Lecture 25, Thurs April 20: Hamiltonians

Now we'll move on to our second-to-last unit…

Hamiltonians and the Adiabatic Algorithm

We've seen how it's an open question whether quantum computers can solve **NP**-complete problems in polynomial time. If this turned out to be possible, it would be world-changing.

Like, it would be time for a Manhattan Project to build scalable quantum computers...

But if it turns out that quantum computers can't solve **NP**-complete problems in polynomial time, the question still remains, "How close can they get to solving?"

We know from the BBBV Theorem that any approach that ignores the structure of **NP**-complete problems will only yield the Grover speedup.

There have been many papers on arXiv.org that claim to solve **NP**-complete problems in polynomial time with a quantum computer, but do so in ways that violate this theorem.

Virtually all quantum computing papers can be found on arXiv.org, but the site has no peer review.

So to do better than Grover, we'd need to exploit problem structure in some way. For example, with Boolean satisfiability, we could imagine devising some quantum algorithm that dealt with certain parts of the formula first, and worried about other parts later.

If we managed to show that *any* **NP**-complete problem was in **BQP** (i.e., solvable in polynomial time by a quantum computer), then by definition, *all* of **NP** would be in **BQP**.

However, if we're talking about small speedups, then the choice of **NP**-complete problem might actually matter, because the process of reduction from one **NP**-complete problem to another might cancel out a speed advantage.

The Adiabatic Algorithm (Farhi, Goldstone, Gutmann, Sipser 2000)

is a famous attempt to do exactly the above--i.e., get a quantum speedup for **NP**-complete problems (conceivably, even an exponential speedup) by actually exploiting their structure.

It's an extremely important quantum algorithm, but unlike (say) Shor's or Grover's algorithms, it doesn't come with any rigorous analysis guaranteeing it will run fast in all cases--and indeed, we now know that it doesn't. To this day, no one really knows how useful this algorithm will be in practice.

It's something that people will eagerly experiment with, as soon as they have reliable large-scale quantum computers to test it on!

For some instances of optimization problems, the adiabatic algorithm might give a huge speed advantage, but for other instances it gives little or no advantage, or is even outperformed by classical algorithms. People are still trying to figure out for which types of instances the algorithm is most useful.

To understand the adiabatic algorithm, we first need to back up, and familiarize ourselves with a central concept in quantum mechanics called **Hamiltonians**.

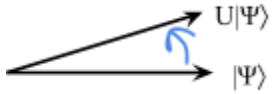
In a physics course on quantum mechanics, Hamiltonians would be day-one material,

while we'd only get to quantum computing and information at the very end (if at all). In this course, it's exactly the opposite!

Recall that unitaries are discrete linear transformations of quantum states:

$|\Psi\rangle \rightarrow U|\Psi\rangle$

But a physicist would treat time as *continuous*, and say that the state $|\Psi\rangle$ rotates continuously to $U|\Psi\rangle$ over some interval of time.



Hamiltonians are just the instantaneous time generators of unitary transformations. I.e., they're things that give rise to unitary transformations when you "leave them running" for some period of time. Like density matrices, Hamiltonians are described by *Hermitian matrices*. (But unlike density matrices, Hamiltonians don't need to be positive semidefinite or to have trace 1.)

Remember: for H to be Hermitian means that $H = H^{\dagger}$

From a physics perspective, the central equation of quantum mechanics is **Schrödinger's Equation:**

$i \frac{d|\Psi\rangle}{dt} = H|\Psi\rangle$ with H being some Hamiltonian.

This equation describes the evolution of an isolated quantum pure state in continuous time.

Well, the full version of Schrödinger's equation also includes the so-called *Planck's constant* $\hbar$, which is needed to convert between units of time and units of energy. But unless we're dealing with actual experimental data, involving meters, seconds, joules, and so forth, it's more convenient just to set $\hbar = 1$ -- the convention we'll adopt throughout! And for future reference: if it ever comes up, the speed of light c is also 1.

We can solve Schrödinger's equation, to find that the state after time t is

$|\Psi(t)\rangle = e^{-iHt}|\Psi(0)\rangle$.

Basically, we have here a whole *system* of linear differential equations--one for each coordinate of the vector $|\Psi\rangle$--but we can formally solve it by pretending that the matrix H is a scalar.

Here, though, we need to back up to address a mathematical point:

What does it mean to raise *e* to the power of a matrix?

The "right" definition turns out to be: you take the standard Taylor series for the exponential function,

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!},$$

and then just plug in a matrix instead of a number, to get a matrix-valued result.

To give some examples:

$$e\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad e\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix}$$

and

More generally, we can say that $$e^{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}} = \begin{bmatrix} e^{\lambda_1} & & \\ & & \\ & & e^{\lambda_n} \end{bmatrix}$$

I.e., you can exponentiate a diagonal matrix by exponentiating each diagonal term individually.

That might look like a very special case, but for Hermitian matrices, in some sense it's really all we need. Suppose we're given a matrix A, which can be written as $A = UDU^{-1}$ where D is diagonal. Then to compute e^A we can write:

$$e^A = \sum_{k=0}^{\infty} \frac{(UDU^{-1})^k}{k!} = \sum_{k=0}^{\infty} \frac{UD^kU^{-1}}{k!} = Ue^DU^{-1}$$

Thus we have a simple algorithm to exponentiate any diagonalizable matrix.

<u>What leverage do we get from H being Hermitian?</u>

For what we've said to make sense--and in particular, for it to be consistent with the discrete-time version of QM we've used in the rest of the course--it better be the case that the matrix e^{-iHt} is unitary. Let's now prove this to be true, by using the fact that H is Hermitian.

First claim: If H is Hermitian, then all its eigenvalues are real.

This is a special property of Hermitian matrices, although it's not "if and only if."

Proof: Suppose λ is an eigenvalue. Then by definition, there's some eigenvector $|v\rangle$ such that

$$H|v\rangle = \lambda|v\rangle,$$
$$\langle v|H|v\rangle = \lambda.$$

What do we get by conjugate-transposing the whole thing?

$$\langle v|H^\dagger|v\rangle = \overline{\lambda}$$

But since $H = H^\dagger$, this is equivalent to

$$\langle v|H|v\rangle = \overline{\lambda},$$
$$\langle v|H^\dagger|v\rangle = \lambda.$$

So $\lambda = \overline{\lambda}$, which means $\lambda \in \mathbb{R}$.

Something stronger is also true: every Hermitian matrix is diagonalizable.

One can prove that by induction on the dimension of the matrix. We won't go through the details here.

Generalizing these facts, the Spectral Theorem says that any Hermitian matrix can be written as

$H = UDU^\dagger$ with U being unitary and
D being diagonal and real

Now, to show that $e^{-i\mathrm{H}t}$ is unitary, just diagonalize H:

$$e^{-i\mathrm{H}t} = e^{-itUDU^{\dagger}}$$

$$= \mathrm{U}\, \boldsymbol{e^{-itD}}\, \mathrm{U}^{\dagger}$$

But $\boldsymbol{e^{-itD}} = \begin{bmatrix} e^{-it\lambda_1} & & \\ & & \\ & & e^{-it\lambda_n} \end{bmatrix}$

which is a diagonal unitary matrix (since $\lambda_1,...,\lambda_n$ are real). Hence $e^{-i\mathrm{H}t}$ is a unitary matrix.

Note that, if $|v\rangle$ is an eigenvector of H associated with the eigenvalue λ,

then $|v\rangle$ is also an eigenvector of $e^{-i\mathrm{H}t}$, associated with the eigenvalue $e^{-i\lambda t}$.

So, eigenvectors of H give rise to eigenvectors of the unitary.

Now, what about going backwards:

Given a unitary U, can we find a Hermitian matrix H such that $\mathrm{U} = e^{-i\mathrm{H}t}$?

Yes, this is not hard.

First diagonalize U, to get $\mathrm{U}=\mathrm{VDV}^{\dagger}$.

We then just need to take a logarithm of each diagonal entry of D:

That is, for each D_{jj}, find a λ such that $\mathrm{D}_{jj} = e^{-i\lambda t}$.

Will the λ that we get by solving this be unique?

No, because by Euler's formula, we can always add $2\pi i$ to the exponent and the equation will still hold.

We saw $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ that is a logarithm of the identity matrix. What else is?

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = e^{i\begin{bmatrix} 2\pi & 0 \\ 0 & 2\pi \end{bmatrix}}$$ and so on.

Thus, any given unitary can arise from infinitely many different Hamiltonians.

Physicists have a special name for the eigenvalues that you get by diagonalizing a Hamiltonian. They call them **energies**. Note that they're all real, and can therefore can be ordered from least to greatest:

$$\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$H = U\,U^{\dagger}$, with $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$

To each energy λ_j, there corresponds an **energy eigenstate** $|v_j\rangle$ such that $H|v_j\rangle = \lambda_j|v_j\rangle$.

Why are they called energies?

Because they *are* energies. These values are amounts of energy that the system can have.

Quantum mechanics gives us one explanation for why the concept of "energy" arises in physics: because unitary matrices arise by exponentiating Hamiltonians, and Hamiltonians can be diagonalized and have real eigenvalues.

If we apply the unitary transformation e^{-iHt} to the energy eigenstate $|v_j\rangle$, we get

$$e^{-iHt}\,|v_j\rangle = e^{-i\lambda_j t}\,|v_j\rangle.$$

Meaning that nothing happened, apart from the state picking up a global phase (unobservable by itself) dependent on the energy.

We can write an arbitrary state as a superposition over the energy eigenstates:

$$|\Psi\rangle = \alpha_1|v_1\rangle + \ldots + \alpha_n|v_n\rangle$$

From this perspective, applying the Hamiltonian H is equivalent to doing:

$$e^{-iHt}|\Psi\rangle = \alpha_1\, e^{-i\lambda_1 t}\,|v_1\rangle + \ldots + \alpha_n\, e^{-i\lambda_n t}\,|v_n\rangle$$

This presents a terrifyingly boring picture of the history of universe! It suggests that all that has ever happened, and all that ever *will* happen, is that the various energy eigenstates of the universe pick up phases, each rotating around the unit circle at a speed proportional to its energy.

From the utter lack of interesting activity when we view the world in the energy eigenbasis, we conclude that life is a basis-dependent phenomenon.

But the above picture is extremely useful. For one thing, it suggests that we simply *define* energy as the speed at which a quantum state picks up a phase.

It's not obvious that this corresponds to the usual conceptions of energy in physics, but it turns out that it does. You'll have to go to the physics department for details though!

One thing that *is* clear, from our definition, is that "energy is conserved." More formally: the expectation value of the energy in the state $|\Psi\rangle$, namely $\sum_j |\alpha_j|^2\,\lambda_j$, stays the same over time.

Let's give a few more definitions.

The energy eigenstate $|v_1\rangle$ corresponding to the lowest energy is the **ground state**, which as we'll see plays an extremely special role. The corresponding energy λ_1 is the **ground state energy**.

The energy eigenstate $|v_2\rangle$ corresponding to the second-lowest energy is the **first excited state**. The energy eigenstate $|v_3\rangle$ corresponding to the third-lowest energy is the **second excited state**, and so forth.

One detail: if $\lambda_1 = \lambda_2$, then we get what's called a **ground state degeneracy**: there's no longer a unique ground state, but a subspace of two or more dimensions, in which every state equally minimizes the energy. We can similarly get degenerate excited states, if $\lambda_j = \lambda_{j+1}$ for some larger j. For the most part, though, we'll be able to ignore this.

The standard game plan for much of modern physics goes like this:

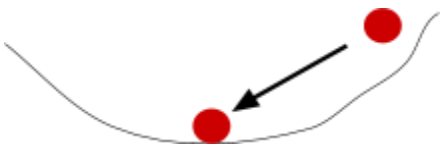
1. Start with the Hamiltonian H of your system.
2. Diagonalize H.
3. Get out the energy eigenstates.
4. Then, as a first guess, see if your system is just sitting in its ground state $|v_1\rangle$ doing nothing.

Why are quantum systems often found sitting in their ground states doing nothing?

Intuitively, because physical systems "like" to minimize their energy, so they tend to get into lower energy states. And the ground state, by definition, is the lowest they can go.

But since quantum mechanics is time-reversible, how is it even possible for a system to be "attracted" to a certain state?

Excellent question! You can thank the Second Law of Thermodynamics and the conservation of energy for this.



Note that the same question arises in classical physics, which is time-reversible too. If you leave a ball rolling around in a basin, then return a while later, you probably won't find it in an "excited state"--i.e., continuing to roll around. Whatever energy it had in its excited state, it could reach a lower energy by rolling downhill, slowing down, and giving off *heat* via *friction*.

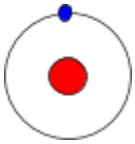
When this happens, the kinetic energy that used to be in the ball dissipates away in the heat.

In principle, it's *possible* that the reverse could happen, and that the heat could coalesce back into the ball and make it spontaneously move. But we essentially never observe that, and the reason comes down to entropy. For all the heat to coalesce back into motion would require an absurdly finely-tuned "conspiracy," whose probability of occurring by chance falls off exponentially with the number of particles in the ball. But the reverse process, motion dissipating into heat, requires no similar conspiracy: it only requires that our universe does contain low-entropy objects like balls. (This, in turn, can ultimately be traced back to the low entropy of the universe at the Big Bang--something that no one has satisfactorily explained in terms of anything deeper, but *we'll* be content to leave it there!)

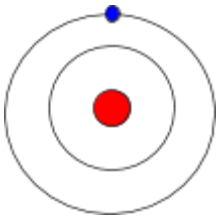
Pretty much exactly the same story works in the quantum case, and explains why, when we find quantum systems in Nature, they're often sitting in their ground states. (Namely because, if they weren't, then their interactions with surrounding systems would tend to carry away excess energy until they were.)

By contrast, all the quantum algorithms and protocols that we've seen in this course are examples of quantum systems that *don't* just sit in their ground states. Stuff happens; the system evolves!

We as people don't just sit in ground states either.



To give an example of these concepts that the physicists *really* love: the ground state of a hydrogen atom has the electron sitting in the lowest shell (the one closest to the nucleus).



The first excited state has the electron in the next shell up.

If the atom is in its first excited state, it can drop back down to its ground state via the electron emitting a photon. The photon carries away an amount of energy that's exactly equal to the difference between the ground and the first excited energies.

Conversely, a hydrogen atom in its ground state can jump up to its first excited state via the electron *absorbing* a photon. For this to happen, though, requires a photon to happen to hit the electron, which makes this process less likely than its reverse, for a hydrogen atom that's just sitting by itself in space somewhere. So, that's why hydrogen atoms in Nature are often found in their ground states.

Our final topic for this lecture is an extremely important operation that we can do with Hamiltonians, even though we could never do it with unitaries. Namely, we can add them!

Addition of Hamiltonians

$$H = H_0 + H_1$$

What does this mean physically?

Intuitively, it just means we've got two things going on at the same time. For example, H_0 and H_1 could correspond to two different forces acting on our system. To illustrate, at an *extremely* high level, we could write the Hamiltonian for the Standard Model of elementary particle physics as

$$H_{SM} = H_{Kinetic} + H_{EM} + H_{Strong} + H_{Weak}$$

Here $H_{Kinetic}$ is the Hamiltonian that would act even if there were no forces (corresponding to Newton's First Law of Motion), and H_{EM}, H_{Strong}, and H_{Weak} are the Hamiltonians corresponding to electromagnetism and to the strong and weak nuclear forces respectively.

Recall that the Standard Model excludes gravity.

This isn't exactly right for all sorts of reasons, but is good enough to get across the point.

Once we're adding Hamiltonians, we immediately face a mathematical question:

If A and B are matrices, is it generally the case that $e^{A+B} = e^A e^B$?

Alas, the answer is no. Indeed, it's not hard to find a 2×2 counterexample (exercise).

On the other hand, you can check using the Taylor series definition that *if* A and B commute (that is, AB=BA), then $e^{A+B} = e^A e^B$ *does* hold.

We'll care about unitary transformations like $e^{-it(H_1+H_2)}$, or their counterparts with many more than two H_i's. In particular, we'll need a way to apply these sorts of unitaries efficiently, given only the ability to apply H_1 and H_2 by themselves---even if H_1 and H_2 don't happen to commute.

Fortunately, there's a trick for this, known as **Trotterization**. The trick is to use the following approximation:

$$e^{A+B} \approx \underbrace{e^{\varepsilon A} e^{\varepsilon B} e^{\varepsilon A} e^{\varepsilon B} \ldots e^{\varepsilon A} e^{\varepsilon B}}_{\frac{1}{\varepsilon} \text{ times}}$$

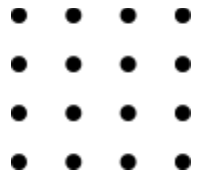
This basically means that we can achieve the same effect as A and B occurring simultaneously, by repeatedly switching between doing a tiny bit of A and a tiny bit of B.

We won't do it here, but it's possible to prove that the approximation improves as ε decreases, becoming an exact equality in the limit $\varepsilon \rightarrow 0$.

This is important for the question of how to simulate a real-world quantum system using a quantum computer. Indeed, the straightforward approach is just:

1. Discretize all the degrees of freedom (positions, momenta, etc.) that aren't already discrete.
2. Write the total Hamiltonian H acting on the system as a sum of "simple" terms (say, terms that act on only 1 or 2 particles as a time).
3. Trotterize H, in order to simulate it by a product of "simple" unitary transformations.

To flesh this out a bit more, we ought to say *something* about what the Hamiltonians of real physical systems tend to look like, at the level of abstraction relevant for this class.



Let's model the universe as a gigantic lattice of qubits, say in 2 or 3 dimensions (hey, it *is* a quantum computing class!). In that case, the total Hamiltonian that acts on the qubits can typically be written

$$H = \sum_j H_j + \sum_{j \sim k} H_{jk}$$

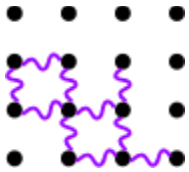
Here H_j is a Hamiltonian that acts only on the qubit j, and trivially on all the others, for example:

$$H_1 = \begin{bmatrix} h_1 & h_2 \\ h_3 & h_4 \end{bmatrix} \otimes \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} \otimes \cdots \otimes \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}$$

We can achieve this by, for each qubit, tensoring a 2×2 Hamiltonian on that qubit with the identity on all the other qubits--similar to what's done with unitary gates.

Meanwhile, H_{jk} is a Hamiltonian that acts on the neighboring qubits j and k, for example:

$$\begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}^{\otimes n-2}$$



This means that each qubit "talks" only to its immediate neighbors in the lattice, but evolving the Hamiltonian over time gives us effects that can propagate arbitrarily far.

$$e^{-i\mathrm{H}t} = \sum_{k=0}^{\infty} \frac{(-iHt)^k}{k!}$$

As soon as we've written this, though, we face a puzzle:
Won't this lead to faster-than-light travel?

Indeed, even when t is arbitrarily small, one can check that the unitary matrix $e^{-i\mathrm{H}t}$ will generically contain effects from *every* qubit in the lattice to every other one. Granted, the magnitude of these effects will fall off exponentially with distance, but causality demands that there should be literally *zero* effects propagating across the lattice faster than light.

So what's the resolution? Basically, it's just that the picture we're using comes from non-relativistic quantum mechanics, so it yields a good approximation only if the relevant speeds are small compared to the speed of light. When the speeds are larger, we need the framework of quantum field theory, which *does* entail that faster-than-light influences are exactly zero.

OK, now we're ready to set things up for the next lecture. Suppose that H, a Hamiltonian acting on n qubits, is the sum of many "simple" Hamiltonians acting on a few qubits each:

$$\mathrm{H} = \mathrm{H}_1 + \ldots + \mathrm{H}_m.$$

Because H is a $2^n \times 2^n$ matrix, figuring out its ground state (or ground states) by brute force could be extremely time-consuming. Which leads to a question:

If I know the ground state of the H_j's individually, can I combine them in some simple way to get the ground state of H itself?

Alas, the answer is almost certainly "no." More precisely, we claim that finding the ground state of a Hamiltonian of this form is an **NP**-hard problem. To prove this, we'll show how to take any instance of the famous 3SAT problem, and encode it into the ground state problem. Thus, suppose we have a Boolean formula in n variables,

$$\varphi(x_1, \ldots, x_n) = c_1 \wedge \ldots \wedge c_m$$

where each clause c_i acts on at most 3 variables.

We can now define a Hamiltonian H as follows: H will act on n qubits, and will contain one term H_i for each clause c_i. The term H_i will act on at most 3 qubits, corresponding to the bits acted on by c_i, and will impose an "energy penalty," say of 1, if and only if the qubits are set in such a way that c_i is violated. For example, you could encode the clause $(v_1 \vee v_2 \vee v_3)$ as the Hamiltonian

$$h = \begin{bmatrix} 1 & & & & & & & \\ & 0 & & & & & & \\ & & 0 & & & & & \\ & & & 0 & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & & & & & 0 & \\ & & & & & & & 0 \end{bmatrix}$$

This Hamiltonian will yield 0 (no energy penalty) unless all three qubits are set to 0, which corresponds to the case where the clause is not satisfied.

It's not hard to see that, to minimize the energy for such a Hamiltonian, without loss of generality you just set all the variables classically, and see which setting leads to the smallest energy.

In other words: you solve 3SAT!

In particular, H's ground-state energy will be 0 if φ is satisfiable, or positive if φ is unsatisfiable.

So not only is this problem **NP**-hard, it's **NP**-hard for reasons that have nothing to do with quantum mechanics.

Could this observation somehow let Nature solve **NP**-hard problems in polynomial time? For more, tune in next time!