CS395T NOTES: QUANTUM COMPLEXITY THEORY

ARUN DEBRAY
AUGUST 29, 2016

These notes were taken in UT Austin's CS395T (Quantum Complexity Theory) class in Fall 2016, taught by Scott Aaronson. I live-TEXed them using `vim`, so there may be typos; please send questions, comments, complaints, and corrections to `a.debray@math.utexas.edu`.

Contents

Lecture 1.

Introduction to Quantum Mechanics: 8/29/16

"The big secret of quantum mechanics is how simple it is once you take the physics out of it."

The course website is `http://www.scottaaronson.com/qct2016/`, and the syllabus is at `http://www.scottaaronson.com/qct2016/syllabus-qct2016.pdf`. We'll be mostly following lecture notes found at `http://www.scottaaronson.com/barbados-2016.pdf`.

This lecture's goal is to acquaint the listener with the basic concepts and notation that we'll use in the rest of the course; it's not presented as review, but everything else in the course depends on it. For this material, there are many excellent references, some of which are listed in the syllabus.

Quantum mechanics has a very underserved reputation for being very complicated. Mysterious, yes; counterintuitive, yes; but complicated is a bit much. All sorts of interesting consequences follow from a single change to the laws of probability, crucial to physics at the subatomic level, but thought to apply to everything in the universe.

A probability of something happening is a real number $p \in [0,1]$: it makes no sense to ask what a probability of $-1/3$ is, much less $i/3$. But quantum mechanics assigns a more general number, an *amplitude* $\alpha \in \mathbb{C}$, to an event. The thesis of quantum mechanics is that any isolated physical system's state can be described by a vector of its amplitudes.

In particular, systems in quantum mechanics have a dimension; intuitively, if there are N different things you can observe, the system is N-dimensional. The simplest quantum systems are two-dimensional, where there are two possinilities) and 1. These systems have a special name: *qubits*.

In general, we think of the state of a quantum system as a unit vector $\psi \in \mathbb{C}^N$ of length 1. These vectors are denoted using a notation that Paul Dirac invented in the 1930s, the *Dirac ket notation*. The syntax looks a little jarring at first, but is convenient in a lot of ways. A *ket* is a vector $|v\rangle$: a qubit has two basis vectors $|0\rangle$, representing an outcome of 0 and $|1\rangle$, similarly an outcome of 1, so a general state is $|v\rangle = \alpha|0\rangle + \beta|1\rangle$, representing a linear combination, or *superposition*, of the two options: $\alpha, \beta \in \mathbb{C}$ are complex numbers, and must satisfy a *normalization rule*: $|\alpha|^2 + |\beta|^2 = 1$. In other words, $|v\rangle$ stands in for the column vector $\left(\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right)$. Usually, ket notation will only be for unit vectors, but sometimes we might use it more generally.

This feels a little schizophrenic. Is it both at the same time? Is it neither? In popular books, these are the only ontological categories the writer can imagine, but these really belong in a different conceptual framework altogether.

In addition to column vectors, we like row vectors too, denoted with a *bra* $\langle v|$. However, since we're in the land of complex vector spaces, taking the transpose comes along with complex conjugation, so $\langle v| = (\alpha^* \ \beta^*)$. Combining these two notations, $\langle \cdot|\cdot\rangle$ is the notation for the inner product. Thus, that v is a unit vector is succinctly expressed in the condition $\langle v|v\rangle = 1$.

Explicitly, if $|\psi\rangle = \alpha_1|1\rangle + \cdots + \alpha_N|N\rangle$ and $|\varphi\rangle = \beta_1|1\rangle + \cdots + \beta_N|N\rangle$, their inner product is $\langle\psi|\varphi\rangle = \alpha_1^*\beta_1 + \cdots + \alpha_N^*\beta_N$. This measures how similar two vectors are: if the inner product is 1, they lie on the same line, and are related, but if it's 0, they're orthogonal, and thus very different.

Relatedly, there is an *outer product* $|\psi\rangle\langle\varphi|$, which is a rank-1 $N \times N$ matrix whose ij^{th} term is $\alpha_i\beta_j^*$.

There are two things one can do to quantum systems.

(1) One option is a *unitary transformation*. These should be thought of as doing something smooth and well-behaved. They are continuous, reversible, and deterministic.
(2) The other choice is a *measurement*. These are useful, especially if you want to actually learn anything about system, but these are discontinuous and irreversible, and famously are probabilistic. Quantum mechanics tells you probabilities, not certainties.

Maybe you're wondering how two so very different systems can coexist in the same universe. This is the *measurement problem*, and people have been discussing it for a century. In some sense, unitary transformations arise from changes of basis, but if you follow that viewpoint far enough, it seems like all of quantum mechanics is a particular change of basis! Yet there are ways in which the choice of basis matters; unitary transformations are information-preserving, relating to the very general physical principle that information cannot be destroyed. A unitary transformation might horribly transform information, but it's still there.

The measurement problem and its metaphysics notwithstanding, we can at least write down the mathematical rules for these transformations. A unitary evolution is multiplication by a matrix: $|\psi\rangle \mapsto U|\psi\rangle$, but U must be norm-preserving, so that all valid quantum states map to valid quantum states.[1] Since unitary transformations should be reversible, we'd like U to be an invertible matrix.

Exercise 1.1. Show that the following are equivalent for a linear transformation $U : \mathbb{C}^n \to \mathbb{C}^n$:

(1) U is norm-preserving and invertible.
(2) U preserves inner products, i.e. $\langle U\psi|U\varphi\rangle = \langle\psi|\varphi\rangle$ for all $|\psi\rangle, |\varphi\rangle \in \mathbb{C}^N$.
(3) $U^\dagger U = I$ (here, † denotes conjugate transpose).
(4) The rows of U are an orthonormal basis for $\mathbb{C}^N$.
(5) The columns of U are an orthonormal basis for $\mathbb{C}^N$.

Such a matrix is called a *unitary matrix*.

Example 1.2 (Qubit)**.** The simplest example is a qubit, whose vector space is spanned by two basis vectors $|0\rangle$ and $|1\rangle$ (so it has two complex dimensions, or four real dimensions). Thus, the possible superpositions are $\alpha|0\rangle + \beta|1\rangle$ such that $|\alpha|^2 + |\beta|^2 = 1$. Often, but not always, α and β will be real, making them easier to draw; in this case, we just need $\alpha^2 + \beta^2 = 1$, defining a circle.

The *plus state* is $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, and the *minus state* is $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. $(|+\rangle, |-\rangle)$ is also an orthonormal basis for this space.

What are some unitary transformations? We have the identity

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

as well as the *NOT gate*

$$\text{NOT} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In general, a *gate* will refer to a unitary matrix that's applied to only one or a few qubits.

The identity and the NOT gate make sense for classical probability too, sending probability vectors to probability vectors. This is not true for the next matrix, called the *phase gate*:

$$\text{Phase} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which sends $(\alpha, \beta) \mapsto (\alpha, -\beta)$. There are other phases, e.g. replacing -1 by another root of unity. Similarly, the *Hadamard matrix* is

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

[1] This is what keeps the probabilities adding up to 1.

Notice that $H|0\rangle = |+\rangle$ and $H|1\rangle = |-\rangle$, but $H|+\rangle = |0\rangle$ and $H|-\rangle = |1\rangle$, so the Hadamard matrix switches the normal basis and the plus-minus basis. Thus, $H^2 = I$.

Finally, there are rotation matrices

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

This rotates counterclockwise by the angle θ.

Every unitary transformation in two dimensions is a product of rotations and reflections. Notice that the Hadamard matrix is not a rotation: we can apply $R = R_{\pi/4}$, which sends $|0\rangle \mapsto |+\rangle$ and $|1\rangle \mapsto -|-\rangle$. In general, $|\psi\rangle$ and $-|\psi\rangle$, as well as $i|\psi\rangle$, produce the same physical behavior: there's no experiment that can tell them apart. The classical analogue would be to move the whole universe twenty feet to the left: does anything actually change?

We can calculate $R|0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $R|1\rangle = (-|0\rangle + |1\rangle)/\sqrt{2}$. Evaluating on $|+\rangle$, we have to cancel out a $|0\rangle/2$ and a $-|0\rangle/2$:

$$R\frac{|0\rangle + |1\rangle}{\sqrt{2}} = \frac{\frac{|0\rangle+|1\rangle}{\sqrt{2}} + \frac{-|0\rangle+|1\rangle}{\sqrt{2}}}{\sqrt{2}} = |1\rangle.$$

This is called *destructive interference*: the two ways to obtain $|0\rangle$ were in opposite amplitudes, so they descrutively interfered, and cancelled out to zero probability. Similarly, the outcomes leading to 1 displayed *constructive interference*. A lot of the weirdness of quantum mechanics comes out of interference phenomena, since they behave so differently from classical mechanics. The *double-slit experiment* is an example: if a photon passes through two slits in an opaque material, there are alternating zones of light and dark, which makes classical sense. But the part that makes no sense classically is that if you close off one of the slits, photons can appear where they hadn't before.

This wasn't just a violation of intuition; it was a violation of the axioms of probability: classically, one assumes that the probabilities p_A and p_B of getting to that point after passing through the two slits A and B, respectively, should add to the total probability of a photon landing at that spot, but the experiment disproved that. This and its analogues in atomic nuclei, etc. are why quantum mechanics works with amplitudes instead of probabilities. In fact, the amplitude of this process is the sum of the amplitude occuring from slit A and the amplitude occuring from slit B. Destructive interference explains why closing slit B affects the answer.

Measurements. Given a qubit $\alpha|0\rangle + \beta|1\rangle$, we want to know whether it's 0 or 1. The rule is $\Pr[0] = |\alpha|^2$ and $\Pr[1] = |\beta|^2$. This is called *Born's rule*, after Max Born (who won his Nobel for work including this!). But the second, and very important, thing that happens is that the state "collapses" to whichever measurement you observed. This is much like some people one encounters: they're not certain about their opinion on a topic, but once they're asked about it, they pick an opinion and stick to it, at least until a unitary transformation is applied to them. This is why one says that measurement in quantum mechanics is an irreversible process.

In general, if we have a superposition of N outcomes $\alpha_1|1\rangle + \cdots + \alpha_N|N\rangle$, then $\Pr[i] = |\alpha_i|^2$. This is why global phase is irrelevant: the only way you can learn anything about a quantum system is measurement. Many of the paradoxes or misunderstandings people make implicitly assume there's some other way to measure the system. This also shows why there's no way to tell apart $|\psi\rangle$ and $-|\psi\rangle$: no measurement can distinguish them. It also explains interference: two amplitudes may both be nonzero, but if they're opposite in sign, the norm-squared of their sum is zero or nearly zero.

Measurement is denoted with a sort of speedometer $\pitchfork$.[2] Precomposing with a unitary transformation allows one to measure with respect to a different basis, e.g. using the Hadamard matrix is measurement in the $\{|+\rangle, |-\rangle\}$-basis. This means we'll get the outcome $+$ with probability $|\langle\psi|+\rangle|^2$ and outcome $-$ with probability $|\langle\psi|-\rangle|^2$. This is really just a rotation.

A pure 0 state always evaluates to 0. A pure 1 state always evaluates to 1. An equal superposition gives 0 half the time, and 1 half the time. But evaluating with respect to the $\{|+\rangle, |-\rangle\}$ basis turns pure states into equal superpositions and vice versa. In other words, $\Pr[|v_i\rangle] = |\langle\psi|v_i\rangle|^2$.

[2]This should be a semicircle with an arrow pointing to the upper right, but I don't know how to TEX that yet.

Example 1.3. We can generalize to systems of multiple qubits, placing them beside each other. A qubit might correspond to an electron with two energy states, or two spin directions (up and down), or any physical system that can be in either of two discrete states: quantum mechanics says there can also be a superposition. The variety of these systems leads to the variety of proposals for the physical architectures of a quantum computer.

Suppose now we have two photons. We refer to the composite of these systems with a tensor product: $(\alpha|0\rangle + \beta|1\rangle) \otimes (\gamma|0\rangle + \delta|1\rangle)$. We can distribute out the $\otimes$: the two-qubit space is actually spanned by the four basis vectors $\{|0\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |0\rangle, |1\rangle \otimes |1\rangle\}$.[3] The amplitudes are

$$(\alpha|0\rangle + \beta|1\rangle) \otimes (\gamma|0\rangle + \delta|1\rangle) = \alpha\gamma|00\rangle + \alpha\delta|01\rangle + \beta\gamma|10\rangle + \beta\delta|11\rangle.$$

This is a unit vector in $\mathbb{C}^4$.

Conversely, one might want to factor a state as a tensor product:

$$\frac{|00\rangle - |01\rangle - |10\rangle + |11\rangle}{2} = \frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

However, not every state can be factored, e.g. $(|00\rangle + |11\rangle)/\sqrt{2}$: if we expanded it out, too many terms would become 0, causing a contradiction. If a state can be written as a tensor product, it's called *separable*; a state that cannot be written in this way is *entangled*.[4] This is all that entanglement is, the quantum-mechanical version of correlation. Entanglement should not be changed by local unitary transformations, though it may be destroyed by measurement (see below). A global unitary transformation, involving both qubits, could entangle or unentangle qubits.

In general, separability arises when we have a state space $\mathbb{C}^{AB} = \mathbb{C}^A \otimes \mathbb{C}^B$. This can get more interesting in infinite-dimensional Hilbert spaces, but most of the spaces we consider in this class will be finite-dimensional, so just $\mathbb{C}^N$ for some N. Some quantum systems appearing in quantum optics arise not as tensor products, but as symmetric products, which can cause people to get tangled up talking about entanglement.

How do we measure in the two-qubit system? It's simple if you present a state as $a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$; the probability of $|00\rangle$ is $|a|^2$. Physically, though, this has surprising implications: the two qubits may be very far apart. If Alice has one and Bob has the other, then it's possible for Alice to only measure her qubit, which has state 0 with probability $\Pr[0] = |a|^2 + |b|^2$. If she observes 0, the two outcomes vanish: even before Bob can make a measurement, the state undergoes a *partial collapse* to $(1/\sqrt{|a|^2 + |b|^2})(a|0\rangle + b|1\rangle)$ (and something similar with c and d, if Alice sees 1). This may be generalized to systems of any dimension.

These statements are true for both separable and entangled qubits, but for separable qubits, they reduce to trivialities.

If Alice and Bob's qubits are in the Bell pair state, and Alice measures a 0, then she know that whenever Bob measures it, he will measure a 0 (and similarly, if she measures a 1, so must he). Bob's qubit "updates" instantaneously as soon as Alice measures it, no matter how far away they are. This is what famously unsettled Einstein, since it violates the relativistic principle that information cannot exceed the speed of light; this is so-called "spooky action at a distance." This doesn't seem useful for creating a faster-than-light telephone, since Alice has no control over the bit she sends. When you pick up a newspaper, that determines the headline on every copy of that newspaper, but that's not as spooky: there are other variables which explain the correlations without FTL travel. A similar result for the two qubits would be called a *local hidden-variable theory*, postulating a shared secret between the two qubits that explains the correlation.

It took about thirty years for the question to be formulated in this way and then to be answered.

Theorem 1.4 (No-communication theorem)**.** *It's not possible to use entangled states for faster-than-light communication.*

---

[3] In practice, to save effort, these are simplified: $|0\rangle \otimes |0\rangle$ is denoted $|0\rangle|0\rangle$ or even $|00\rangle$.

[4] There are ways to quantify the amount of entanglement; this state, sometimes called the *Bell pair*, *EPR pair*, or *singlet state*, happens to be maximally entangled.

So quantum mechanics does not break relativity. However, there is no hidden-variable theory, either: quantum mechanics is an intermediate point between the hidden-variable theory and true FTL communication. Yet if you wanted to simulate quantum mechanics in a classical universe, the simulation would need FTL communication.

Bell conducted an experiment that led to this conclusion, which was really an early phenomenon of a familiar concept in theoretical computer science, the two-prover game. There are three actors: Alice, Bob, and a referee. Alice and Bob cannot communicate, but the referee can send challenges to Alice and Bob and collect their responses. Alice and Bob are trying to cooperate, trying to get the referee to accept with the largest probability. They may plan a strategy in advance, but cannot communicate during the experiment, just like in a separated police interrogation.

In the modern reformulation of Bell's theorem, this game is called the *CHSH game*. The referee sends a random bit $x \in \{0,1\}$ to Alice and an independent random bit $y \in \{0,1\}$ to Bob. Alice sends back a random bit $a = a(x, r_a)$ and Bob sends back a random bit $b = b(y, r_b)$ (here, r_a and r_b are the sources of randomness for Alice and Bob, respectively). Alice and Bob win the game if $a + b = xy \pmod 2$.

Clearly this is not a game many people play for fun. Classically, Alice and Bob can win 3/4 of the time by always responding 0 — and one can prove that, classically, there is no strategy that does better, a fact called *Bell's inequality*.

But if Alice and Bob shared a Bell pair of two qubits in advance, there is a way of correlating their measurements in this state such that their probability of winning is $\cos^2(\pi/8) \approx 0.85$. This is a lot more subtle than sending messages back and forth: by themselves, Alice and Bob don't notice anything special, since you need both of their answers. In this case, there can be no local hidden-variable theory, and entanglement is not just shared classical randomness.

The protocol takes a little time to explain, but Alice measures her qubit in a specific basis if she sees a 0, and in a different basis if she sees a 1, and Bob does something similar. One can show that the probability of winning is $\cos^2$ of the difference of their measurement angles, which can be as high as $\pi/8$. A second inequality, called *Tsirelson's inequality*, says that no matter how many qubits they share, Alice and Bob cannot do better.

Any theory with local hidden variables predicts that the success probability is at most 3/4, but quantum mechanics doesn't, so quantum mechanics is not a hidden-variable theory. Bell never imagined his experiment to be actually carried out, but in the 1980s, people actually did this, and the universe is consistent with the predictions given by quantum mechanics. Most physicists weren't surprised: this was not the first experiment testing quantum mechanics, and it's passed all of them, and wasn't the last.

A neat slogan is that, like everything else, Bell's theorem comes down to interference influencing correlation. Bell's theorem, rather than getting into metaphysical questions, uses quantum entanglement to solve problems, and in this way anticipates the field of quantum communication.

Mixed states. Suppose Alice and Bob have qubits in a Bell pair. What state does Alice see? Naïvely, one might expect Alice to end up with $|+\rangle$, but if this were the case, then if she measured it in the $\{|+\rangle, |-\rangle\}$ basis, she should always get $|+\rangle$. So what does it mean for her to apply a Hadamard gate H to her qubit only? We take the tensor product $H \otimes I$: the Hadamard for Alice, and the identity for Bob. Often, the unitary operators we care about can be broken up into smaller components. One way of thinking about this: for all possible states of Bob's qubit ($|0\rangle$ and $|1\rangle$), Alice applies the Hadamard gate.

When Alice does this, the state looks like $(1/2)(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$: Alice observes $|0\rangle$ and $|1\rangle$ with equal probability. That's weird. The takeaway is that a rose (Bell pair) by any other name (basis) still smells as sweet (is still a Bell pair). Something new is happening: Alice's qubit is behaving more like a classical random bit than a quantum state.

To talk about a piece of an entangled system, one needs a more general description of states, called *mixed states*: these are probability distributions over different states. For example, as we just saw for the Bell pair, Alice's state is $1/2\ |0\rangle$ and $1/2\ |1\rangle$. To be clear: this is not a superposition, just a plain old random bit. This is a surprisingly classical form of uncertainly!

There's an important subtlety in mixed states, which is why people don't always think of them as probability distributions over pure states.[5] Specifically, there are different probability distributions that give rise to the same mixed state: Alice's mixed state is $1/2\ |0\rangle$ and $1/2\ |1\rangle$, but is indistinguishable from the

[5]A pure state is a degenerate mixed state, which assigns probability 1 to a single state.

mixed state $1/2\ |+\rangle$ and $1/2\ |-\rangle$: in any orthonormal basis, each of these produces each outcome half of the time. Writing out a mixed state as a distribution over pure states is redundant. Fortunately, there's a representation for mixed states that's not redundant, using what's called *density matrices*. These are a whole new (equivalent) way to view quantum mechanics itself, and is usually preferred by experimentalists.

Suppose I have a probability distribution of pure states $\{p_i, |\psi_i\rangle\}_{i=1}^n$; then, the corresponding density matrix is

$$\rho = \sum_{i=1}^{n} p_i |\psi_i\rangle\langle\psi_i|.$$

This is an $n \times n$ complex matrix.

For example, the density matrix for Alice's mixed state in the Bell pair is

$$\frac{1}{2}\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}.$$

You can compute that if we started with $(1/2, |+\rangle)$ and $(1/2, |-\rangle)$, we end up with the same matrix.

Generally, if we have an entangled system of the form $\sum \alpha_i |i\rangle|\psi_i\rangle$, then Bob's density matrix is

$$\rho_{\text{Bob}} = \sum_i |\alpha_i|^2 |\psi_i\rangle\langle\psi_i|.$$

This also eliminates global phase.

If U is a unitary matrix, then it acts on ρ by conjugation:

$$\rho \longmapsto \sum_i \rho_i U|\psi_i\rangle\langle\psi_i|U^\dagger = U\rho U^\dagger.$$

Another advantage of the density matrix is that measuring the mixed state in this basis just requires the diagonal entries: $\Pr[|i\rangle] = \rho_{ii}$.

That is, along the diagonal of a density matrix ρ, there's a probability distribution. Sometimes, that's all we have, and the density matrix is diagonal (including the Bell state). But density matrices may also have off-diagonal entries, e.g. the superposition $(1/\sqrt{2})(|0\rangle + |1\rangle)$. This is not the same, because in the $\{|+\rangle, |-\rangle\}$-basis, its outcome is always $|+\rangle$. Its density matrix is

$$\frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}.$$

Experimentalists regard off-diagonal entries as the signature of quantum behavior. In practice, the diagonal has the largest terms, but the bigger the off-diagonal terms are, the better the experiment was (according to the experimentalists).

The no-communication theorem says that quantum entanglement still preserves locality. More precisely, if Alice and Bob have entangled quantum systems, there is no combination of unitary transformations and measurements that Alice can make to her system that changes Bob's density matrix, unless we condition on Alice's measurement outcomes. This is very similar to how measurements affect classical correlation. Since Bob's density matrix can be used to calculate every possible outcome of every possible measurement Bob can make, this theorem encompasses anything Alice and Bob can do.

Density matrices don't provide us any new physics. Given a density matrix $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, there's an equivalent pure state

$$\sum_i \sqrt{p_i}|i\rangle \otimes |\psi_i\rangle,$$

and from the perspective of the second observer, these look the same.