CS 294-2
10/28/04
Fall 2004

Hamiltonians, Dirac Eqn., Entropic Uncertainty Principle

Lecture 15

Recall that the Schrödinger equation that governs the time evolution of a system is:

$$i\frac{d|\Psi(t)\rangle}{dt} = H|\Psi(t)\rangle \ ,$$

where $|\Psi(t)\rangle$ is the state of the system at time t, and H, the Hamiltonian, is a Hermitian operator, i.e. $H = H^\dagger$, and therefore H has an orthogonal set of eigenvectors $\left|e_j\right\rangle$ with real eigenvalues λ_j. For example, if system has n qubits H is 2^nx2^n Hermitian matrix. If $|\Psi(0)\rangle = \left|e_j\right\rangle$ then:

$$|\Psi(t)\rangle = \alpha(t)\left|e_j\right\rangle$$

where $i\frac{d\alpha(t)}{dt} = \lambda_j$. Therefore:

$$|\Psi(t)\rangle = e^{i\lambda_i t}\left|e_j\right\rangle .$$

Thus the time evolution operator U is unitary and can be written in its eigenbasis $\left|e_j\right\rangle$ as:

$$U = \begin{bmatrix} e^{i\lambda_0 t} & \cdots & \cdots \\ \vdots & \ddots & \vdots \\ \cdots & \cdots & e^{i\lambda_{n-1}t} \end{bmatrix} = e^{iHt} = \sum_k \frac{(iHt)^k}{k!}$$

Recall that in coined quantum walk we applied same unitary transformation U, such as the Hadamard transform:

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

to the coin qubit $|b\rangle$ at each step of the walk. This is a quantum coin flip in the sense that it transforms $|0\rangle$ to some superposition of $|0\rangle$ and $|1\rangle$ (actually to $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$) and $|1\rangle$ to $\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$. In terms of Hamiltonians, we can achieve a similar effect by applying Hamiltonian:

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

since H transform $|0\rangle$ to $|1\rangle$ and vice versa.

1 Dirac Equation

Recall that for a free particle on a line the Schrödinger's equation says:

$$i\frac{d|\Psi(x,t)\rangle}{dt} = -\frac{\hbar}{2m}\frac{d^2|\Psi(x,t)\rangle}{dx^2}.$$

Here, the energy operator $\frac{d^2|\Psi\rangle}{dx^2}$ specifies the kinetic energy of the particle since we included no potential term. To understand this further let us consider the classical situation. There

$$E = \frac{mV^2}{2} = \frac{p^2}{2m},$$

where $p = mV$ is particle's momentum.

Quantum analog of this is the operator:

$$\hat{p} = \frac{\hbar d}{i\, dx}.$$

Aside: Why is $\hat{p} \sim \frac{\hbar d}{i\, dx}$? Should not it be the derivative with respect to time? This is because Schrödinger equation tells us how the particle's instantaneous state $|\Psi\rangle$ evolves. Recall from previous lecture that the states of definite velocity are Fourier basis functions, i.e $|\Psi(x,t)\rangle = e^{i(\lambda x+\omega t)}|x\rangle$. Thus particle's velocity is $\frac{d\Psi(x,t)}{dx}$, making the momentum operator to be the derivative of wave function with respect to position.

But so far we only had the had the non-relativistic picture both classically and quantumly. Classically relativistic correction is given by equation:

$$E^2 = p^2c^2 + m^2c^4.$$

If we assume $p << mc$ then:

$$E = \sqrt{p^2c^2 + m^2c^4} = mc^2\sqrt{1 + \frac{p^2c^2}{m^2c^4}} \simeq mc^2(1 + \frac{p^2c^2}{2m^2c^4}) = \frac{p^2}{2m} + mc^2.$$

By analogy in quantum case the square of Hamiltonian satisfies:

$$H^2 = \hat{p}^2c^2 + m^2c^4\hat{I}.$$

But the question is how to find the the square root of this operator? This puzzled physicists until Dirac suggested the following novel solution. To find the square root, expand the Hilbert space by tensoring it additional qubit, i.e. spin of the particle, and consider the operator:

$$H = \begin{pmatrix} \hat{p}^2 + mI & 0 \\ 0 & \hat{p}^2 + mI \end{pmatrix},$$

assuming that the variables are scaled such that $c = 1$. This has a square root:

$$H = \begin{pmatrix} \hat{p} & mI \\ mI & -\hat{p} \end{pmatrix}.$$

To understand the behavior of the particle let's first understand the case $m = 0$. Then Schrödinger equation becomes:

$$i\frac{d|\Psi\rangle}{dt} = H|\Psi\rangle = \begin{pmatrix} \hat{p}^2 & 0 \\ 0 & -\hat{p}^2 \end{pmatrix}\begin{pmatrix} |\Psi_0\rangle|0\rangle \\ |\Psi_1\rangle|1\rangle \end{pmatrix}$$

So, we can decompose this equation to separately consider the evolution of $|\Psi_0\rangle$ and $|\Psi_1\rangle$. The states now evolve in accord to:

$$\frac{d|\Psi_0\rangle}{dt} = \frac{d|\Psi_0\rangle}{dx} \Longrightarrow |\Psi_0\rangle = f(x-t).$$

For an arbitrary function of f, i.e. $|\Psi_0\rangle$ moves toward the right at the unit speed. Since we assumed $c = 1$ the particle moves toward right at the speed of light. And similarly we get the second solution:

$$|\Psi_1\rangle = f(x+t) \ .$$

In the general case when $m \neq 0$ Hamiltonian can be divided in two parts:

$$H = H_1 + H_2 = \begin{pmatrix} \hat{p}^2 & 0 \\ 0 & -\hat{p}^2 \end{pmatrix} + mI \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} ,$$

the second part of the Hamiltonian does not change the position of the particle and operates as a coin flip with the intensity proportional with the mass of the particle. So after time t our evolution is exactly a quantum walk on a line at expected distance t from the origin. However mass of the particle will actually determine how much the coin is flipped. Classical walks make only $\sqrt{n}$ distance in n steps. Dirac showed that quantum walk move $\Omega(n)$ away from origin. The actual distance depends on constant drift velocity determined with frequency of coin flips which is dependent upon mass. As quantum walk moves away from the origin step length converges to zero. Note also that the explained quantum walk would be exactly the same as its classical counterpart if we were measuring after each application of H. However in quantum case we get the interference effects since phases cancel out and we end up far from the origin. In classical case we would end up with binomial distribution around the origin. This interference effects in quantum case are the main mechanism behind the quantum speed up.

Simulating Schrödinger's equation: What is the state after time t? Suppose that given Hamiltonian H operating on initial state $|\Psi\rangle$ of n qubits gives us after time t the following state:

$$|\Psi(t)\rangle = e^{iHt}|\Psi_0\rangle = \begin{bmatrix} e^{i\lambda_0 t} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{i\lambda_{n-1}t} \end{bmatrix} |\Psi_0\rangle \ ,$$

assuming that we start in eigenbasis of H in which H is diagonal. Since eigenstates $|x\rangle$ are known to us we can easily compute eignevalues λ_j. Then $|\Psi(t)\rangle$ can be represented as follows:

$$|\Psi(t)\rangle = \sum_{x\varepsilon\{0,1\}^n} \alpha_x |x\rangle |\lambda_x\rangle$$

Since eigenvalues are known to us we can apply controlled phase gate to get: in order to the superposition:

$$|\Psi(t+\delta t)\rangle = \sum \alpha_x e^{i\lambda_x \delta t} |x\rangle |\lambda_x\rangle \ ,$$

from this if we trace out $|\lambda_x\rangle$ we acquire the update $|\Psi(t+\delta t)\rangle$.

How does the wave function updates from one superposition to the other as the time evolves? Since in general Hamiltonian operates on the particle in the particle on the line in a potential field, we will have:

$$i\frac{d|\Psi(x,t)\rangle}{dt} = -\frac{\hbar}{2m}\frac{d^2|\Psi(x,t)\rangle}{dx^2} + V(x)|\Psi\rangle \ .$$

Now the simulation of this two term Hamiltonian would be hard and nonlocal if we were to simulate it in the basis in which total Hamiltonian is diagonal. Instead, since Hamiltonian now consists of two terms each of which is diagonal either in the position basis or velocity basis we can do the following. We can update the evolution locally in the position basis, than update in the momentum basis, in which the update corresponds to phase change and then we can move back to the position basis and so forth. To update the evolution locally we would need to look only to the neighboring particles since we operate in diagonal basis. This change of update basis would actually take place even in the classical simulation of given Hamiltonian. In general Hamiltonians are diagonal in more than two basis.

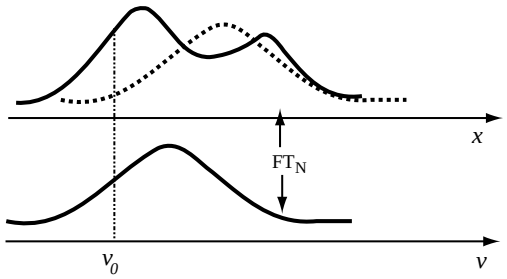

Figure 1: *Evolution of the Hamiltonian in two diagonal basis.*

2 Entropic version of uncertainty principle

The uncertainty principle is a statement about the fact that a superposition and its Fourier transform can not both be sharply concentrated. This holds under various measures of concentration. Thus we will show an entropic version of this statement for the Fourier transform over any finite abbelian group $G = \mathbb{Z}_N$. Given $|\Psi(t)\rangle$ and its Fourier transformation $\left|\hat{\Psi}(t)\right\rangle$:

$$|\Psi(t)\rangle = \sum_{x\varepsilon\{0,1\}^n} \alpha_x |x\rangle \xrightarrow{FT_N} \sum \hat{\alpha}_x |x\rangle$$

we will have probability distribution D that results from measuring $|\Psi(t)\rangle$ in standard basis, and distribution $\hat{D}$ that results from measuring $\left|\hat{\Psi}(t)\right\rangle$ in Fourier basis.

Theorem: Sum of entropies satisfies inequality:

$$S(D) + S(\hat{D}) \geq \log_2 N \, .$$

Example: If $G = \mathbb{Z}_N$ Fourier transformation on a subgroup $H \subseteq G$, $|H| = k$ gives us $H^\perp$ where $|H^\perp| = N/k$. In general we would apply FT_N on coset H with phases X_l. Than FT_N moves phases information X_l to cosets H_l and coset information H_g to phase X_g. So we would have corresponding entropies: $S(D) = \log_2 k$ and $S(\hat{D}) = \log_2(N/k)$, that will sum to:

$$S(D) + S(\hat{D}) = \log_2 k + \log_2 \frac{N}{k} = \log_2 N \, .$$

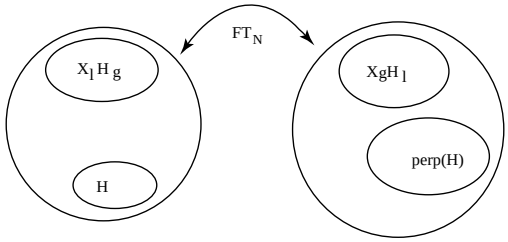

Figure 2: *Relationship between standard and Fourier domain.*

Proof: The proof consists of two parts:

(1) $S(D) + S(\hat{D}) \geq \log_2 N - \log_2 \log_2 N$,

(2) the lower order term is zero so that: $S(D)+S(\hat{D}) \geq \log_2 N$

Proof of the second fact: Consider group $K = G^n$ and n independent copies of state $|\Psi\rangle$. This tensor product will have Fourier transformation: $|\Psi\rangle^{\otimes n} \overset{FT_N}{\longleftrightarrow} |\Psi\rangle^{\otimes n}$, i.e. distribution D^n gets transformed to distribution $\hat{D}^n$. Now the sum of new entropies is:

$$nS(D)+nS(\hat{D}) = \log_2 N^n - \log_2 \log_2 N^n =$$

$$= n\log_2 N - \log_2(n\log_2 N) = n\log_2 N - \log_2 n - \log_2\log_2 N\,,$$

$$S(D)+S(\hat{D}) \geq \frac{n\log_2 N - \log_2 n - \log_2\log_2 N}{n} =$$

$$= \log_2 N - \frac{\log_2 n}{n} - \frac{\log_2\log_2 N}{n} \overset{n\to\infty}{\longrightarrow} \log_2 N\,.$$

Sketch of the proof for the first fact: set a bipartite graph according the distributions $D \sim \|\alpha_x\|^2$, and $\hat{D} \sim \|\hat{\alpha}_x\|^2$. We put an edge between say verteces x and y only if:

$$\|\alpha_x\|^2\|\beta_x\|^2 \leq \frac{\gamma}{N}\,,$$

for some $\gamma > 1$. Assume that `minCut` $\simeq 1$ then according to `maxFlow minCut` theorem `maxFlaw` $=$ `minCut` $\simeq 1$. Now since:

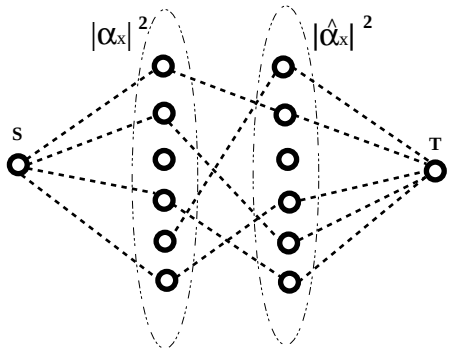

Figure 3: *MaxFlow = MinCut between S and T.*

$$S(D)+S(\hat{D}) = \mathbb{E}_x\left[\log_2\frac{1}{\|\alpha_x\|^2}\right] + \mathbb{E}_y\left[\log_2\frac{1}{\|\hat{\alpha}_x\|^2}\right]$$

Then every therm is paired and expectation over all edges will give us almost $\log_2 N$.

Why is the flow large? Look only at large terms $\|\alpha_x\|$, say: $\|\alpha_x\|^2 \geq p$. Worst case would be when this coefficients are connected to all the small coefficients, making the `mincut` to be large.