at the same time it represents a number (in superposition) describing the physical quantity of interest (energy level, spin, polarization, etc). In the context of quantum computation and quantum information it is data (0 or 1) to be processed. The $\{\left|0\right\rangle, \left|1\right\rangle, \dots, \left|k-1\right\rangle\}$ basis is called the standard basis.

2.2 Measurement Principle

This linear superposition $\left|\psi\right\rangle = \sum_{j=0}^{k-1} \alpha_j \left|j\right\rangle$ is part of the private world of the electron. For us to know the electron's state, we must make a measurement. Measuring $\left|\psi\right\rangle$ in the standard basis yields j with probability $|\alpha_j|^2$.

One important aspect of the measurement process is that it alters the state of the quantum system: the effect of the measurement is that the new state is exactly the outcome of the measurement. I.e., if the outcome of the measurement is j, then following the measurement, the qubit is in state $\left|j\right\rangle$. This implies that you cannot collect any additional information about the amplitudes α_j by repeating the measurement. This property forms the basis of quantum cryptography where the presence of an eavesdropper necessarily alters the quantum state being transmitted.

Intuitively, a measurement provides the only way of reaching into the Hilbert space to probe the quantum state vector. In general this is done by selecting an orthonormal basis $\left|e_0\right\rangle,\dots,\left|e_{k-1}\right\rangle$. The outcome of the measurement is $\left|e_j\right\rangle$ (or j) with probability equal to the square of the length of the projection of the state vector ψ on $\left|e_j\right\rangle$. A consequence of performing the measurement is that the new state vector is $\left|e_j\right\rangle$. Thus measurement may be regarded as a probabilistic rule for projecting the state vector onto one of the vectors of the orthonormal measurement basis.

2.3 Qubits

The basic entity of quantum information is a qubit (pronounced "cue-bit"), or a quantum bit. This corresponds to a 2-state quantum system, and its state can be written as a unit (column) vector $\binom{\alpha}{\beta} \in \mathscr{C}^2$. In Dirac notation, this may be written as:

$$\left|\psi\right\rangle = \alpha\left|0\right\rangle + \beta\left|1\right\rangle \quad \alpha,\beta \in \mathscr{C} \quad \text{and} \quad |\alpha|^2 + |\beta|^2 = 1$$

This linear superposition $\left|\psi\right\rangle = \alpha\left|0\right\rangle + \beta\left|1\right\rangle$ is part of the private world of the electron. For us to know the electron's state, we must make a measurement. Measuring $\left|\psi\right\rangle$ in the $\{\left|0\right\rangle, \left|1\right\rangle\}$ basis yields 0 with probability $|\alpha|^2$, and 1 with probability $|\beta|^2$.

As we noted above, the measurement process alters the state of the qubit: the outcome of the measurement is a single classical bit of information, and the effect of the measurement is that the new state is exactly the outcome of the measurement. I.e., if the outcome of the measurement of$\left|\psi\right\rangle = \alpha\left|0\right\rangle + \beta\left|1\right\rangle$ yields 0, then following the measurement, the qubit is in state $\left|0\right\rangle$. This implies that you cannot collect any additional information about α, β by repeating the measurement.

More generally, we may choose any orthogonal basis $\{\left|v\right\rangle, \left|w\right\rangle$ and measure the qubit in it. To do this, we rewrite our state in that basis: $\left|\psi\right\rangle = \alpha'\left|v\right\rangle + \beta'\left|w\right\rangle$. The outcome is v with probability $|\alpha'|^2$, and $\left|w\right\rangle$ with probability $|\beta'|^2$. If the outcome of the measurement on $\left|\psi\right\rangle$ yields $\left|v\right\rangle$, then as before, the the qubit is then in state $\left|v\right\rangle$.

2.3.1 Measurement example I.

Q: We measure $\left|\psi\right\rangle = \alpha\left|0\right\rangle + \beta\left|1\right\rangle$ in the $\left|v\right\rangle, \left|v^\perp\right\rangle$ basis, where $\left|v\right\rangle = a\left|0\right\rangle + b\left|1\right\rangle$. What is the probability of measuring $\left|v\right\rangle$?

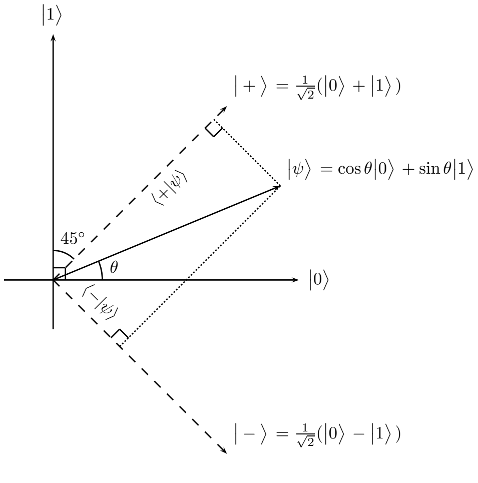

$|1\rangle$
$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$
$|\psi\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$
$\langle +|\psi\rangle$
45°
θ
$|0\rangle$
$\langle -|\psi\rangle$
$|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$

Figure 1:

A: First let's do the simpler case $a = b = \frac{1}{\sqrt{2}}$, so $\left|v\right\rangle = \frac{1}{\sqrt{2}}(\left|0\right\rangle + \left|1\right\rangle) \equiv \left|+\right\rangle$, $\left|v^\perp\right\rangle = \frac{1}{\sqrt{2}}(\left|0\right\rangle - \left|1\right\rangle) \equiv \left|-\right\rangle$.

See Figure 1. We express $\big|\psi\big\rangle$ in the $\big|+\big\rangle, \big|-\big\rangle$ basis:

$$\begin{aligned}\big|\psi\big\rangle &= \alpha\big|0\big\rangle + \beta\big|1\big\rangle \\ &= \alpha\tfrac{1}{\sqrt{2}}(\big|+\big\rangle + \big|-\big\rangle) + \beta\tfrac{1}{\sqrt{2}}(\big|+\big\rangle - \big|-\big\rangle) \\ &= \Big(\tfrac{\alpha+\beta}{\sqrt{2}}\big|+\big\rangle + \tfrac{\alpha-\beta}{\sqrt{2}}\big|-\big\rangle\Big) \quad .\end{aligned}$$

Therefore the probability of measuring $\big|+\big\rangle$ is $|\frac{1}{\sqrt{2}}(\alpha+\beta)|^2 = |\alpha+\beta|^2/2$. The probability of measuring $\big|-\big\rangle$ is $|\alpha+\beta|^2/2$. We will do the general case in §2.3.2.

2.3.2 Measurement example II.

The notation $\langle v|$ (“bra v”) denotes a row vector, the conjugate-transpose of $|v\rangle$, or $|v\rangle^\dagger$. For example, $\langle 0| = (1\ 0)$ and $\langle 1| = (0\ 1)$. More generally,

$$\langle\psi| = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}^\dagger = (\bar\alpha\ \bar\beta) = \bar\alpha\langle 0| + \bar\beta\langle 1| \quad .$$

The Dirac notation can be handy. For example, let

$$\big|v_1\big\rangle = a_1\big|0\big\rangle + b_1\big|1\big\rangle, \quad \big|v_2\big\rangle = a_2\big|0\big\rangle + b_2\big|1\big\rangle \quad .$$

Then $\big\langle v_1\big|v_2\big\rangle$ (shorthand for $\big\langle v_1\big|\,\big|v_2\big\rangle$) is a matrix product of the 1×2 matrix $\big\langle v_1\big|$ and the 2×1 matrix $\big|v_2\big\rangle$, or just a scalar:

$$\big\langle v_1\big|v_2\big\rangle = (\bar a_1\ \bar b_1)\begin{pmatrix}a_2\\ b_2\end{pmatrix} = \bar a_1 a_2 + \bar b_1 b_2 \quad .$$

$\big\langle v_1\big|v_2\big\rangle = \overline{\big\langle v_2\big|v_1\big\rangle}$ is an inner product. Note that $\big\langle 0\big|0\big\rangle = \big\langle 1\big|1\big\rangle = 1$ and $\big\langle 0\big|1\big\rangle = \overline{\big\langle 1\big|0\big\rangle} = 0$. Thus the above equation could have been expanded,

$$\begin{aligned}\big\langle v_1\big|v_2\big\rangle &= (\bar a_1\big\langle 0\big| + \bar b_1\big\langle 1\big|)(a_2\big|0\big\rangle + b_2\big|1\big\rangle) \\ &= \bar a_1 a_2\big\langle 0\big|0\big\rangle + \bar a_1 b_2\big\langle 0\big|1\big\rangle + \bar b_1 a_2\big\langle 1\big|0\big\rangle + \bar b_1 b_2\big\langle 1\big|1\big\rangle \\ &= \bar a_1 a_2\cdot 1 + \bar a_1 b_2\cdot 0 + \bar b_1 a_2\cdot 0 + \bar b_1 b_2\cdot 1 \\ &= \bar a_1 a_2 + \bar b_1 b_2 \quad .\end{aligned}$$

In this notation, $\alpha = \big\langle 0\big|\,\psi\big\rangle$, $\beta = \big\langle 1\big|\,\psi\big\rangle$. The normalization condition $|\alpha|^2 + |\beta|^2 = 1$ is

$$\begin{aligned}1 = |\alpha|^2 + |\beta|^2 &= \bar\alpha\alpha + \bar\beta\beta \\ &= \big\langle\psi\big|0\big\rangle\big\langle 0\big|\psi\big\rangle + \big\langle\psi\big|1\big\rangle\big\langle 1\big|\psi\big\rangle \\ &= \big\langle\psi\big|(\big|0\big\rangle\big\langle 0\big| + \big|1\big\rangle\big\langle 1\big|)\big|\psi\big\rangle \\ &= \big\langle\psi\big|\,\psi\big\rangle \quad .\end{aligned}$$

The last equality above follows since $\big|0\big\rangle\big\langle 0\big| = \begin{pmatrix}1&0\\0&0\end{pmatrix}$, $\big|1\big\rangle\big\langle 1\big| = \begin{pmatrix}0&0\\0&1\end{pmatrix}$, so $\big|0\big\rangle\big\langle 0\big| + \big|1\big\rangle\big\langle 1\big|$ is the 2×2 identity matrix. (This trick is important enough to have its own name, the “resolution of the identity.”)

In the next lecture, we will introduce tensor product spaces, where the advantages of this notation increase.

With the new notation, it is simple to solve the general case of the question asked in §2.3.1. Recall $|v\rangle = a|0\rangle + b|1\rangle$ and choose $|v^{\perp}\rangle = \overline{b}|0\rangle - \overline{a}|1\rangle$. Indeed, $\langle v|v^{\perp}\rangle = ab - ba = 0$.

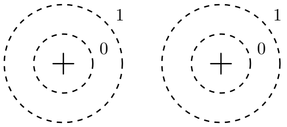
$$\begin{aligned}
|\psi\rangle &= \Big(|v\rangle\langle v| + |v^{\perp}\rangle\langle v^{\perp}|\Big)\,|\psi\rangle \\
&= \alpha(|v\rangle\langle v|0\rangle + |v^{\perp}\rangle\langle v^{\perp}|0\rangle) + \beta(|v\rangle\langle v|1\rangle + |v^{\perp}\rangle\langle v^{\perp}|1\rangle) \\
&= (\alpha\langle v|0\rangle + \beta\langle v|1\rangle)|v\rangle + (\alpha\langle v^{\perp}|0\rangle + \beta\langle v^{\perp}|1\rangle)|v^{\perp}\rangle \\
&= (\alpha\overline{a} + \beta\overline{b})|v\rangle + (\alpha b + \beta a)|v^{\perp}\rangle \ \ .
\end{aligned}$$

The probability of measuring $|v\rangle$ in a measurement in the $v, v^{\perp}$ basis is therefore

$$|\langle v|\psi\rangle|^2 = |\alpha\overline{a} + \beta\overline{b}|^2 \ \ .$$

3 Two qubits:

Now let us examine the case of two qubits. Consider the two electrons in two hydrogen atoms:



Since each electron can be in either of the ground or excited state, classically the two electrons are in one of four states – 00, 01, 10, or 11 – and represent 2 bits of classical information. Quantum mechanically, they are in a superposition of those four states:

$$\big|\psi\big\rangle = \alpha_{00}\big|00\big\rangle + \alpha_{01}\big|01\big\rangle + \alpha_{10}\big|10\big\rangle + \alpha_{11}\big|11\big\rangle \ \ ,$$

where $\sum_{ij}|\alpha_{ij}|^2 = 1$. Again, this is just Dirac notation for the unit vector in $\mathscr{C}^4$:

$$\begin{pmatrix} \alpha_{00} \\ \alpha_{01} \\ \alpha_{10} \\ \alpha_{11} \end{pmatrix}$$

where $\alpha_{ij} \in \mathscr{C}$, $\sum|\alpha_{ij}|^2 = 1$.

Measurement:

If the two electrons (qubits) are in state $\big|\psi\big\rangle$ and we measure them, then the probability that the first qubit is in state i, and the second qubit is in state j is $P(i,j) = |\alpha_{ij}|^2$. Following the measurement, the state of the two qubits is $\big|\psi'\big\rangle = \big|ij\big\rangle$. What happens if we measure just the first qubit? What is the probability that the first qubit is 0? In that case, the outcome is the same as if we had measured both qubits: $\Pr\{\text{1st bit } = 0\} = |\alpha_{00}|^2 + |\alpha_{01}|^2$. The new state of the two qubit system now consists of those terms in the superposition that are consistent with the outcome of the measurement – but normalized to be a unit vector:

$$\big|\phi\big\rangle = \frac{\alpha_{00}\big|00\big\rangle + \alpha_{01}\big|01\big\rangle}{\sqrt{|\alpha_{00}|^2 + |\alpha_{01}|^2}}$$

.

Tensor products (informal):

Suppose the first qubit is in the state $\left|\phi_1\right\rangle = \alpha_1\left|0\right\rangle + \beta_1\left|1\right\rangle$ and the second qubit is in the state $\left|\phi_2\right\rangle = \alpha_2\left|0\right\rangle + \beta_2\left|1\right\rangle$. How do we describe the joint state of the two qubits?

$$\begin{aligned}\left|\phi\right\rangle &= \left|\phi_1\right\rangle \otimes \left|\phi_2\right\rangle \\ &= \alpha_1\alpha_2\left|00\right\rangle + \alpha_1\beta_2\left|01\right\rangle + \beta_1\alpha_2\left|10\right\rangle + \beta_1\beta_2\left|11\right\rangle \quad .\end{aligned}$$

We have simply multiplied together the amplitudes of $|0\rangle_1$ and $|0\rangle_2$ to determine the amplitude of $|00\rangle_{12}$, and so on. The two qubits are not entangled with each other and measurements of the two qubits will be distrbuted independently.

Given a general state of two qubits can we say what the state of each of the individual qubits is? The answer is usually no. For a random state of two qubits is entangled — it cannot be decomposed into state of each of two qubits. In next section we will study the Bell states, which are maximally entangled states of two qubits.

Example:

State $\frac{1}{1}(\left|00\right\rangle - \left|01\right\rangle + \left|10\right\rangle - \left|11\right\rangle)$ can be decomposed into $\frac{1}{\sqrt{2}}(\left|0\right\rangle + \left|1\right\rangle)\frac{1}{\sqrt{2}}(\left|0\right\rangle - \left|1\right\rangle)$, but for state $\frac{1}{\sqrt{2}}(\left|00\right\rangle + \left|11\right\rangle)$ there is no such decomposition - the states of the two qubits are not independent but entangled/correlated.

4 EPR Paradox:

In 1935, Einstein, Podolsky and Rosen (EPR) wrote a paper “Can quantum mechanics be complete?” [Phys. Rev. 47, 777, Available online via PROLA: `http://prola.aps.org/abstract/PR/v47/i10/p777_1`]

For example, consider coin-flipping. We can model coin-flipping as a random process giving heads 50% of the time, and tails 50% of the time. This model is perfectly predictive, but incomplete. With a more accurate experimental setup, we could determine precisely the range of initial parameters for which the coin ends up heads, and the range for which it ends up tails.

For Bell state $\frac{1}{\sqrt{2}}(\left|00\right\rangle + \left|11\right\rangle)$, when you measure first qubit, the second qubit is determined. However, if two qubits are far apart, then the second qubit must have had a determined state in some time interval before measurement, since the speed of light is finite. Moreover this holds in any basis. This appears analogous to the coin flipping example. EPR therefore suggested that there is a more complete theory where “God does not throw dice.”

What would such a theory look like? Here is the most extravagant framework…When the entangled state is created, the two particles each make up a (very long!) list of all possible experiments that they might be subjected to, and decide how they will behave under each such experiment. When the two particles separate and can no longer communicate, they consult their respective lists to coordinate their actions.

But in 1964, almost three decades later, Bell showed that properties of EPR states were not merely fodder for a philosophical discussion, but had verifiable consequences: local hidden variables are not the answer.

5 Bell’s Inequality

Bell’s inequality states: There does not exist any local hidden variable theory consistent with these outcomes of quantum physics.

Consider the following communication protocol in the classical world: Alice (A) and Bob (B) are two parties who share a common string S. They receive independent, random bits X_A, X_B, and try to output bits a, b respectively, such that $X_A \wedge X_B = a \oplus b$. (The notation $x \wedge y$ takes the AND of two binary variables x and y, i.e., is one if $x = y = 1$ and zero otherwise. $x \oplus y \equiv x + y$ mod 2, the XOR.)

In the quantum mechanical analogue of this protocol, A and B share the EPR pair $\frac{1}{\sqrt{2}}(\left|00\right\rangle + \left|11\right\rangle)$. As before, they receive bits X_A, X_B, and try to output bits a, b respectively, such that $X_A \wedge X_B = a \oplus b$.

If the odd behavior of the Bell state can be explained using some hidden variable theory, then for any protocol in the quantum world there should be a corresponding classical protocol that achieves the same results.

However, Alice and Bob's best protocol for the classical (hidden variable) game, as you will prove in the homework, is to output $a = 0$ and $b = 0$, respectively. Then $a \oplus b = 0$, so as long as the inputs $(X_A, X_B) \neq (1, 1)$, they are successful: $a \oplus b = 0 = X_A \wedge X_B$. If $X_A = X_B = 1$, then they fail. Therefore they are successful with probability exactly $3/4$.

We will show that the quantum mechanical system can do better. Specifically, if Alice and Bob share an EPR pair, we will describe a protocol for which the probability $\Pr\{X_A \wedge X_B = a \oplus b\}$ is greater than 3/4, and is about $cos^2\pi/8 = 0.85$.

Here is the protocol:

- if $X_A = 0$, then Alice measures in the $-\pi/16$ basis.
- if $X_A = 1$, then Alice measures in the $3\pi/16$ basis.
- if $X_B = 0$, then Bob measures in the $\pi/16$ basis.
- if $X_B = 1$, then Bob measures in the $-3\pi/16$ basis.

Now an easy calculation shows that in each of the four cases $X_A = X_B = 0$, etc, the success probability is $cos^2\pi/8$. This is because in the three cases where $x_A \cdot x_B = 0$, Alice and Bob measure in bases that differ by $/pi/8$. In the last case they measure in bases that differ by $3\pi/8$, but in this case they must output different bits.