8.06 Spring 2016 Lecture Notes

2. Time-dependent approximation methods

Aram Harrow

Last updated: March 12, 2016

Contents

1 Time-dependent perturbation theory

Perturbation theory can also be used to analyze the case when we have a large static Hamiltonian H_0 and a small, possibly time-dependent, perturbation $\delta H(t)$. In other words

$$H(t) = H_0 + \delta H(t). \tag{1}$$

However, the more important difference from time-independent perturbation theory is in our *goals*: we will seek to analyze the dynamics of the wavefunction (i.e. find $|\psi(t)\rangle$ as a function of t) rather than computing the spectrum of H. In fact, when we use a basis, we will work in the eigenbasis of H_0. For example, one common situation that we will analyze is that we start in an eigenstate of H_0, temporarily turn on a perturbation $\delta H(t)$ and then measure in the eigenbasis of H_0. This is a bit abstract, so here is a more concrete version of the example. H_0 is the natural Hamiltonian of

the hydrogen atom and $\delta H(t)$ comes from electric and/or magnetic fields that we temporarily turn on. If we start in the $1s$ state, then what is the probability that after some time we will be in the $2p$ state? (Note that the very definition of the states depends on H_0 and not the perturbation.) Time-dependent perturbation theory will equip us to answer these questions.

1.1 Rotating frame

We want to solve the time-dependent Schrödinger equation $i\hbar\partial_t|\psi(t)\rangle = H(t)|\psi\rangle$. We will assume that the dynamics of H_0 are simple to compute and that the computational difficulty comes from $\delta H(t)$. At the same time, if H_0 is much larger than $\delta H(t)$ then most of the change in the state will come from H_0. In classical dynamics when an object is undergoing two different types of motion, it is often useful to perform a change of coordinates to eliminate one of them. We will do the same thing here. Define the state

$$|\tilde{\psi}(t)\rangle = e^{\frac{iH_0 t}{\hbar}}|\psi(t)\rangle. \tag{2}$$

We say that $|\tilde{\psi}(t)\rangle$ is in the *rotating frame* or alternatively the *interaction picture*. Multiplying by $e^{\frac{iH_0 t}{\hbar}}$ cancels out the natural time evolution of H_0. In particular, if $\delta H(t) = 0$ then we would have $|\tilde{\psi}(t)\rangle = |\tilde{\psi}(0)\rangle = |\psi(0)\rangle$. Thus, any change in $|\tilde{\psi}(t)\rangle$ must come from $\delta H(t)$.

Aside: comparison to Schrödinger and Heisenberg pictures. In 8.05 we saw the Schrödinger picture and the Heisenberg picture. In the former, states evolve according to H and operators remain the same; in the latter, states stay the same and operators evolve according to H. The interaction picture can be thought of as intermediate between these two. We pick a frame rotating with H_0, which means that the operators evolve according to H_0 and the states evolve with the remaining piece of the Hamiltonian, namely δH. As we will see below, to calculate this evolution correctly we need δH to rotate with H_0, just like all other operators. This is a little vague but below we will perform an exact calculation to demonstrate what happens.

Now let's compute the time evolution of $|\tilde{\psi}(t)\rangle$.

$$\begin{aligned}
i\hbar\frac{d}{dt}|\tilde{\psi}(t)\rangle &= i\hbar\frac{d}{dt}\left(e^{\frac{iH_0 t}{\hbar}}|\psi(t)\rangle\right) \\
&= -H_0 e^{\frac{iH_0 t}{\hbar}}|\psi(t)\rangle + e^{\frac{iH_0 t}{\hbar}}(H_0 + \delta H(t))|\psi(t)\rangle \\
&= e^{\frac{iH_0 t}{\hbar}}\delta H(t)|\psi(t)\rangle && \text{since } H_0 \text{ and } e^{\frac{iH_0 t}{\hbar}} \text{ commute} \\
&= \underbrace{e^{\frac{iH_0 t}{\hbar}}\delta H(t)e^{-\frac{iH_0 t}{\hbar}}}_{\widetilde{\delta H}(t)}|\tilde{\psi}(t)\rangle
\end{aligned}$$

Thus we obtain an effective Schrödinger equation in the rotating frame

$$\boxed{i\hbar\frac{d}{dt}|\tilde{\psi}(t)\rangle = \widetilde{\delta H}(t)|\tilde{\psi}(t)\rangle} \tag{3}$$

where we have defined

$$\widetilde{\delta H}(t) = e^{\frac{iH_0 t}{\hbar}}\delta H(t)e^{-\frac{iH_0 t}{\hbar}}.$$

This has a simple interpretation as a matrix. Suppose that the eigenvalues and eigenvectors of H_0 (reminder: we work with the eigenbasis of H_0 and *not* H(t)) are given by

$$H_0|n\rangle = E_n|n\rangle.$$

Define $\delta H_{mn}(t) = \langle m|\delta H(t)|n\rangle$. Then

$$\widetilde{\delta H}_{mn}(t) = \langle m|e^{\frac{iH_0 t}{\hbar}}\delta H(t)e^{-\frac{iH_0 t}{\hbar}}|n\rangle = e^{\frac{i(E_m-E_n)t}{\hbar}}\delta H_{mn}(t) \equiv e^{i\omega_{mn}t}\delta H_{mn},$$

where we have defined $\omega_{mn} = \frac{E_m-E_n}{\hbar}$. If we define $c_n(t)$ according to

$$|\tilde{\psi}(t)\rangle = \sum_n c_n(t)|n\rangle \qquad \implies \qquad |\psi(t)\rangle = \sum_n e^{-\frac{iE_n t}{\hbar}}c_n(t)|n\rangle$$

then we obtain the following coupled differential equations for the $\{c_n\}$.

$$i\hbar\dot{c}_m(t) = \sum_n \widetilde{\delta H}_{mn}(t)c_n(t) = \sum_n e^{i\omega_{mn}t}\delta H_{mn}c_n(t).$$

1.2   Perturbation expansion

So far everything has been exact, although sometimes this is already enough to solve interesting problems. But often we will need approximate solutions. So assume that $\delta H(t) = O(\lambda)$ and expand the wavefunction in powers of λ, i.e.

$$
\begin{array}{ccccccc}
c_m(t) & = & c_m^{(0)}(t) & + & c_m^{(1)}(t) & + & c_m^{(2)}(t) & +\ldots \\
 & & O(1) & & O(\lambda) & & O(\lambda^2) & \\
|\tilde{\psi}(t)\rangle & = & |\tilde{\psi}^{(0)}(t)\rangle & + & |\tilde{\psi}^{(1)}(t)\rangle & + & |\tilde{\psi}^{(2)}(t)\rangle & +\ldots
\end{array}
$$

We can solve these order by order. Applying (3) we obtain

$$\underbrace{i\hbar\partial_t|\tilde{\psi}^{(0)}(t)\rangle}_{O(1)} + \underbrace{i\hbar\partial_t|\tilde{\psi}^{(1)}(t)\rangle}_{O(\lambda)} + \underbrace{i\hbar\partial_t|\tilde{\psi}^{(2)}(t)\rangle}_{O(\lambda^2)} + \ldots = \underbrace{\widetilde{\delta H}(t)|\tilde{\psi}^{(0)}(t)\rangle}_{O(\lambda)} + \underbrace{\widetilde{\delta H}(t)|\tilde{\psi}^{(1)}(t)\rangle}_{O(\lambda^2)} + \ldots \quad (4)$$

The solution is much simpler than in the time-dependent case. There is no zeroth order term on the RHS, so the zeroth order approximation is simply that nothing happens:

$$\boxed{|\tilde{\psi}^{(0)}(t)\rangle = |\tilde{\psi}^{(0)}(0)\rangle = |\psi(0)\rangle} \qquad (5)$$

The first-order terms yield

$$i\hbar\partial_t|\tilde{\psi}^{(1)}(t)\rangle = \widetilde{\delta H}(t)|\tilde{\psi}^{(0)}(t)\rangle = \widetilde{\delta H}(t)|\psi(0)\rangle. \qquad (6)$$

Integrating, we find

$$\boxed{|\tilde{\psi}^{(1)}(t)\rangle = \int_0^t dt' \frac{\widetilde{\delta H}(t')}{i\hbar}|\psi(0)\rangle}. \qquad (7)$$

This leads to one very useful formula. If we start in state $|n\rangle$, turn on $H(t)$ for times $0 \le t \le T$ and then measure in the energy eigenbasis of H_0, then we find that the probability of ending in state $|m\rangle$ is

$$P_{n\to m} = \left|\int_0^t dt' \frac{\delta H_{mn}(t')e^{i\omega_{mn}t'}}{i\hbar}\right|^2$$

We can also continue to higher orders. The second-order solution is

$$\boxed{|\tilde{\psi}^{(2)}(t)\rangle = \int_0^t dt' \int_0^{t'} dt'' \frac{\widetilde{\delta H}(t')}{i\hbar}\frac{\widetilde{\delta H}(t'')}{i\hbar}|\psi(0)\rangle}. \qquad (8)$$

3

1.3 NMR

In some cases the rotating frame already helps us solve nontrivial problems *exactly* without going to perturbation theory. Suppose we have a single spin-1/2 particle in a magnetic field pointing in the $\hat{z}$ direction. This field corresponds to a Hamiltonian

$$H_0 = \omega_0 S_z = \frac{\hbar}{2}\omega_0 \sigma_z.$$

If the particle is a proton (i.e. hydrogen nucleus) and the field is typical for NMR, then ω_0 might be around 500 MHz.

Static magnetic field Now let's add a perturbation consisting of a magnetic field in the $\hat{x}$ direction. First we will consider the *static* perturbation

$$\delta H(t) = \Omega S_x,$$

where we will assume $\Omega \ll \omega_0$, e.g. Ω might be on the order of 20 KHz. (Why are we considering a time-independent Hamiltonian with time-dependent perturbation theory? Because really it is the time-dependence of the *state* and not the Hamiltonian that we are after.)

We can solve this problem exactly without using any tools of perturbation theory, but it will be instructive to compare the exact answer with the approximate one. The exact evolution is given by precession about the

$$\frac{\omega_0 \hat{z} + \Omega \hat{x}}{\sqrt{\omega_0^2 + \Omega^2}}$$

axis at an angular frequency of $\sqrt{\omega_0^2 + \Omega^2}$. If $\Omega \ll \omega_0$ then this is very close to precession around the $\hat{z}$ axis.

Now let's look at this problem using first-order perturbation theory.

$$\widetilde{\delta H}(t) = e^{i\frac{\omega_0 t}{2}\sigma_z}\Omega S_x e^{-i\frac{\omega_0 t}{2}\sigma_z} = \Omega\left(\cos(\omega_0 t)S_x - \sin(\omega_0 t)S_y\right)$$

$$|\tilde{\psi}^{(1)}(t)\rangle = \int_0^t dt' \frac{\widetilde{\delta H}(t')}{i\hbar}|\psi(0)\rangle$$

$$= \int_0^t dt' \frac{\Omega}{i\hbar}\left(\cos(\omega_0 t')S_x - \sin(\omega_0 t')S_y\right)|\psi(0)\rangle$$

$$= \frac{1}{i\hbar}\frac{\Omega}{\omega_0}\left(\sin(\omega_0 t)S_x + (\cos(\omega_0 t) - 1)S_y\right)|\psi(0)\rangle$$

We see that the total change is proportional to Ω/ω_0, which is $\ll 1$. Since this is the difference between pure rotating around the $\hat{z}$ axis, this is consistent with the exact answer we obtained.

The result of this calculation is that if we have a strong $\hat{z}$ field, then adding a weak static $\hat{x}$ field doesn't do very much. If we want to have a significant effect on the state, we will need to do something else. The rotating-frame picture suggests the answer: the perturbation should rotate along with the frame, so that in the rotating frame it appears to be static.

Rotating magnetic field Now suppose we apply the perturbation

$$\delta H(t) = \Omega\left(\cos(\omega_0 t)S_x + \sin(\omega_0 t)S_y\right).$$

We have already computed $\tilde{S}_x$ above. In the rotating frame we have

$$\tilde{S}_x = (\cos(\omega_0 t)S_x - \sin(\omega_0 t)S_y)$$
$$\tilde{S}_y = (\cos(\omega_0 t)S_y + \sin(\omega_0 t)S_x)$$

Thus

$$\widetilde{\delta H}(t) = \Omega S_x.$$

The rotating-frame solution is now very simple:

$$|\tilde{\psi}(t)\rangle = e^{-\frac{i\Omega t}{2}\sigma_x}|\psi(0)\rangle.$$

This can be easily translated back into the stationary frame to obtain

$$|\psi(t)\rangle = e^{-\frac{i\omega_0 t}{2}\sigma_z}e^{-\frac{i\Omega t}{2}\sigma_x}|\psi(0)\rangle.$$

1.4 Periodic perturbations

The NMR example suggests that transitions between eigenstates of H_0 happens most effectively when the perturbation rotates at the frequency ω_{mn}. We will show that this holds more generally in first-order perturbation theory. Suppose that

$$\delta H(t) = V\cos(\omega t),$$

for some time-independent operator V. If our system starts in state $|n\rangle$ then at time t we can calculate

$$c_m^{(1)}(t) = \langle m|\tilde{\psi}^{(1)}(t)\rangle = \int_0^t dt' \frac{\widetilde{\delta H}_{mn}(t')}{i\hbar}$$
$$= \int_0^t dt' \frac{\delta H_{mn}(t')}{i\hbar}e^{i\omega_{mn}t'}$$
$$= \int_0^t dt' \frac{V_{mn}}{i\hbar}\cos(\omega t)e^{i\omega_{mn}t'}$$
$$= \frac{V_{mn}}{2i\hbar}\int_0^t dt' \left(e^{i(\omega_{mn}+\omega)t'} + e^{i(\omega_{mn}-\omega)t'}\right)$$
$$= \frac{V_{mn}}{2i\hbar}\left[\frac{e^{i(\omega_{mn}+\omega)t}-1}{\omega_{mn}+\omega} + \frac{e^{i(\omega_{mn}-\omega)t}-1}{\omega_{mn}-\omega}\right]$$

The $\omega_{mn}\pm\omega$ terms in the denominator mean that we will get the largest contribution when $\omega \approx |\omega_{mn}|$. (A word about signs. By convention we have $\omega > 0$, but ω_{mn} is a difference of energies and so can have either sign.) For concreteness, let's suppose that $\omega \approx \omega_{mn}$; the $\omega \approx -\omega_{mn}$ case is similar. Then we have

$$c_m^{(1)}(t) \approx \frac{V_{mn}}{2i\hbar}\frac{e^{i(\omega_{mn}-\omega)t}-1}{\omega_{mn}-\omega}.$$

If we now measure, then the probability of obtaining outcome m is

$$P_{n\to m}(t) \approx |c_m^{(1)}(t)|^2 = \frac{|V_{mn}|^2}{\hbar^2}\frac{\sin^2\left(\frac{(\omega_{mn}-\omega)t}{2}\right)}{(\omega_{mn}-\omega)^2} = \frac{|V_{mn}|^2}{\hbar^2}\frac{\sin^2\left(\frac{\alpha t}{2}\right)}{\alpha^2},$$

where we have defined the *detuning* $\alpha \equiv \omega_{mn} - \omega$. The t, α dependence is rather subtle, so we examine it separately. Define

$$f(t, \alpha) = \frac{\sin^2\left(\frac{\alpha t}{2}\right)}{\alpha^2} \tag{9}$$

For fixed α, $f(t, \alpha)$ is periodic in t.

It is more interesting to consider the case of fixed t, for which f has a sinc-like appearance (see Fig. 1).
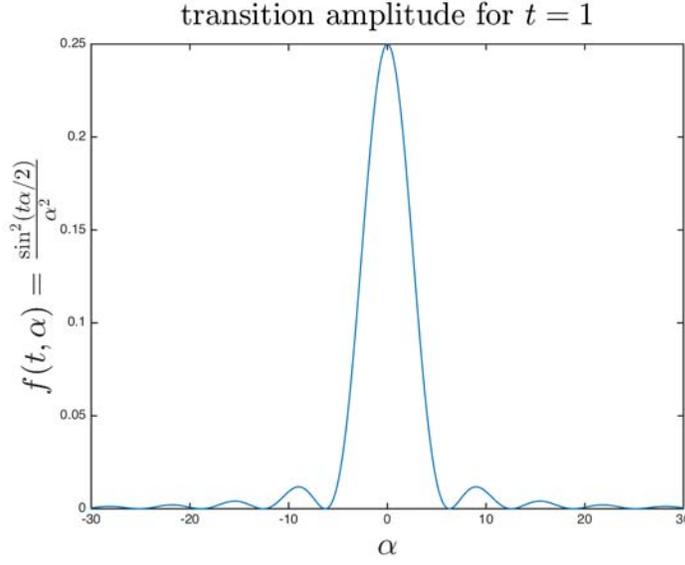

Figure 1: The function $f(\alpha, t)$ from (9), representing the amount of amplitude transfered at a fixed time t as we vary the detuning $\alpha \equiv \omega_{mn} - \omega$.

It has zeroes as $\alpha = \frac{2\pi}{t} n$ for integers $n \neq 0$. Since the closest zeros to the origin are at $\pm 2\pi/t$, we call the region $\alpha \in [-2\pi/t, 2\pi/t]$ the "peak" and the rest of the real line (i.e. $|\alpha| \geq \frac{2\pi}{t}$) the "tails." For $\alpha \to \infty$, $f(t, \alpha) \leq 1/\alpha^2$. Thus, the tail has total area bounded by $2 \int_{2\pi/t}^{\infty} 1/\alpha^2 = O(t)$.

For the peak, as $\alpha \to 0$, $f(t, \alpha) \to t^2/4$. On the other hand, sin is concave, so for $0 \leq \theta \leq \pi/2$ we have $\sin(\theta) \geq \frac{\sin(\pi/2)}{\pi/2}\theta = \frac{2}{\pi}\theta$. Thus for $|\alpha| \leq \frac{\pi}{t}$ we have $f(\alpha, t) \geq \frac{t^2}{\pi}$. While these crude bounds do not determine the precise multiplicative constants, this does show that there is a region of width $\sim 1/t$ and height $\sim t^2$, and so the peak also has area $O(t)$.

We conclude that $\int_{-\infty}^{\infty} d\alpha f(t, \alpha) \sim t$. Dividing by t, we obtain

$$\int_{-\infty}^{\infty} d\alpha \frac{f(t, \alpha)}{t} \sim 1.$$

On the other hand, $\frac{f(t,\alpha)}{t} \to 0$ as $t \to \infty$ for all $\alpha \neq 0$. So as $t \to \infty$ we see that $\frac{f(t,\alpha)}{t}$ is always nonnegative, always has total mass roughly independent of t, but approaches zero for all nonzero α. This means that it approaches a delta function. A more detailed calculation (omitted, but it uses complex analysis) shows that

$$\lim_{t \to \infty} \frac{f(t, \alpha)}{t} = \frac{\pi}{2}\delta(\alpha).$$

The reason to divide by t is that this identifies the rate of transitions per unit time. Define $R_{n\to m} = \frac{P_{n\to m}}{t}$. Then the above arguments imply that

$$R_{n\to m} \approx \frac{\pi}{2}\frac{|V_{mn}|^2}{\hbar^2}\delta(|\omega_{mn}| - \omega) \qquad \text{for large } t. \tag{10}$$

Linewidth In practice the frequency dependence is not a delta function. The term "linewidth" refers to the width of the region of ω that drives a transition; more concretely, FWHM stands for "full-width half-maximum" and denotes the width of the region that achieves $\geq 1/2$ the peak transition rate. The above derivation already suggests some reasons for nonzero linewidth.

1. *Finite lifetime.* If we apply the perturbation for a limited amount of time, or if the state we are driving to/from has finite lifetime, then this will contribute linewidth on the order of $1/t$.

2. *Power broadening.* If $|V_{mn}|$ is large, then we will still see transitions for larger values of $|\alpha|$. For this to prevent us from seeing the precise location of a peak, we need also the phenomenon of *saturation* in which transition rates all look the same above some threshold. (For example, we might observe the fraction of a beam that is absorbed by some sample, and by definition this cannot go above 1.)

There are many other sources of linewidth. In general we can think of both the driving frequency ω and the gap frequency ω_{mn} as being distributions rather than δ functions. The driving frequency might come from a thermal distribution or a laser, both of which output a distribution of frequencies. The linewidth of a laser is much lower but still nonzero. The energy difference $\hbar\omega_{mn}$ seems like a universal constant, can also be replaced by a distribution by phenomena such as Doppler broadening, in which the thermal motion of an atom will redshift or blueshift the incident light. This is just one example of a more general phenomenon in which interactions with other degrees of freedom can add to the linewidth; e.g. consider the hyperfine splitting, which measures the small shifts in an electron's energy from its interaction with the nuclear spin. This can be thought of as adding to linewidth in two different, roughly equivalent, ways. We might think of the nuclear spins as random and thus the interaction adds a random term to the electon's Hamiltonian. Alternatively, we might view the interaction with the nuclear spin as a source of decoherence and thus as contributing to the finite lifetime of the electron's excited state. We will not explore those issues further here.

The other contribution to the rate is the matrix element $|V_{mn}|$. This depends not only on the strength of the perturbation, but also expresses the important point that we only see transitions from $n \to m$ if $V_{mn} \neq 0$. This is called a *selection rule*. In Griffiths it is proved that transitions from electric fields (see the next section) from Hydrogen state $|n, l, m\rangle$ to $|n', l', m'\rangle$ are only possible when $|l - l'| = 1$ and $|m - m'| \leq 1$ (among other restrictions). Technically these constraints hold only for first-order perturbation theory, but still selection rules are important, since they tell us when we need to go to higher-order perturbation theory to see transitions (known as "forbidden transitions"). In those cases transition rates are much lower. One dramatic example is that $2p \to 1s$ transition in hydrogen takes 1.6ns because it occurs at first order while the $2s \to 1s$ transition takes 0.12 seconds. For this reason states such as the 2s states are called "metastable."

We now consider the most important special case, which gets its own top-level section, despite being an example of a periodic perturbation, which itself is an example of first-order perturbation theory.

2 Light and atoms

Light consists of oscillating $\vec{E}$ and $\vec{B}$ fields. The effects of the $\vec{B}$ fields are weaker by a factor $O(v/c) \sim \alpha$, so we will focus on the $\vec{E}$ fields. Let

$$\vec{E}(\vec{r}) = E_0 \hat{z} \cos(\omega t - kx).$$

However, optical wavelengths are $4000\text{-}8000\mathring{A}$, while the Bohr radius is $\approx 0.5\mathring{A}$, so to leading order we can neglect the x dependence. Thus we approximate

$$\delta H(t) = eE_0\, z\, \cos(\omega t). \tag{11}$$

We now can apply the results on transition rates from the last section with $V_{mn} = eE_0 \langle m|z|n \rangle$. (This term is responsible for selection rules and for the role of polarization.) Thus the rate of transitions is

$$R_{n \to m} = \frac{\pi}{2} \frac{e^2 E_0^2}{\hbar^2} |\langle m|z|n \rangle|^2 \delta(|\omega_{mn}| - \omega). \tag{12}$$

We get contributions at $\omega_{mn} = \pm\omega$ corresponding to both absorption and stimulated emission.

Aside: quantizing light What about spontaneous emission? This does not appear in the semiclassical treatment we've described here. Nor do the photons. "Absorption" means jumping from a low-energy state to a higher-energy state, and "stimulated emission" means jumping from high energy to low energy. In the former case, we reason from energy conservation that a photon must have been absorbed, and in the latter, a photon must have been emitted. However, these arguments are rather indirect. A much more direct explanation of what happens to the photon comes from a more fully quantum treatment. This also yields the phenomenon of spontaneous emission. Recall from 8.05 that oscillating electromagnetic fields can be quantized as follows:

$$E_0 = \mathcal{E}_0(\hat{a} + \hat{a}^\dagger) \qquad \mathcal{E}_0 = \sqrt{\frac{2\pi\hbar\omega}{V}} \text{ (Gaussian units)} \qquad = \sqrt{\frac{\hbar\omega}{\epsilon_0 V}} \text{ (SI units)}$$

Using $\delta H = eE_0 z$, we obtain

$$\delta H = eE_0 z \otimes (\hat{a} + \hat{a}^\dagger).$$

If we look at the action of z in the $\{1s, 2p_z\}$ basis, then it has the form $\begin{pmatrix} 0 & \alpha \\ \alpha & 0 \end{pmatrix}$ with $\alpha = \langle 1s|z|2p_z \rangle$.

We then obtain the form of the Hamiltonian examined on pset 3.

This perspective also can be used to give a partial derivation of the Lamb shift, which can be thought of as the interaction of the electron with fluctuations in the electric field of the vacuum. In the vacuum (i.e. ground state of the photon field) we have $\langle E_0 \rangle \sim \langle \hat{a} + \hat{a}^\dagger \rangle = 0$ but $\langle E_0^2 \rangle \sim \langle (\hat{a} + \hat{a}^\dagger)^2 \rangle > 0$. These vacuum fluctuations lead to a small separation in energy between the 2s and 2p levels of hydrogen.

Dipole moment In the Stark effect we looked at the interaction of the hydrogen atom with an electric field. This was a special case of the interaction between a $\vec{E}$ field and the dipole moment of a collection of particles. Here we discuss the more general case.

Suppose that we have charges $q_1, \ldots, q_N$ at positions $\vec{x}^{(1)}, \ldots, \vec{x}^{(N)}$, and we apply an electric field $\vec{E}(\vec{x})$. The energy is determined by the scalar potential $\phi(\vec{x})$ which is related to the electric

field by $\vec{E} = -\vec{\nabla}\phi$. If $\vec{E}(\vec{x}) = \vec{E}$ (i.e. independent of position $\vec{x}$) then one possible solution is $\phi(\vec{x}) = -\vec{x} \cdot \vec{E}$. In this case the Hamiltonian will be

$$H = \sum_{i=1}^{N} q_i \phi\big(\vec{x}^{(i)}\big) = -\sum_{i=1}^{N} q_i \vec{x}^{(i)} \cdot \vec{E} = -\vec{d} \cdot \vec{E}$$

where we have defined the dipole moment $\vec{d} = \sum_{i=1}^{N} q_i \vec{x}^{(i)}$. Our choice of ϕ was not unique, and we could have chosen $\phi(\vec{x}) = C - \vec{x} \cdot \vec{E}$ for any constant C. However, this would only have added an overall constant to the Hamiltonian, which would have no physical effect.

What if the electric field is spatially varying? If this spatial variation is small and we are near the origin, we use the first few terms of the Taylor expansion to approximate the field:

$$\vec{E}(\vec{x}) = \vec{E}(0) + \sum_{i,j=1}^{3} \frac{\partial E_i}{\partial x_j} \hat{e}_i x_j + \dots.$$

This corresponds to a scalar potential of the form

$$\phi = \sum_{i=1}^{3} x_i E_i(0) + \frac{1}{2} \sum_{i,j=1}^{3} x_i x_j \frac{\partial E_i}{\partial x_j} x_i x_j + \dots$$

For the quadratic terms we see that the field couples not to the dipole moment, but to the *quadrupole moment*, defined to be $\sum_{i=1}^{N} q_i \vec{x}^{(i)} \otimes \vec{x}^{(i)}$. This is related to emission lines such as $1s \to 3d$ in which ℓ may change by up to ± 2. Of course higher moments such as octupole moments can be also be considered. We will not explore these topics further in 8.06.

2.1 Incoherent light

While we have so far discussed monochromatic light with a definite polarization, it is easier to produce light with a wide range of frequencies and with random polarization. To analyze the rate of transitions this causes we will average (12) over frequencies and polarizations.

Begin with polarizations. Instead of the field being $E_0 \hat{z}$, let the electric field be $E_0 \hat{P}$ for some random unit vector $\hat{P}$. We then replace V with $-E_0 \hat{P} \cdot \vec{d}$. The only operator here is the dipole moment $\vec{d} = (d_1, d_2, d_3)$, so the matrix elements of V are given by

$$V_{mn} = -E_0 \hat{P} \cdot \vec{d}_{mn} = -E_0 \sum_{i=1}^{3} P_i \langle m|d_i|n\rangle.$$

Since the transition rate depends on $|V_{mn}|^2$, we will average this quantity over the choice of polar-

ization. Denote the average over all unit vectors $\hat{P}$ by $\langle \cdot \rangle_{\hat{P}}$.

$$\left\langle |V_{mn}|^2 \right\rangle_{\hat{P}} = E_0^2 \left\langle |\hat{P} \cdot \vec{d}_{mn}|^2 \right\rangle_{\hat{P}}$$

$$= E_0^2 \sum_{i,j=1}^{3} \left\langle \langle m|P_i d_i|n \rangle \langle n|P_j d_j|m \rangle \right\rangle_{\hat{P}}$$

$$= E_0^2 \sum_{i,j=1}^{3} \left\langle P_i P_j \right\rangle_{\hat{P}} \langle m|d_i|n \rangle \langle n|d_j|m \rangle$$

$$= E_0^2 \sum_{i,j=1}^{3} \frac{\delta_{ij}}{3} \langle m|d_i|n \rangle \langle n|d_j|m \rangle \qquad \text{explained below}$$

$$= \frac{E_0^2}{3} \sum_{i} |\langle m|d_i|n \rangle|^2$$

$$\equiv \frac{E_0^2}{3} |\vec{d}_{mn}|^2$$

How did we calculate $\langle P_i P_j \rangle_{\hat{P}}$? This can be done by explicit calculation, but it is easier to use symmetry. First, observe that the uniform distribution over unit vectors is invariant under reflection. Thus, if $i \neq j$, then $\langle P_i P_j \rangle = \langle (-P_i) P_j \rangle = 0$. On the other hand rotation symmetry means that $\langle P_i^2 \rangle$ should be independent of i. Since $P_1^2 + P_2^3 + P_3^2 = 1$, we also have $\langle P_1^2 + P_2^3 + P_3^2 \rangle = 1$ and thus $\langle P_i^2 \rangle = 1/3$. Putting this together we obtain

$$\langle P_i P_j \rangle_{\hat{P}} = \frac{\delta_{ij}}{3}. \tag{13}$$

Next, we would like to average over different frequencies. The energy density of an electric field is $U = \frac{E_0^2}{8\pi}$ (using Gaussian units). Define $U(\omega)$ to be the energy density at frequency ω, so that $U = \int U(\omega)\,d\omega$. If we consider light with this power spectrum, then we should integrate the rate times this distribution over $U(\omega)$ to obtain

$$R_{n \to m} = \int d\omega\, U(\omega) \frac{4\pi^2}{3\hbar^2} |\vec{d}_{mn}|^2 \delta(\omega - |\omega_{mn}|)$$

$$= \boxed{\frac{4\pi^2}{3\hbar^2} |\vec{d}_{mn}|^2 U(|\omega_{mn}|)}$$

This last expression is known as *Fermi's Golden Rule*. (It was discovered by Dirac, but Fermi called it "Goldren Rule #2".)

2.2 Spontaneous emission

The modern description of spontaneous emission requires QED, but the first derivation of it predates even modern quantum mechanics! In a simple and elegant argument, Einstein:

(a) derived an exact relation between rates of spontaneous emission, stimulated emission and absorption; and

(b) proposed the phenomenon of stimulated emission, which was not observed until 1960.

He did this in 1917, more than a decade before even the Schrödinger equation!

Here we will reproduce that argument. It assumes a collection of atoms that can be in either state a or state b. Suppose that there are N_a atoms in state a and N_b atoms in state b, and that the states have energies E_a, E_b with $E_b > E_a$. Define $\omega_{ba} = \frac{E_b - E_a}{\hbar}$ and $\beta = 1/k_B T$. Assume further that the atoms are in contact with a bath of photons and that the entire system is in thermal equilibrium with temperature T. From this we can deduce three facts:

Fact 1. Equilibrium means no change: $\dot{N}_a = \dot{N}_b = 0$.

Fact 2. At thermal equilibrium we have $\frac{N_b}{N_a} = \frac{e^{-\beta E_b}}{e^{-\beta E_a}} = e^{-\beta \hbar \omega_{ba}}$.

Fact 3. At thermal equilibrium the black-body radiation spectrum is

$$U(\omega) = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\beta \hbar \omega} - 1}. \tag{14}$$

We would like to understand the following processes:

Process	Explanation	Rate
Absorption	A photon of frequency ω_{ba} is absorbed and an atom changes from state a to state b.	$B_{ab} N_a U(\omega_{ba})$
Spontaneous emission	A photon of frequency ω_{ba} is emitted and an atom changes from state b to state a.	$A N_b$
Stimulated emission	A photon of frequency ω_{ba} is emitted and an atom changes from state b to state a.	$B_{ba} N_b U(\omega_{ba})$

These processes depend on the *Einstein coefficients* A, B_{ab} and B_{ba} for spontaneous emission, absorption and stimulated emission respectively. They also depend on the populations of atoms and/or photons that they involve; e.g. absorption requires an atom in state a and a photon of frequency ω_{ba}, so its rate is proportional to $N_a U(\omega_{ba})$. Here it is safe to posit the existence of stimulated emission because we have not assumed that B_{ba} is nonzero.

Having set up the problem, we are now almost done! Adding these processes up, we get

$$\dot{N}_b = -N_b A - N_b B_{ba} U(\omega_{ba}) + N_a B_{ab} U(\omega_{ba}). \tag{15}$$

From Fact 1, $\dot{N}_b = 0$ and so we can rearrange (15) to obtain

$$U(\omega_{ba}) = \frac{A}{\frac{N_a}{N_b} B_{ab} - B_{ba}} \overset{\text{Fact 2}}{=} \frac{A}{e^{\beta \hbar \omega_{ba}} B_{ab} - B_{ba}} \overset{\text{Fact 3}}{=} \frac{\hbar \omega^3}{\pi^2 c^3} \frac{1}{e^{\beta \hbar \omega_{ba}} - 1}$$

Since this relation should hold for all values of β, we can equate coefficients and find

$$B_{ab} = B_{ba} \tag{16a}$$

$$A = \frac{\hbar \omega_{ba}^3}{\pi^2 c^3} B_{ab} \tag{16b}$$

We see that these three processes are tightly related! All from a simple thought experiment, and not even the one that Einstein is most famous for.

Today we can understand this as the fact that the electric field enters into the Hamiltonian as a Hermitian operator proportional to $\hat{a} + \hat{a}^\dagger$, and so the photon-destroying processes containing $\hat{a}$

are inevitably accompanied by photon-creating processes containing $\hat{a}^\dagger$. Additionally the relation between spontaneous and stimulated emission can be seen in the fact that both involve an $a^\dagger$ operator acting on the photon field. If there are no photons, then the field is in state $|0\rangle$ and we get the term $a^\dagger|0\rangle = |1\rangle$, corresponding to spontaneous emission. If there are already n photon in the mode, then we get the term $a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$. Since the probabilities are the square of the amplitudes, this means that we see photons emitted at $n+1$ times the spontaneous emission rate. In Einstein's terminology, the n here is from stimulated emission and the $+1$ is from the spontaneous emission which always occurs independent of the number of photons present.

Returning to (16), we plug in Fermi's Golden Rule and obtain the rates

$$B_{ab} = B_{ba} = \frac{4\pi^2}{3\hbar^2}|\vec{d}_{ab}|^2 \quad \text{and}$$

$$A = \frac{4\omega_{ba}^3}{3\hbar c^3}|\vec{d}_{ab}|^2.$$

2.3 The photoelectric effect

So far we have considered transitions between individual pairs of states. Ionization (aka the photoelectric effect) involves a transition from a bound state of some atom to one in which the electron is in a free state. This presents a few new challenges. First, we are used to treating unbound states as unnormalized, e.g. $\psi(\vec{x}) = e^{i\vec{k}\cdot\vec{x}}$. Second, to calculate the ionization rate, we need to sum over final states, since the quantity of physical interest is the total rate of electrons being dislodged from an atom, and not the rate at which they transition to any specific final state. (A more refined analysis might look at the angular distribution of the final state.)

Suppose our initial state is our old friend, the ground state of the hydrogen atom, and the final state is a plane wave. If the final state is unnormalized, then matrix elements such as $\langle\psi_{\text{final}}|V|\psi_{\text{initial}}\rangle$ become less meaningful. One way to fix this is to put the system in a box of size $L \times L \times L$ with $L \gg a_0 \equiv \frac{\hbar^2}{me^2}$ and to impose periodic boundary conditions. The resulting plane-wave states are now

$$\psi_{\vec{k}}(\vec{x}) = \frac{\exp(i\vec{k}\cdot\vec{x})}{L^{3/2}},$$

where $|k\rangle = \frac{2\pi}{L}\vec{n}$ and $\vec{n} = (n_1, n_2, n_3)$ is a vector of integers. (We use $\vec{k}$ instead of $\vec{p} = \hbar\vec{k}$ to keep the notation more compact.) We will assume that $L \gg a_0$ and also that the final energy of the electron is $\gg 13.6$ eV. This means that we can approximate the final state as a free electron and can ignore the interaction with the proton.

Apply an oscillating electric field to obtain the time-dependent potential

$$\delta H = eE_0 x_3 \cos(\omega t) \equiv V \cos(\omega t).$$

The rate of the transitions will be governed by the matrix element

$$\langle \vec{k}|V|1,0,0\rangle = \frac{eE_0}{\sqrt{\pi a_0^3 L^3}} \underbrace{\int d^3x\, x_3 \exp\left(-\frac{r}{a_0} - i\vec{k}\cdot\vec{x}\right)}_{A} \equiv \frac{eE_0}{\sqrt{\pi a_0^3 L^3}}A.$$

We have defined A to equal the difficult-to-evaluate integral. The factor of x_3 can be removed by

writing $A = i\frac{\partial}{\partial k_3}B$, where

$$B = \int d^3x \, \exp\left(-\frac{r}{a_0} - i\vec{k}\cdot\vec{x}\right)$$

$$= 2\pi \int_0^\infty r^2 \int_{-1}^1 d\mu \exp\left(-\frac{r}{a_0} - ikr\mu\right) \qquad \text{defining } \mu = \frac{\vec{k}\cdot\vec{x}}{kr}, k = |\vec{k}|, r = |\vec{x}|$$

$$= 4\pi \int_{-1}^1 d\mu \frac{1}{\left(\frac{1}{a_0} + ik\mu\right)^3} \qquad \text{using } \int_0^\infty r^n e^{-\alpha r} = \frac{n!}{\alpha^{n+1}}$$

$$= \frac{4\pi i}{k^3} \int_{-1}^1 d\mu \frac{1}{\left(\mu + \frac{i}{ka_0}\right)^3}$$

$$= -\frac{2\pi i}{k^3}\left[\frac{1}{\left(1 + \frac{i}{ka_0}\right)^2} - \frac{1}{\left(1 - \frac{i}{ka_0}\right)^2}\right] = \frac{2\pi i}{k^3}\left[\frac{\left(1 + \frac{1}{ika_0}\right)^2 - \left(1 - \frac{1}{ika_0}\right)^2}{\left(1 + \frac{1}{k^2 a_0^2}\right)^2}\right]$$

$$= \frac{8\pi i}{k^3}\frac{\frac{1}{ika_0}}{\left(1 + \frac{1}{k^2 a_0^2}\right)^2} = \frac{8\pi}{k^4 a_0 \left(1 + \frac{1}{k^2 a_0^2}\right)^2} = \boxed{\frac{8\pi}{a_0 \left(k^2 + a_0^{-2}\right)^2}}$$

To compute A, we use the fact that $\frac{\partial}{\partial k_3}k^2 = 2k_3$. Thus

$$A = i\frac{\partial}{\partial k_3}B = \frac{-32\pi i k_3}{a_0 \left(k^2 + a_0^{-2}\right)^3}.$$

We can simplify this expression using our assumption that the photon energy (and therefore also the final state energy) is much larger than the binding energy. The final energy is $\frac{\hbar^2 k^2}{2m}$ and the binding energy is $\frac{\hbar^2}{2ma_0^2}$. Thus

$$\frac{\hbar^2 k^2}{2m} \gg \frac{\hbar^2}{2ma_0^2} \qquad \Longrightarrow \qquad ka_0 \gg 1.$$

We can use this to simplify A by dropping the a_0^{-2} term in the dominator:

$$A \approx \frac{-32\pi i k_3}{a_0 k^6} = \frac{-32\pi i \cos(\theta)}{a_0 k^5},$$

where θ is the angle between $\vec{k}$ and the z-axis.

We can now compute the squared matrix element to be (canceling a factor of π from numerator and denominator)

$$|\langle\vec{k}|V|1,0,0\rangle|^2 = \frac{e^2 E_0^2}{a_0^3 L^3}\frac{1024\pi \cos^2(\theta)}{a_0^2 k^{10}}.$$

To compute the average rate over some time t, we will multiply this by $\frac{1}{\hbar^2}\frac{f(t,\alpha)}{t}$, where $f(t,\alpha) \equiv \frac{\sin^2(\alpha t/2)}{\alpha^2}$ and $\hbar\alpha$ is the difference in energy between the initial and final states. If we neglect the energy of the initial state, we obtain

$$R_{1,0,0\to\vec{k}} = \frac{1024\pi e^2 E_0^2 \cos^2(\theta)}{\hbar^2 a_0^5 k^{10} L^3}\frac{f\left(t, \frac{\hbar k^2}{2m} - \omega\right)}{t}.$$

13

We can simplify this a bit by averaging over all the polarizations of the light. (In fact, the angular dependance of the free electron can often carry useful information, but here it will help simplify some calculations.) The average of $\cos^2(\theta)$ over the sphere is $1/3$ (by the same arguments we used in the derivation of Fermi's golden rule), so we obtain

$$\langle R_{1,0,0 \to \vec{k}} \rangle = \frac{1024\pi e^2 E_0^2}{3\hbar^2 a_0^5 k^{10} L^3} \frac{f\left(t, \frac{\hbar k^2}{2m} - \omega\right)}{t}.$$

Let's pause for a minute to look at what we've derived. One strange feature is the $1/L^3$ term, because the rate of ionization should not depend on how much empty space surrounds the atom. Another strange thing appears to happen when we take t large, so that $f(t,\alpha)/t$ will approach $\frac{\pi}{2}\delta(\alpha)$. This would cause the transition rates to be nonzero only when $2m\omega/\hbar$ exactly equals k^2 for some valid vector $\vec{k}$ (i.e. of the form $\frac{2\pi}{L}\vec{n}$). We do not generally expect physical systems to have such sensitive dependence on their parameters.

As often happens when two things look wrong, these difficulties can be made to "cancel each other out." Let us take t to be large but finite. It will turn out that t needs to be large only relative to $\frac{\hbar^2}{L^2 m}$, which is not very demanding when L is large. In this case, we can approximate $f(t,\alpha)$ with a step function:

$$f(t,\alpha) \approx \tilde{f}(t,\alpha) \equiv \begin{cases} \frac{\pi}{2}t^2 & \text{if } 0 \leq \alpha \leq \frac{1}{t} \\ 0 & \text{otherwise} \end{cases}$$

In what sense is this a good approximation? We argue that for large t, $\tilde{f}(t,\alpha)/t \approx \frac{\pi}{2}\delta(\alpha)$, just like $f(t,\alpha)/t$. Suppose that $g(\alpha)$ is a function satisfying $|g'(\alpha)| \leq C$ for all α. Then

$$\left| \int_{-\infty}^{\infty} d\alpha \left(\frac{\tilde{f}(t,\alpha)}{\frac{\pi}{2}t} - \delta(\alpha) \right) g(\alpha) \right| = \left| \left(t \int_0^{1/t} d\alpha\, g(\alpha) \right) - g(0) \right|$$

$$= \left| t \int_0^{1/t} d\alpha\, (g(\alpha) - g(0)) \right|$$

$$= \left| t \int_0^{1/t} d\alpha \int_0^{\alpha} d\beta\, g'(\beta) \right|$$

$$\leq t \int_0^{1/t} d\alpha \int_0^{\alpha} d\beta\, |g'(\beta)| \qquad \text{triangle inequality}$$

$$\leq t \frac{1}{2t^2} C = \frac{C}{2t}$$

This tends to 0 as $t \to \infty$. (This is an example of a more general principle that the "shape" of a δ function doesn't matter. For example, the limit of a Gaussian distribution with $\sigma^2 \to 0$ would also work.)

Now using $\tilde{f}(t,\alpha)$, we get a nonzero contribution for $\vec{k}$ satisfying

$$0 \leq \frac{\hbar}{k^2} 2m - \omega \leq \frac{1}{t}$$

$$\Leftrightarrow \quad \frac{2m\omega}{\hbar} \leq k^2 \leq \frac{2m\omega}{\hbar}\left(1 + \frac{1}{t\omega}\right)$$

$$\Leftrightarrow \quad \sqrt{\frac{2m\omega}{\hbar}} \leq k \leq \sqrt{\frac{2m\omega}{\hbar}\left(1 + \frac{1}{t\omega}\right)} \approx \sqrt{\frac{2m\omega}{\hbar}}\left(+\frac{1}{2t\omega}\right) \tag{17}$$

14

How many $\vec{k}$ satisfy (17)? Valid $\vec{k}$ live on a cubic lattice with spacing $2\pi/L$, and thus have density $(L/2\pi)^3$. Thus we can estimate the number of $\vec{k}$ satisfying (17) by $(L/2\pi)^3$ times the volume of k-space satisfying (17). This in turn corresponds to a spherical shell of inner radius $\sqrt{\frac{2m\omega}{\hbar}}$ and thickness $\sqrt{\frac{2m\omega}{\hbar}}\frac{1}{2t\omega}$. Thus we have

$$\# \text{ valid } \vec{k} = \left(\frac{L}{2\pi}\right)^3 4\pi \left(\frac{2m\omega}{\hbar}\right)^{3/2}\frac{1}{2t\omega} = \frac{L^3 m}{2\pi^2 \hbar t}\sqrt{\frac{2m\omega}{\hbar}} = \frac{L^3 mk}{2\pi^2 \hbar t}.$$

In the last step we use the fact that spherical shell is thin to approximate $k \approx \sqrt{\frac{2m\omega}{\hbar}}$. Thus, when we sum $\tilde{f}(t,\alpha)/t$ over $\vec{k}$ we obtain

$$\sum_{\vec{k}} \frac{\tilde{f}(t,\alpha)}{t} = \frac{\pi}{2}t \cdot \# \text{ valid } \vec{k} = \frac{L^3 mk}{4\pi\hbar}.$$

We have obtained our factor of L^3 that removes the unphysical dependence on the boundary conditions. Putting everything together we get

$$R_{1,0,0 \to \text{all } \vec{k}} = \frac{256 me^2 E_0^2}{3\hbar^3 a_0^5 k^9}.$$

3 Adiabatic evolution

3.1 The adiabatic approximation

We now turn to a different kind of approximation, in which we consider slowly varying Hamiltonians. We will consider a time-dependent Hamiltonian $H(t)$. Let $|\psi_n(t)\rangle$ and $E_n(t)$ be the "instantaneous" eigenbases and eigenenergies, defined by

$$H(t)|\psi_n(t)\rangle = E_n(t)|\psi_n(t)\rangle \qquad E_1(t) \le E_2(t) \le \ldots \tag{18}$$

We also define $|\Psi(t)\rangle$ to be the solution of Schrödinger's equation: i.e.

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = H(t)|\Psi(t)\rangle. \tag{19}$$

Beware that (18) and (19) are *not* the same. You might think of (18) as a naive attempt to solve (19). If the system starts in $|\psi_n(0)\rangle$ at time 0, there is of course no reason in general to expect that $|\psi_n(t)\rangle$ will be the correct solution for later t. And yet, the adiabatic theorem states that in some cases this is exactly what happens.

Theorem 1 (Adiabatic theorem). *Suppose at $t = 0$, $|\Psi(0)\rangle = |\psi_n(0)\rangle$ for some n. Then if H is changed slowly for $0 \le t \le T$, then at time T we will have $|\Psi(T)\rangle \approx |\psi_n(T)\rangle$.*

This theorem is stated in somewhat vague terms, e.g. what does "changed slowly" mean? $\dot{H}$ should be small, but relative to what? One clue is the reference to the nth eigenstate $|\psi_n(t)\rangle$. This is only well defined if $E_n(t)$ is unique, so clearly the theorem fails in the case of degenerate eigenvalues. And since the theorem should not behave discontinuously with respect to $H(t)$, it should also fail for "nearly" degenerate eigenvalues. This gives us another energy scale to compare with $\dot{H}$ (which has units of energy/time, or energy squared once we multiply by $\hbar$). We will see later the sense in which this can be shown to be the right comparison.

Example. Suppose we have a spin-1/2 particle in a magnetic field $\vec{B}(t)$. Then the Hamiltonian is $H(t) = g_e \mu_B \frac{\vec{S}}{\hbar} \cdot \vec{B}(t)$. The adiabatic theorem says that if we start with the spin and $\vec{B}$ both pointing in the $+\hat{z}$ direction and gradually rotate $\vec{B}$ to point in the $\hat{x}$ direction, then the spin will follow the magnetic field and also point in the $\hat{x}$ direction. Given that the Schrödinger equation prescribes instead that the spin precess around the magnetic field, this behavior appears at first somewhat strange.

Derivation We will not rigorously prove the adiabatic theorem, but will describe most of the derivation. Begin by writing

$$|\Psi(t)\rangle = \sum_n c_n(t)|\psi_n(t)\rangle.$$

Taking derivatives of both sides we obtain

$$i\hbar\frac{d}{dt}|\Psi(t)\rangle = i\hbar\sum_n \dot{c}_n(t)|\psi_n(t)\rangle + c_n(t)|\dot{\psi}_n(t)\rangle = \sum_n c_n(t)E_n(t)|\psi_n(t)\rangle.$$

Multiply both sides by $\langle\psi_k(t)|$ and we obtain

$$i\hbar\dot{c}_k = E_k c_k - i\hbar\sum_n \langle\psi_k|\dot{\psi}_n\rangle c_n. \tag{20}$$

Now we need a way to evaluate $\langle\psi_k|\dot{\psi}_n\rangle$ in terms of more familiar quantities.

$$\text{Start with} \qquad H|\psi_n\rangle = E_n|\psi_n\rangle$$
$$\text{Apply } \frac{d}{dt} \qquad \dot{H}|\psi_n\rangle + H|\dot{\psi}_n\rangle = \dot{E}_n|\psi_n\rangle + E_n|\dot{\psi}_n\rangle$$
$$\text{Apply } \langle\psi_k| \qquad \langle\psi_k|\dot{H}|\psi_n\rangle + E_k\langle\psi_k|\dot{\psi}_n\rangle = \dot{E}_n\delta_{kn} + E_n\langle\psi_k|\dot{\psi}_n\rangle$$

This equation has two interesting cases: $k = n$ and $k \neq n$. The former will not be helpful in estimating $\langle\psi_k|\dot{\psi}_n\rangle$, but does give us a useful result, called the Hellmann-Feynman theorem.

$$k = n \qquad \dot{E}_n = \langle\psi_n|\dot{H}|\psi_n\rangle$$
$$k \neq n \qquad \langle\psi_k|\dot{\psi}_n\rangle = \frac{\langle\psi_k|\dot{H}|\psi_n\rangle}{E_n - E_k} \equiv \frac{\dot{H}_{kn}}{E_n - E_k}$$

In the last step, we used $\dot{H}_{kn}$ to refer to the matrix elements of $\dot{H}$ in the $\{|\psi_n\rangle\}$ basis.

Plugging this into (20) we find

$$i\hbar\dot{c}_k = \underbrace{(E_k - i\hbar\langle\psi_k|\dot{\psi}_k\rangle)c_k}_{\text{adiabatic approximation}} - \underbrace{i\hbar\sum_{n\neq k} \frac{\dot{H}_{kn}}{E_n - E_k}c_n}_{\text{error term}}. \tag{21}$$

If the part of the equation denoted "error term" did not exist, then $|c_k|$ would be independent of time, which would confirm the adiabatic theorem. Furthermore, the error term is suppressed by a factor of $1/\Delta_{nk}$, where $\Delta_{nk} \equiv E_n - E_k$ is the energy gap. So naively it seems that if $\dot{H}$ is small relative to Δ_{nk} then the error term should be small. On the other hand, these two quantities do not even have the same units, so we will have to be careful.

16

Phases Before we analyze the error term, let's look at the phases we get if the error term were not there. i.e. suppose that $i\hbar\dot{c}_k = (E_k - i\hbar\langle\psi_k|\dot{\psi}_k\rangle)c_k$. The solution of this differential equation is

$$c_k(t) = c_k(0)e^{i\theta_k(t)}e^{i\gamma_k(t)} \tag{22a}$$

$$\theta_k(t) \equiv -\frac{1}{\hbar}\int_0^t E_k(t')dt' \qquad \gamma_k(t) \equiv \int_0^t \nu_k(t')dt' \qquad \nu_k(t) \equiv i\langle\psi_k|\dot{\psi}_k\rangle \tag{22b}$$

The $\theta_k(t)$ term is called the "dynamical" phase and corresponds to exactly what you'd expect from a Hamiltonian that's always on; namely the phase of state k rotates at rate $-E_k/\hbar$. The $\gamma_k(t)$ is called the "geometric phase" or "Berry phase" and will be discussed further in the next lecture. At this point, observe only that it is independent of $\hbar$ and that $\nu_k(t)$ can be seen to be real by applying d/dt to the equation $\langle\psi_k|\psi_k\rangle = 1$.

Validity of the adiabatic approximation Let's estimate the magnitude of the error term in a toy model. Suppose that $H(t) = H_0 + \frac{t}{T}V$, where H_0, V are time-independent and T is a constant that sets the timescale on which V is turned on. Then $\dot{H} = V/T$. An important prediction about the adiabatic theorem is that if the more slowly H changes from H_0 to $H_0 + V$, the lower the probability of transition should be; i.e. increasing T should reduce the error term, even if we integrate over time from 0 to T.

Let's see how this works. If the gap is always $\gtrsim \Delta$, then we can upper-bound the transition rate by some matrix element of $\frac{V}{T\Delta}$. This decreases as T and Δ increase, which is good. But if we add up this rate of transitions over time T, then the total transition amplitude can be as large as $\sim V/\Delta$. Thus, going more slowly appears *not* to reduce the total probability of transition!

What went wrong? Well, we assumed that amplitude from state n simply added up in state k. But if the states have different energies, then over time the terms we add will have different phases, and may cancel out. This can be understood in terms of time-dependent perturbation theory. Define $\tilde{c}_k(t) = e^{-i\theta_k(t)}c_k(t)$. Observe that

$$i\hbar\frac{d}{dt}\tilde{c}_k(t) = \hbar\dot{\theta}_k(t)e^{-i\theta_k(t)}c_k(t) + i\hbar e^{-i\theta_k(t)}\dot{c}_k(t)$$

$$= -E_k(t)e^{-i\theta_k(t)}c_k(t) + e^{-i\theta_k(t)}(E_k(t) - \hbar\nu_k(t))c_k(t) - i\hbar\sum_{n\neq k}\frac{\dot{H}_{kn}}{E_n - E_k}e^{-i\theta_k(t)}c_n(t)$$

$$= -\hbar\nu_k(t)\tilde{c}_k(t) - i\hbar\sum_{n\neq k}\frac{\dot{H}_{kn}}{E_n - E_k}e^{i(\theta_n(t)-\theta_k(t))}\tilde{c}_n(t)$$

In the last step we have used $c_n(t) = e^{i\theta_n(t)}\tilde{c}_n(t)$. Let's ignore the $\nu_k(t)$ geometric phase term (since our analysis here is somewhat heuristic). We see that the error term is the same as in (21) but with an extra phase of $e^{i(\theta_n(t)-\theta_k(t))}$. Analyzing this in general is tricky, but let's suppose that the energy levels are roughly constant, so we can replace it with $e^{-i\omega_{nk}t}$, where $\omega_{nk} = (E_n - E_k)/\hbar$. Now when we integrate the contribution of this term from $t = 0$ to $t = T$ we get

$$\int_0^T \frac{\dot{H}_{kn}e^{-i\omega_{nk}t}}{E_n - E_k}dt \sim \frac{V}{T}\frac{e^{-i\omega_{nk}T} - 1}{\hbar\omega_{nk}^2} \sim \frac{V}{T\hbar\omega_{nk}^2} \sim \frac{\hbar V}{\Delta^2 T}$$

Finally we obtain that the probability of transition decreases with T. This can be thought of as a rough justification of the adiabatic theorem, but it of course made many simplifying assumptions and in general it will be only qualitatively correct.

This was focused on a specific transition. In general adiabatic transitions between levels m and n are suppressed if

$$\boxed{\hbar|\dot{H}_{mn}| \ll \Delta^2 = \min_t (E_m(t) - E_n(t))^2}. \tag{23}$$

Landau-Zener transitions One example that can be solved exactly is a two-level system with a linearly changing Hamiltonian. Suppose a spin-1/2 particle experiences a magnetic field resulting in the Hamiltonian

$$H(t) = \Delta\sigma_x + \frac{vt}{T}\sigma_z,$$

for some constants Δ, v, T. The eigenvalues are $\pm\sqrt{\Delta^2 + (vt/T)^2}$. Assuming $v > 0$, then when $t = -\infty$ the top eigenstate is $|-\rangle$ and the bottom eigenstate is $|+\rangle$. When $t = \infty$ these are reversed; $|+\rangle$ is the top eigenstate and $|-\rangle$ is the bottom eigenstate. When $t = 0$, the eigenstates are $\frac{|+\rangle \pm |-\rangle}{\sqrt{2}}$. See diagram on black-board for energy levels.

Suppose that $\Delta = 0$ and we start in the $|-\rangle$ at $t = -\infty$. Then at $t = \infty$ we will still be in the $|-\rangle$ state, with only the phase having changed. But if $\Delta > 0$ and we move slowly enough then the adiabatic approximation says we will remain in the top eigenstate, which for $t = \infty$ will be $|+\rangle$. Thus, the presence of a very small transverse field can completely change the state if we move slowly enough through it.

In this case, the error term in the adiabatic approximation can be calculated rather precisely and is given by the Landau-Zener formula (proof omitted):

$$\Pr[\text{transition}] \approx \exp\left(-\frac{2\pi^2\Delta^2 T}{\hbar v}\right).$$

Observe that it has all the qualitative features that we expect in terms of dependence on Δ, v, T, but that it corresponds to a rate of transitions exponentially smaller than our above estimate from first-order perturbation theory. Note that here "transition" refers to transitions between energy level. Thus starting in $|-\rangle$ and ending in $|+\rangle$ corresponds to "no transition" while ending in $|-\rangle$ would correspond to "transition," since it means starting in the higher energy level and ending in the lower energy level.

3.2 Berry phase

Recall that the adiabatic theorem states that if we start in state $|\psi_n(0)\rangle$ and change the Hamiltonian slowly, then we will end in approximately the state

$$e^{i\theta_n(t)}e^{i\gamma_n(t)}|\psi_n(t)\rangle \tag{24a}$$

$$\theta_n(t) \equiv -\frac{1}{\hbar}\int_0^t E_n(t')dt' \qquad \gamma_n(t) \equiv \int_0^t \nu_n(t')dt' \qquad \nu_n(t) \equiv i\langle\psi_n|\dot{\psi}_n\rangle \tag{24b}$$

The phase $\gamma_n(t)$ is called the geometric phase, or the Berry phase, after Michael Berry's 1984 explanation of it.

Do the phases in the adiabatic approximation matter? This is a somewhat subtle question. Of course an overall phase cannot be observed, but a relative phase can lead to observable interference effects. The phases in (22) depend on the eigenstate label n, and so in principle interference is possible. But solutions to the equation $H(t)|\psi_n(t)\rangle = E_n(t)|\psi_n(t)\rangle$ are not uniquely defined, and we can in general redefine $|\psi_n(t)\rangle$ by multiplying by a phase that can depend on both n and t.

To see how this works, let us consider the example of a spin-1/2 particle in a spatially varying magnetic field. If the particle moves slowly, we can think of the position $\vec{r}(t)$ as a classical variable causing the spin to experience the Hamiltonian $H(\vec{r}(t))$. This suggests that we might write the state as a function of $\vec{r}(t)$, as $|\psi_n(\vec{r}(t))\rangle$ or even $|\psi_n(\vec{r})\rangle$. If the particle's position is a classical function of time, then we need only consider interference between states with the same value of $\vec{r}$, and so we can safely change $|\psi_n(\vec{r})\rangle$ by any phase that is a function of n and $\vec{r}$.

In fact, even if the particle were in a superposition of positions as in the two-slit experiment, then we could still only see interference effects between branches of the wavefunction with the same value of $\vec{r}$. Thus, again we can define an arbitrary $(n, \vec{r})$-dependent phase.

More generally, suppose that H depends on some set of coordinates $\vec{R}(t) = (R_1(t), \ldots, R_N(t))$. The eigenvalue equation (18) becomes $H(\vec{R})|\psi_n(\vec{R})\rangle = E_n(\vec{R})|\psi_n(\vec{R})\rangle$ where we leave the time-dependence of $\vec{R}$ implicit. This allows us to compute even in situations where $\vec{R}$ is in a superposition of coordinates at a given time t.

To express $\gamma_n(t)$ in terms of $|\psi_n(\vec{R})\rangle$, we compute

$$\frac{d}{dt}|\psi_n(\vec{R})\rangle = \sum_{i=1}^{N} \frac{d|\psi_n(\vec{R})\rangle}{dR_i}\frac{dR_i}{dt} = \vec{\nabla}_{\vec{R}}|\psi_n(\vec{R})\rangle \cdot \frac{d\vec{R}}{dt}$$

$$\gamma_n(t) = i\int_0^t \langle\psi_n|\vec{\nabla}_{\vec{R}}|\psi_n\rangle \cdot \frac{d\vec{R}}{dt}dt = \int_{\vec{R}(0)}^{\vec{R}(t)} i\langle\psi_n|\vec{\nabla}_{\vec{R}}|\psi_n\rangle \cdot d\vec{R}$$

The answer is in terms of a line integral, which depends only on the path and *not* on time (unlike the dynamical phase).

How does this change if we reparameterize $|\psi_n(\vec{R})\rangle$? Suppose we replace $|\psi_n(\vec{R})\rangle$ with $|\tilde{\psi}_n(\vec{R})\rangle = e^{-i\beta(\vec{R})}|\psi_n(\vec{R})\rangle$. Then the Berry phase becomes

$$\tilde{\gamma}_n(t) = i\int_{\vec{R}(0)}^{\vec{R}(t)} \langle\tilde{\psi}_n(\vec{R})|\vec{\nabla}_{\vec{R}}|\tilde{\psi}_n(\vec{R})\rangle \cdot d\vec{R} = i\int_{\vec{R}(0)}^{\vec{R}(t)} \langle\psi_n(\vec{R})|e^{i\beta(\vec{R})}\vec{\nabla}_{\vec{R}}e^{-i\beta(\vec{R})}|\psi_n(\vec{R})\rangle \cdot d\vec{R}$$

$$= \gamma_n(t) + \beta(\vec{R}(t)) - \beta(\vec{R}(0))$$

Changing β only changes phases as a function of the endpoints of the path. Thus, we can eliminate the phase for any fixed path with $\vec{R}(t) \neq \vec{R}(0)$, but not simultaneously for all paths. In particular, if a particle takes two different paths to the same point, the difference in their phases cannot be redefined away. More simply, suppose the path is a loop, so that $\vec{R}(0) = \vec{R}(t)$. Then regardless of β we will have $\gamma_n = \tilde{\gamma}_n$. This suggests an important point about the Berry phase, which is that it is uniquely defined on closed paths, but not necessarily open ones.

Suppose that $\vec{R}(t)$ follows a closed curve C. Then we can write

$$\gamma_n[C] = \oint \underbrace{i\langle\psi_n|\vec{\nabla}_{\vec{R}}|\psi_n\rangle}_{\vec{\mathcal{A}}_n(\vec{R})} \cdot d\vec{R} = \oint \vec{\mathcal{A}}_n(\vec{R}) \cdot d\vec{R},$$

where we have defined the *Berry connection* $\vec{\mathcal{A}}_n(\vec{R}) = i\langle\psi_n|\vec{\nabla}_{\vec{R}}|\psi_n\rangle$. Note that it is real for the same reason that $\nu_n(t)$ is real.

In some cases, we can simplify $\gamma_n[C]$. If $N = 1$ then the integral is always zero, since the line integral of a closed curve in 1-d is always zero. In 2-d or 3-d we can use Green's theorem or Stokes's theorem respectively to simplify the computation of $\gamma_n[C]$. Let's focus on 3-d, because it contains 2-d as a special case. Then if S denotes the surface enclosed by curve C, we have

$$\oint_C \vec{\mathcal{A}}_n(\vec{R}) \cdot d\vec{R} = \iint_S (\vec{\nabla}_{\vec{R}} \times \vec{\mathcal{A}}_n) \cdot d\vec{a} \equiv \iint_S \vec{D}_n \cdot d\vec{a}.$$

19

Here we define the *Berry curvature* $\vec{D}_n = \vec{\nabla}_{\vec{R}} \times \vec{\mathcal{A}}_n$ and the infinitesimal unit of area $d\vec{a}$. We can write $\vec{D}_n$ in a more symmetric way as follows:

$$(D_n)_i = i \sum_{j,k} \epsilon_{ijk} \frac{d}{dR_j} \langle \psi_n | \frac{d}{dR_k} | \psi_n \rangle = i \sum_{j,k} \epsilon_{ijk} \left(\frac{d\langle \psi_n |}{dR_j} \frac{d|\psi_n\rangle}{dR_k} + \langle \psi_n | \frac{d}{dR_j} \frac{d}{dR_k} | \psi_n \rangle \right).$$

Because ϵ_{ijk} is antisymmetric in j,k and $\frac{d}{dR_j}\frac{d}{dR_k}$ is symmetric, the second term vanishes and we are left with

$$\vec{D}_n = i(\vec{\nabla}_r \langle \psi_n |) \times (\vec{\nabla}_r | \psi_n \rangle). \tag{25}$$

Example: electron spin in a magnetic field. The Hamiltonian is

$$H = \mu \vec{\sigma} \cdot \vec{B} \qquad \mu = \frac{e\hbar}{mc}.$$

Suppose that $\vec{B} = B\vec{r}$ where B is fixed and we slowly trace out a closed path in the unit sphere with $\vec{r}$. Suppose that we start in the state

$$|\vec{r}; +\rangle = |\vec{r}\rangle = \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi} \sin(\theta/2) \end{pmatrix} \qquad \text{with} \qquad \vec{r} = \begin{pmatrix} \sin(\theta)\cos(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\theta) \end{pmatrix}$$

Then the adiabatic theorem states that we will remain in the state $|\vec{r}\rangle$ at later points, up to an overall phase. To compute the geometric phase observe that

$$\vec{\nabla} = \frac{d}{dr}\hat{r} + \frac{1}{r}\frac{d}{d\theta}\hat{\theta} + \frac{1}{r\sin\theta}\frac{d}{d\phi}\hat{\phi}.$$

Since $\frac{d}{dr}|\vec{r}\rangle = 0$ we have

$$\vec{\nabla}|\vec{r}\rangle = \frac{1}{2r} \underbrace{\begin{pmatrix} -\sin(\theta/2) \\ e^{i\phi}\cos(\theta/2) \end{pmatrix}}_{|-\vec{r}\rangle} \hat{\theta} + \frac{1}{r\sin\theta} \begin{pmatrix} 0 \\ ie^{i\phi}\sin(\theta/2) \end{pmatrix} \hat{\phi}.$$

This first term will not contribute to the Berry connection, and so we obtain

$$\vec{\mathcal{A}}_+(\vec{r}) = i\langle \vec{r} | \vec{\nabla} | \vec{r} \rangle = -\frac{1}{r}\frac{\sin^2(\theta/2)}{\sin(\theta)}\hat{\phi}.$$

Finally the Berry curvature is

$$\vec{D}_+ = \vec{\nabla} \times \vec{\mathcal{A}}_+ = \frac{1}{r\sin\theta}\frac{d}{d\theta}\left(\sin\theta \mathcal{A}_{+,\phi}\right)\hat{r} = -\frac{\hat{r}}{r^2\sin\theta}\frac{d}{d\theta}\sin^2(\theta/2) = -\frac{1}{2r^2}\hat{r}.$$

For this last computation, observe that $\frac{d}{d\theta}\sin^2(\theta/2) = \frac{d}{d\theta}\frac{1-\cos\theta}{2} = \sin\theta$. We can now compute the Berry phase as

$$\gamma_+[C] = \iint_S \vec{D}_+ \cdot \underbrace{d\vec{a}}_{r^2 d\Omega \hat{r}} = -\frac{1}{2}\Omega.$$

20

Here $d\Omega$ is a unit of solid angle, and Ω is the solid angle contained by C.

What if we used a different parameterization for $|\vec{r}\rangle$? An equally valid choice is

$$|\vec{r}\rangle = \begin{pmatrix} e^{-i\phi}\cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}. \tag{26}$$

If we carry through the same computation we find that now

$$\vec{\mathcal{A}}_+ = \frac{1}{r}\frac{\cos^2(\theta/2)}{\sin(\theta)}\hat{\phi} \qquad \text{and} \qquad \vec{\mathcal{D}}_+ = -\frac{1}{2r^2}\hat{r}.$$

We see that the $\frac{d}{d\theta}\sin^2(\theta/2)$ was replaced by a $\frac{d}{d\theta}(-\cos^2(\theta/2))$ which gives the same answer. This is an example of the general principle that the Berry connection is sensitive to our choice of phase convention but the Berry curvature is not. Accordingly the Berry curvature can be observed in experiments.

What if we started instead with the state $|\vec{r}; -\rangle = |-\vec{r}\rangle$? Then a similar calculation would find that

$$\gamma_-[C] = \frac{1}{2}\Omega.$$

Since the two states pick up different phases, this can be seen experimentally if we start in a superposition of $|\vec{r}; +\rangle$ and $|\vec{r}; -\rangle$.

More generally, if we have a spin-s particle, then its z component of angular momentum can be anything in the range $-s \leq m \leq s$ and one can show that

$$\gamma_m[C] = -m\Omega.$$

There is much more that can be said about Berry's phase. An excellent treatment is found in the 1989 book *Geometric phases in physics* by Wilczek and Shapere. There is a classical analogue called Hannay's phase. Berry's phase also has applications to molecular dynamics and to understanding electrical and magnetic properties of Bloch states. We will see Berry's phase again when we discuss the Aharonov-Bohm effect in a few weeks.

3.3 Neutrino oscillations and the MSW effect

In this section we will discuss the application of the adiabatic theorem to a phenomenon involving solar neutrinos. The name *neutrino* means "little neutral one" and neutrinos are spin-1/2, electrically neutral, almost massless and very weakly interacting particles. Neutrinos were first proposed by Pauli in 1930 to explain the apparent violation of energy, momentum and angular momentum conservation in beta decay. (Since beta decay involves the decay of a neutron into a proton, an electron and an electron antineutrino, but only the proton and electron could be readily detected, there was an apparent anomaly.)

Their almost complete lack of interaction with matter (it takes 100 lightyears of lead to absorb 50% of a beam of neutrinos) has made many properties of neutrinos remain mysterious. Corresponding to the charged leptons e^-, e^+ (electron/position), μ^-/μ^+ (muon/antimuon) and τ^-/τ^+ (tau/antitau), neutrinos (aka neutral leptons) also exist in three flavors: ν_e, ν_μ, ν_τ, with antineutrinos denoted $\bar{\nu}_e, \bar{\nu}_\mu, \bar{\nu}_\tau$. Most, but not all, interactions preserve lepton number, defined to be the number of leptons minus the number of antileptons. Indeed, most interactions preserve $\#e^- + \#\nu_e - \#e^+ - \#\bar{\nu}_e$ (electronic number) and similarly for muons and taus. However, these quantities are not conserved by neutrino oscillations. Even the total lepton number is violated by a phenomenon known as the chiral anomaly.

Solar neutrinos Solar neutrinos are produced via the p-p chain reaction, which converts (via a series of reactions)

$$4\,^1H = 4p^+ + 4e^- \mapsto \underbrace{2p^+ + 2n + 2e^-}_{^4He} + 2\nu_e.$$
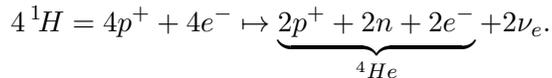
The resulting neutrinos are produced with energies in the range 0.5-20MeV. Almost all of neutrinos produced in the sun are electron neutrinos.

Detection Neutrinos can be detected via inverse beta decay, corresponding to the reaction

$$A + \nu_e \mapsto A' + e^-,$$

where A, A' are different atomic nuclei. For solar neutrinos this will only happen for electron neutrinos because the reaction $A + \nu_\mu \mapsto A' + \mu^-$ will only happen for mu neutrinos carrying at least 108 MeV of kinetic energy. So it is easiest to observe electron neutrinos. However other flavors of neutrinos can also be detected via more complicated processes, such as neutrino-mediated disassociation of deuterium.

Observations of solar neutrinos The first experiment to detect cosmic neutrinos was the 1968 Homestake experiment, led by Ray Davis, which used 100,000 gallons of dry-cleaning fluid (C_2Cl_4) to detect neutrinos via the process $^{37}Cl + \nu_e \mapsto\, ^{37}Ar + e^-$. However, this only found about 1/3 as many neutrinos as standard solar models predicted.

In 2002, the Sudbery Neutrino Observatory (SNO) measured the total neutrino flux and found that once mu- and tau-neutrinos were accounted for, the total number of neutrinos was correct. Thus, somehow electron neutrinos in the sun had become mu and tau neutrinos by the time they reached the Earth.

Neutrino oscillations The first high-confidence observations of neutrino oscillations were by the Super Kamiokande experiment in 1998, which could distinguish electron neutrinos from muon neutrinos. Since neutrinos oscillate, they must have energy, which means they must have mass (if we wish to exclude more speculative theories, such as violations of the principle of relativity). This means that a neutrino, in its rest frame, has a Hamiltonian with eigenstates $|\nu_1\rangle, |\nu_2\rangle, |\nu_3\rangle$ that in general will be different from the flavor eigenstates $|\nu_e\rangle, |\nu_\mu\rangle, |\nu_\tau\rangle$ that participate in weak-interaction processes such as beta decay.

We will treat this in a simplified way by neglecting $|\nu_3\rangle$ and $|\nu_\tau\rangle$. So the Hamiltonian can be modeled as (in the $|\nu_e\rangle, |\nu_\mu\rangle$ basis)

$$H = \begin{pmatrix} E_e & \Delta \\ \Delta & E_\mu \end{pmatrix}, \tag{27}$$

where E_e, E_μ are the energies (possibly equal) of the electron and muon neutrinos and Δ represents a mixing term. Unfortunately, plugging in known parameter estimates for the terms in (27) would predict that roughly a 0.57 fraction of solar neutrinos would end up in the $|\nu_e\rangle$ state, so this still cannot fully explain our observations.

The MSW effect It turns out that this puzzle can be resolved by a clever use of the adiabatic theorem. Electron neutrinos scatter off of electrons and thus the Hamiltonian in (27) should be

modified to add a term proportional to the local density of electrons. Thus after some additional rearranging, we obtain

$$H = E_0 + \begin{pmatrix} -\Delta_0 \cos(2\theta) & \Delta_0 \sin(2\theta) \\ \Delta_0 \sin(2\theta) & \Delta_0 \cos(2\theta) \end{pmatrix} + \begin{pmatrix} CN_e & 0 \\ 0 & 0 \end{pmatrix}, \tag{28}$$

where Δ_0, θ come from (27) ($\theta \approx \pi/6$ is the "mixing angle" that measures how far the flavor states are from being eigenstates), C is a constant and $N_e = N_e(\vec{r})$ is the local electron density. If the neutrino is traveling at speed $\approx c$ in direction $\hat{x}$, then $\vec{r} \approx ct\hat{x}$. Thus we can think of N_e as time-dependent. We then can rewrite H as

$$\text{const} \cdot I + \left(\frac{CN_e(t)}{2} - \Delta_0 \cos(2\theta) \right) \sigma_z + \Delta_0 \sin(2\theta)\sigma_x. \tag{29}$$

This looks like the adiabatic Landau-Zener transition we studied in the last lecture, although here the σ_z term is no longer being swept from $-\infty$ to $+\infty$. Instead, near the center of the sun, $N_e(0)$ is large and the eigenstates are roughly $|\nu_e\rangle, |\nu_\mu\rangle$. For large t, the neutrinos are in vacuum, where their eigenstates are $|\nu_1\rangle, |\nu_2\rangle$.

If the conditions of the adiabatic theorem are met, then neutrinos that start in state $|\nu_e\rangle$ (in the center of the sun) will emerge in state $|\nu_2\rangle$ (at the surface of the sun). They will then remain in this state as they propagate to the Earth. It turns out that this holds for neutrinos of energies $\gtrsim 2MeV$. In this case, the probability of observing the neutrino on Earth in the $|\nu_e\rangle$ state (thinking of neutrino detectors as making measurements in the flavor basis) is $\sin^2(\theta)$, which gives more or less the observed value of 0.31.

3.4 Born-Oppenheimer approximation

Consider a system with N nuclei and n electrons. Write the Hamiltonian as

$$H = -\sum_{j=1}^{N} \frac{\hbar^2}{2M_j} \vec{\nabla}_{\vec{R}_j}^2 + H_{\text{el}}(\mathbf{R}). \tag{30}$$

Here $\mathbf{R} = (\vec{R}_1, \ldots, \vec{R}_N)$ denotes the positions of the N nuclei and $H_{\text{el}}(\mathbf{R})$ includes all the other terms, i.e. kinetic energy of the electrons as well as the potential energy terms which include electron-electron, nuclei-nuclei and electron-nuclei interactions. Let $\mathbf{r}$ denote all of the coordinates of the electrons. While (30) may be too hard to solve exactly, we can use a version of the adiabatic theorem to derive an approximate solution.

We will consider a product ansatz:

$$\Psi(\mathbf{R}, \mathbf{r}) = \gamma(\mathbf{R})\Phi_{\mathbf{R}}(\mathbf{r}), \tag{31}$$

where the many-electron wavefunction is an eigenstate of the reduced Hamiltonian:

$$H_{\text{el}}(\mathbf{R})\Phi_{\mathbf{R}}(\mathbf{r}) = E_{\text{el}}(\mathbf{R})\Phi_{\mathbf{R}}(\mathbf{r}). \tag{32}$$

(Typically this eigenstate wil be simply the ground state.) This is plausible because of the adiabatic theorem. If the nuclei move slowly then as this happens the electrons can rapidly adjust to remain in their ground states. Then once we have solved (32) we might imagine that we can substitute

back to solve for the nuclear eigenstates. We might guess that they are solutions to the following eigenvalue equation

$$\left(-\sum_{j=1}^{N}\frac{\hbar^2}{2M_j}\vec{\nabla}^2_{\vec{R}_j} + E_{\text{el}}(\mathbf{R})\right)\gamma(\mathbf{R}) = E\gamma(\mathbf{R}). \tag{33}$$

However, this is not quite right. If we apply $\vec{\nabla}_{\vec{R}_j}$ to (31) we obtain

$$\vec{\nabla}_{\vec{R}_j}\Psi(\mathbf{R},\mathbf{r}) = (\vec{\nabla}_{\vec{R}_j}\gamma(\mathbf{R}))\Phi_{\mathbf{R}}(\mathbf{r}) + \gamma(\mathbf{R})\vec{\nabla}_{\vec{R}_j}\Phi_{\mathbf{R}}(\mathbf{r}). \tag{34}$$

Using the adiabatic approximation we neglect the overlap of $\vec{\nabla}_{\vec{R}_j}\Psi(\mathbf{R},\mathbf{r})$ with all states to $|\Phi_{\mathbf{R}}\rangle$. Equivalently we can multiply on the left by $\langle\Phi_{\mathbf{R}}|$. This results in

$$\int d^{3n}\mathbf{r}\Phi_{\mathbf{R}}(\mathbf{r})^*\vec{\nabla}_{\vec{R}_j}\Psi(\mathbf{R},\mathbf{r}) = \vec{\nabla}_{\vec{R}_j}\gamma(\mathbf{R}) + \gamma(\mathbf{R})\int d^{3n}\mathbf{r}\Phi_{\mathbf{R}}(\mathbf{r})^*\vec{\nabla}_{\vec{R}_j}\Phi_{\mathbf{R}}(\mathbf{r})$$
$$= (\vec{\nabla}_{\vec{R}_j} - i\vec{\mathcal{A}}_j)\gamma(\mathbf{R}),$$

where $\vec{\mathcal{A}}_j$ is the familiar Berry connection

$$\vec{\mathcal{A}}_j = i\langle\Phi_{\mathbf{R}}|\vec{\nabla}_{\vec{R}_j}|\Phi_{\mathbf{R}}\rangle. \tag{35}$$

We conclude that the effective Hamiltonian actual experienced by the nuclei should be

$$H_{\text{eff}} = \sum_{j=1}^{N}\frac{\hbar^2}{2M_j}(\vec{\nabla}_{\vec{R}_j} - i\vec{\mathcal{A}}_j)^2 + E_{\text{el}}(\mathbf{R}). \tag{36}$$

We will see these $\vec{\mathcal{A}}_j$ terms again when we discuss electomagnetism later in the semester. In systems of nuclei and atoms we need at least three nuclei before the $\vec{\mathcal{A}}_j$ terms can have an effect, for the same reason that we do not see a Berry phase unless we trace out a loop in a parameter space of dimension ≥ 2.

The Born-Oppenheimer applies not just to nuclei and electrons but whenever we can divide a system into fast and slow-moving degrees of freedom; e.g. we can treat a proton as a single particle and ignore (or "integrate out") the motion of the quarks within the proton. This is an important principle that we often take for granted. Some more general versions of Born-Oppenheimer are called "effective field theory" or the renormalization group.

4 Scattering

4.1 Preliminaries

One of the most important types of experiments in quantum mechanics is *scattering*. A beam of particles is sent into a potential and scatters off it in various directions. The angular distribution of scattered particles is then measured. In 8.04 we studied scattering in 1-d, and here we will study scattering in 3-d. This is an enormous field, and we will barely scratch the surface of it. In particular, we will focus on the following special case:

- *Elastic scattering.* The outgoing particle has the same energy as the incoming particle. This means we can model the particles being scattered off semi-classically, as a static potential $V(\vec{r})$. The other types of scattering are inelastic scattering, which can involve transformation of the particles involved or creation of new particles, and absorption, in which there is no outgoing particle.

- *Non-relativistic scattering.* This is by contrast with modern accelerators such as the LHC. However, non-relativistic scattering is still relevant to many cutting-edge experiments, such as modern search for cosmic dark matter (which is believed to be traveling at non-relativistic speeds).

Even this special case can teach us a lot of interesting physics. For example, Rutherford scattering[1] showed that atoms have nuclei, thereby refuting the earlier "plum pudding" model of atoms. This led to a model of atoms in which electrons orbit nuclei like planets, and resolving the problems of this model in turn was one of the early successes of quantum mechanics.

Scattering cross section: In scattering problems it is important to think about which physical quantities can be observed. The incoming particles have a flux that is measured in terms of number of particles per unit area per unit time, i.e. $\frac{d^2 N_{\text{in}}}{dA\,dt}$. If we just count the total number of scattered particles, then this is measured in terms of particles per time: $\frac{dN_{\text{scat}}}{dt}$. The ratio of these quantities has units of area and is called the *scattering cross section*:

$$\frac{\frac{dN_{\text{scat}}}{dt}}{\frac{d^2 N_{\text{in}}}{dA\,dt}} = \sigma. \tag{37}$$

To get a sense of why these are the right units, consider scattering of classical particles off of a classical hard sphere of radius a. If a particle hits the sphere it will scatter, and if it does not hit the sphere it will not scatter. Assume that the beam of particles is much wider than the target, i.e. each particle has trajectory $\vec{r} = (x_0, y_0, z_0 + vt)$ with $\sqrt{x_0^2 + y_0^2}$ given by a distribution with standard deviation that is $\gg a$. The particles that scatter will be the ones with $\sqrt{x_0^2 + y_0^2} \leq a$ which corresponds to a region with area πa^2, which is precisely the cross-sectional area of the sphere. Since we have $\frac{dN_{\text{scat}}}{dt} = \frac{d^2 N_{\text{in}}}{dA\,dt} \pi a^2$, it follows that $\sigma = \pi a^2$. This simple example is good to keep in mind to have intuition about the meaning of scattering cross sections.

Differential cross-section: We can get more information out of an experiment by measuring the angular dependence of the scattered particles. The number of scattered particles can then be measured in terms of a rate *per solid angle*, i.e. $\frac{d^2 N_{\text{scat}}}{d\Omega\,dt}$. The resulting *differential cross-section* $\frac{d\sigma}{d\Omega}$ is defined to be

$$\frac{d\sigma}{d\Omega}(\theta, \phi) \equiv \frac{\frac{d^2 N_{\text{scat}}}{d\Omega\,dt}}{\frac{d^2 N_{\text{in}}}{dA\,dt}} \tag{38}$$

Here the spherical coordinates (θ, ϕ) denote the direction of the outgoing particles. It is conventional to define the axes so that the incoming particles have momentum in the $\hat{z}$ direction, so θ is the angle between the scattered particle and the incoming beam (i.e. $\theta = 0$ means no change in direction and $\theta = \pi$ means backwards scattering) while ϕ is the azimuthal angle. Integrating over all angles gives us the full cross-section, i.e.

$$\sigma = \int d\Omega \frac{d\sigma}{d\Omega}. \tag{39}$$

Quantum mechanical scattering: Assume that the incoming particle states are wavepackets that are large relative to the target. This allows us to approximate the incoming particles as plane wave, i.e.

$$\psi_{\text{in}} \propto e^{ikz - \frac{iEt}{\hbar}}, \tag{40}$$

---

[1]Rutherford scattering is named after Ernest Rutherford for his 1911 explanation of the 1909 experiment which was carried out by Geiger and Marsden.

where $E = \frac{\hbar^2 k^2}{2m}$. Here we need to assume that the potential $V(\vec{r}) \to 0$ as $r \to \infty$ so that plane waves are solutions to the Schrödinger equation for large r. For the scattered wave, we should seek solutions satisfying

$$-\frac{\hbar^2}{2m}\vec{\nabla}^2\psi_{\text{scat}} = E\psi_{\text{scat}} \qquad\qquad \text{as } r \to \infty \qquad (41)$$

$$-\left(\frac{1}{r}\frac{d^2}{dr^2}r + \frac{1}{r^2}\hat{L}^2\right)\psi_{\text{scat}} = k^2\psi_{\text{scat}} \qquad\qquad \text{in spherical coordinates} \qquad (42)$$

A general solution can be written as a superposition of separable solutions. Separable solutions to (42) can in turn be written as

$$r\psi(r,\theta,\phi) = u(r)f(\theta,\phi), \qquad (43)$$

in terms of some functions $u(r), f(\theta,\phi)$. In terms of these (42) becomes

$$u''f + \underbrace{\frac{1}{r^2}u\hat{L}^2 f}_{\to 0 \text{ as } r \to \infty} + k^2 uf = 0. \qquad (44)$$

Thus, for large r, we can cancel the f from each side and simply have $u'' = -k^2 u$, which has solutions $e^{\pm ikr}$. The e^{ikr} solution corresponds to outgoing waves and the e^{-ikr} solution to incoming waves. A scattered wave should be entirely outgoing, and so we obtain

$$\psi_{\text{scat}} \overset{r\to\infty}{=} \frac{f(\theta,\phi)}{r}e^{ikr-\frac{iEt}{\hbar}} \qquad (45)$$

or more precisely

$$\psi_{\text{scat}} = \frac{f(\theta,\phi)}{r}e^{ikr-\frac{iEt}{\hbar}} + O\left(\frac{1}{r^2}\right). \qquad (46)$$

Because the scattering is elastic, the k and E here are the same as for the incoming wave.

Time-independent formulation: As with 1-d scattering problems, the true scattering process is of course time-dependent, but the quantities of interest (transmission/reflection in 1-d, differential cross section in 3-d) can be extracted by solving the time-independent Schrödinger equation with suitable boundary conditions. In the true process, the incoming wave should really be a wavepacket with well-defined momentum $\approx (0,0,k)$ and therefore delocalized position. The outgoing wave will be a combination of an un-scattered part, which looks like the original wave packet continuing forward in the $\hat{z}$ direction, and a scattered part, which is a spherical outgoing wavepacket with a $f(\theta,\phi)$ angular dependence. However, we can treat the incoming wave instead as the static plane wave e^{ikz} and the scattered wave instead as the static outgoing wave $\frac{f(\theta,\phi)}{r}e^{ikr}$. (Both of these are when $r \to \infty$.) Thus we can formulate the entire scattering problem as a time-independent boundary-value problem. The high-level strategy is then to solve the Schrödinger equation subject to the boundary conditions

$$\psi(\vec{r}) \overset{r\to\infty}{=} e^{ikz} + \frac{f(\theta,\phi)}{r}e^{ikr}. \qquad (47)$$

This is analogous to what we did in 1-D scattering, where the boundary conditions were that $\psi(x)$ should approach $e^{ikx} + Re^{-ikx}$ for $x \to -\infty$ and should approach Te^{ikx} for $x \to \infty$. As in the 1-D case, we have to remember that this equation is an approximation for a time-dependent problem. As a result when calculating observable quantities we have to remember not to include interference terms between the incoming and reflected waves, since these never both exist at the same point in time.

Relation to observables. In the 1-D case, the probabilities of transmission and reflection are $|T|^2$ and $|R|^2$ respectively. In the 3-d case, the observable quantities are the differential cross sections $\frac{d\sigma}{d\Omega}$. To compute these, first evaluate the incoming flux

$$\vec{S}_{\text{in}} = \frac{\hbar}{m} \operatorname{Im} \psi_{\text{in}}^* \vec{\nabla} \psi_{\text{in}} = \frac{\hbar}{m} \operatorname{Im}\left(e^{-ikz} \vec{\nabla} e^{ikz} \right) = \frac{\hbar k}{m} \hat{z} = v\hat{z} \tag{48}$$

In the last step we have used $v = \hbar k / m$. The units here are off because technically the wavefunction should be not e^{ikz} but something more like $e^{ikz}/\sqrt{V}$, where V has units of volume. But neglecting this factor in both the numerator and denominator of (38) will cause this to cancel out. Keeping that in mind, we calculate the denominator to be

$$\frac{d^2 N_{\text{in}}}{dA\,dt} = |\vec{S}_{\text{in}}| = v. \tag{49}$$

Similarly the outgoing flux is (using $\vec{\nabla} = \frac{\partial}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial}{\partial\theta}\hat{\theta} + \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\hat{\phi}$)

$$\vec{S}_{\text{scat}} = \frac{\hbar}{m} \operatorname{Im}\left(\frac{e^{-ikr}}{r} f^* ik\hat{r} \frac{e^{ikr}}{r} f + O\left(\frac{1}{r^3}\right) \right) = v\frac{\hat{r}}{r^2}|f|^2 + O(1/r^3). \tag{50}$$

To relate this to the flux per solid angle we use $d\vec{a} = r^2 d\Omega \hat{r}$ to obtain

$$\frac{d^2 N_{\text{scat}}}{d\Omega\,dt} = \vec{S}_{\text{scat}} \cdot d\vec{a} = (v\frac{\hat{r}}{r^2}|f|^2 + O(r^{-3})) \cdot r^2\,d\Omega\,\hat{r} = |f|^2 v\,d\Omega + O(1/r). \tag{51}$$

We can neglect the $O(1/r)$ term as $r \to \infty$ (and on a side note, we see now why the leading-order term in $\vec{S}_{\text{scat}}$ was $O(1/r^2)$) and obtain the simple formula

$$\boxed{\frac{d\sigma}{d\Omega} = |f(\theta,\phi)|^2}. \tag{52}$$

The Optical Theorem. (52) is valid everywhere except at $\theta = 0$. There we have to also consider interference between the scattered and the unscattered wave. (Unlike the incoming wave, the outgoing unscattered does coexist with the scattered wave.) The resulting flux is

$$\vec{S}_{\text{out}} = \frac{\hbar}{m} \operatorname{Im}\left(e^{-ikz} + \frac{f^*}{r}e^{-ikr} \right)\left(ik\hat{z}e^{ikz} + \frac{ik\hat{r}}{r}fe^{ikr} + O(1/r^2) \right) \tag{53a}$$

$$= \underbrace{v\hat{z}}_{\vec{S}_{\text{unscat}}} + \underbrace{v\frac{\hat{r}}{r^2}|f|^2}_{\vec{S}_{\text{scat}}} + \underbrace{v\operatorname{Re}\left(\frac{\hat{z}f^*}{r}e^{ik(z-r)} + \frac{\hat{r}f}{r}e^{ik(r-z)} \right)}_{\vec{S}_{\text{interference}}} \tag{53b}$$

This last term can be thought of as the effects of interference. We will evaluate it for large r and for $\theta \approx 0$. Here $\hat{r} \approx \hat{z}$ and we define $\rho = \sqrt{x^2 + y^2}$ so that (to leading order) $r = z + \frac{\rho^2}{2z}$. Then

$$\int \vec{S}_{\text{inter}} \cdot d\vec{a} = v \int_0^{2\pi} d\phi \int_0^\infty \rho\,d\rho\frac{\hat{z}}{z}\left(f^* e^{-ik\frac{\rho^2}{2z}} + fe^{ik\frac{\rho^2}{2z}} \right) \tag{54a}$$

$$= 4\pi\frac{v}{z}\operatorname{Re}\int_0^\infty \frac{d\rho^2}{2}fe^{\frac{ik\rho^2}{2z}} \tag{54b}$$

$$= \frac{2\pi v}{z}\operatorname{Re}\int_0^\infty dy\,e^{\frac{iky}{2z}}f(0) = -\frac{4\pi v}{k}\operatorname{Im} f(0) \tag{54c}$$

27

Since the outgoing flux should equal the incoming flux, we can define A to be the beam area and find

$$Av = Av + v \int d\Omega |f(\theta, \phi)|^2 - \frac{4\pi v}{k} \operatorname{Im} f(0). \tag{55}$$

Thus we obtain the following identity, known as the *optical theorem*:

$$\int d\Omega |f(\theta, \phi)|^2 = \frac{4\pi}{k} \operatorname{Im} f(0) \tag{56}$$

All this is well and good, but we have made no progress at all in computing $f(\theta, \phi)$. We will discuss two approximation methods: the partial wave method (which is exact, but yields nice approximations when k is very small) and the Born approximation, which is a good approximation when we are scattering off a weak potential. This is analogous to approximations we have seen before (see Table 1). We will discuss the Born approximation in Section 4.2 and the partial-wave

perturbation type	time-independent	time-dependent	scattering
small	TIPT	TDPT	Born
slow	WKB	adiabatic	partial wave

Table 1: Summary of approximation techniques for scattering problems.

technique in Section 4.3.

4.2 Born Approximation

Zooming out, we want to solve the following eigenvalue equation:

$$(\vec{\nabla}^2 + k^2)|\psi\rangle = U|\psi\rangle \qquad \text{where} \qquad U \equiv \frac{2m}{\hbar^2} V. \tag{57}$$

This looks like a basic linear algebra question. Can we solve it by inverting $(\vec{\nabla}^2 + k^2)$ to obtain

$$\psi \stackrel{?}{=} (\vec{\nabla}^2 + k^2)^{-1} U|\psi\rangle? \tag{58}$$

To answer this, we first review some basic linear algebra. Suppose we want to solve the equation

$$A\vec{x} = \vec{b} \tag{59}$$

for some normal matrix A. We can write $\vec{b} = A^{-1}\vec{x}$ only if A is invertible. Otherwise the solution will not be uniquely defined. More generally, suppose that our vectors live on a space V. Then we can divide up V as

$$V = \operatorname{Im} A \oplus \ker A \qquad \text{where} \qquad \ker A = \{\vec{x}_0 : A\vec{x}_0 = 0\}. \tag{60}$$

If we restrict A to the subspace $\operatorname{Im} A$ then it is indeed invertible. The solutions to (59) are then given by

$$\vec{x} = (A|_{\operatorname{Im} A})^{-1}\vec{b} + \vec{x}_0 \qquad \text{where} \qquad \vec{x}_0 \in \ker A. \tag{61}$$

Returning now to the quantum case, the operator $(\vec{\nabla}^2 + k^2)$ is certainly not invertible. States satisfying $(\vec{\nabla}^2 + k^2)|\psi_0\rangle = 0$ exist, and are plane waves with momentum $\hbar k$. But if we restrict

$(\vec{\nabla}^2 + k^2)$ to the subspace of states with momentum $\neq \hbar k$ then it is invertible. Define the Green's operator G to be $(\vec{\nabla}^2 + k^2)|_{p \neq \hbar k}^{-1}$. The calculation of G is rather subtle and the details can be found in Griffiths. However, on general principles we can make a fair amount of progress. Since $(\vec{\nabla}^2 + k^2)$ is diagonal in the momentum basis, then G should be as well. Thus G should be written as an integral over $|\vec{p}\rangle\langle\vec{p}|$ times some function of $\vec{p}$. By Fourier transforming this function we can equivalently write G in terms of translation operators as

$$G = \int d^3\vec{r}\, G(\vec{r}) T_{\vec{r}} \qquad \text{where} \qquad T_{\vec{r}} \equiv e^{-i\frac{\vec{r}\cdot\vec{p}}{\hbar}}. \tag{62}$$

Let's go through this more concretely. In the momentum basis we have the completeness relation $\int d^3\vec{p}\, |\vec{p}\rangle\langle\vec{p}| = I$ which implies

$$\vec{\nabla}^2 + k^2 = \int d^3\vec{p}\, (-\hbar^2 p^2 + k^2)|\vec{p}\rangle\langle\vec{p}|. \tag{63}$$

To invert this we might naively write $G = \int' d^3\vec{p}\, (-\hbar^2 p^2 + k^2)^{-1}|\vec{p}\rangle\langle\vec{p}|$ where $\int'$ denotes the integral over all $\vec{p}$ with $p \neq \hbar k$. To handle the diverging denominator, one method is to write

$$G = \lim_{\epsilon \to 0} \int d^3\vec{p}\, (-\hbar^2 p^2 + k^2 + i\epsilon)^{-1}|\vec{p}\rangle\langle\vec{p}|. \tag{64}$$

Finally we can write this in the position basis according to (62) and obtain the position-space Green's function $G(\vec{r})$ by Fourier-transforming $(-\hbar^2 p^2 + k^2 + i\epsilon)^{-1}$. In Griffiths this integral is carried out obtaining the answer:

$$G(\vec{r}) = -\frac{e^{ikr}}{4\pi r}. \tag{65}$$

This function $G(\vec{r})$ is called a Green's function. We can thus write

$$G = \int d^3\vec{r}\, \frac{-e^{ikr}}{4\pi r} T_{\vec{r}}. \tag{66}$$

Having computed G, we can now solve (57) and obtain

$$|\psi\rangle = |\psi_0\rangle + GU|\psi\rangle \tag{67}$$

for some free-particle solution $|\psi_0\rangle$. Indeed for a scattering problem, we should have $\psi_0(\vec{r}) = e^{ikz}$. (67) is exact, but not very useful because $|\psi\rangle$ appears on both the LHS and RHS. However, it will let us expand $|\psi\rangle$ in powers of U.

The *first Born approximation* consists of replacing the $|\psi\rangle$ on the RHS of (67) by $|\psi_0\rangle$, thus yielding

$$|\psi\rangle = |\psi_0\rangle + GU|\psi_0\rangle. \tag{68}$$

The second Born approximation consists of using (68) to approximate $|\psi\rangle$ in the RHS of (67), which yields

$$|\psi\rangle = |\psi_0\rangle + GU(|\psi_0\rangle + GU|\psi_0\rangle) = |\psi_0\rangle + GU|\psi_0\rangle + GUGU|\psi_0\rangle. \tag{69}$$

Of course we could also rewrite (67) as $|\psi\rangle = (I - GU)^{-1}|\psi_0\rangle = \sum_{n \geq 0}(GU)^n|\psi_0\rangle$ (since $(I - GU)$ is formally invertible) and truncate this sum at some finite value of n.

These results so far have been rather abstract. Plugging in (65) and $\psi_0(\vec{r}) = e^{ikz}$ we find that the first Born approximation is

$$\psi(\vec{r}) = \psi_0(\vec{r}) + \int d^3\vec{r}' \psi_0(\vec{r}') G(\vec{r} - \vec{r}') U(\vec{r}') \tag{70a}$$

$$= e^{ikz} - \int d^3\vec{r}' e^{ikz'} \frac{e^{ik|\vec{r}-\vec{r}'|}}{4\pi|\vec{r}-\vec{r}'|} U(\vec{r}') \tag{70b}$$

If we assume that the potential is short range and we evaluate this quantity for r far outside the range of U, then we will have $r \gg r'$ for all the points where the integral has a nonzero contribution. In this case $|\vec{r} - \vec{r}'| \approx r - \hat{r} \cdot \vec{r}'$. Let us further define

$$\vec{k} = k\hat{r} \qquad \text{and} \qquad \vec{k}' = k\hat{z}, \tag{71}$$

corresponding to the outgoing and incoming wavevectors respectively. Then we have (still in the first Born approximation)

$$\psi_{\text{scat}}(\vec{r}) = - \int d^3\vec{r}' e^{ikz'} \frac{e^{ik|\vec{r}-\vec{r}'|}}{4\pi|\vec{r}-\vec{r}'|} U(\vec{r}') \tag{72a}$$

$$\approx - \int d^3\vec{r}' e^{ikz'} \frac{e^{ikr-ik\hat{r}\cdot\vec{r}'}}{4\pi r} U(\vec{r}') \tag{72b}$$

$$= \left(- \int d^3\vec{r}' e^{i(\vec{k}'-\vec{k})\cdot\vec{r}'} \frac{U(\vec{r}')}{4\pi} \right) \frac{e^{ikr}}{r} \tag{72c}$$

The quantity in parentheses is then $f(\theta, \phi)$. If we define $\tilde{V}(\vec{q})$ to be the Fourier transform of $V(\vec{r})$ then we obtain

$$f_1(\theta, \phi) = -\frac{m}{2\pi\hbar^2} \tilde{V}(\vec{k}' - \vec{k}), \tag{73}$$

where f_1 refers to the first Born approximation. One very simple example is when $V(\vec{r}) = V_0 \delta(\vec{r})$. Then we simply have $f_1 = -\frac{mV_0}{2\pi\hbar^2}$.

A further simplification occurs in the case when $V(\vec{r})$ is centrally symmetric. Then

$$\tilde{V}(\vec{q}) \equiv \int d^3\vec{r} V(r) e^{i\vec{q}\cdot\vec{r}} = 2\pi \int_0^\infty dr \int_{-1}^1 d\mu r^2 V(r) e^{iqr\mu} = \frac{4\pi}{q} \int_0^\infty dr\, rV(r)\sin(qr). \tag{74}$$

Finally the momentum transfer $\vec{q} = \vec{k}' - \vec{k}$ satisfies $q = 2k\sin(\theta/2)$.

One application of this (see Griffiths for details) is the Yukawa potential: $V(r) = -\beta e^{-\mu r}/r$. The first Born approximation yields

$$f_1^{\text{Yukawa}}(\theta) = -\frac{2m\beta^2}{\hbar^2(\mu^2 + q^2)}.$$

Taking $\beta = -eQ$ and $\mu = 0$ recovers Rutherford scattering, with

$$f_1^{\text{Rutherford}}(\theta) = \frac{2meQ}{\hbar^2 q^2} = \frac{meQ}{2\hbar^2 k^2 \sin^2(\theta/2)}.$$

A good exercise is to rederive the Born approximation by using Fermi's Golden Rule and counting the number of outgoing states in the vicinity of a given $\vec{k}$. See section 7.11 of Sakurai for details. Another version is in Merzbacher, section 20.1.

30

Rigorous derivation of Green's functions. The above derivation of G was somewhat informal. However, once the form of $G(\vec{r})$ is derived informally or even guessed, it can be verified that $(\vec{\nabla}^2 + k^2)G$ acts as the identity on all states with no component in the null space of $(\vec{\nabla}^2 + k^2)$. This is the content of Griffiths problem 11.8. Implicit is that we are working in a Schwartz space which rules out exponentially growing wavefunctions, and in turn implies that $\vec{\nabla}^2$ has only real eigenvalues and is in fact Hermitian. A more rigorous derivation of the Born approximation can also be obtained by using time-dependent perturbation theory as described by Sakurai section 7.11 and Merzbacher section 20.1. I particularly recommend the discussion in Merzbacher.

4.3 Partial Waves

In this section assume we have a central potential, i.e. $V(\vec{r}) = V(r)$. Since our initial conditions are invariant under rotation about the $\hat{z}$ axis, our scattering solutions will also be independent of ϕ (but not θ).

Assume further that

$$\lim_{r \to \infty} r^2 V(r) = 0. \tag{75}$$

We will see the relevance of this condition shortly.

A wavefunction with no ϕ dependence can be written in terms of Legendre polynomials[2] (corresponding to the $m = 0$ spherical harmonics)as

$$\psi(r, \theta) = \sum_{l=0}^{\infty} R_l(r) P_l(\cos \theta). \tag{76}$$

If we define $u_l(r) = r R_l(r)$, then the eigenvalue equation becomes (for each l)

$$-u_l'' + V_{\text{eff}} u_l = k^2 u_l \qquad \text{where} \qquad V_{\text{eff}} \equiv \frac{2mV}{\hbar^2} + \frac{l(l+1)}{r^2}. \tag{77}$$

If (75) holds, then for sufficiently large r, we can approximate $V_{\text{eff}} \approx l(l+1)/r^2$. In this region the solutions to (77) are given by the *spherical Bessel functions*. Redefining $x \equiv kr$ and using the assumption $V_{\text{eff}} \approx l(l+1)/r^2$, (77) becomes

$$u_l'' - \frac{l(l+1)}{x^2} u_l = -u_l. \tag{78}$$

This has two linearly independent solutions: $x j_l(x)$ and $x n_l(x)$ where $j_l(x), n_l(x)$ are the spherical Bessel functions of the first and second kind respectively, and are defined as

$$j_l(x) = (-x)^l \left(\frac{1}{x} \frac{d}{dx} \right)^l \frac{\sin(x)}{x} \qquad \text{and} \qquad n_l(x) = -(-x)^l \left(\frac{1}{x} \frac{d}{dx} \right)^l \frac{\cos(x)}{x} \tag{79}$$

These can be thought of as "sin-like" and "cos-like" respectively. For $l = 0, 1$, (79) becomes

$$j_0(x) = \frac{\sin(x)}{x} \qquad\qquad j_1(x) = \frac{\sin(x)}{x^2} - \frac{\cos(x)}{x} \tag{80a}$$

$$n_0(x) = -\frac{\cos(x)}{x} \qquad\qquad n_1(x) = -\frac{\cos(x)}{x^2} - \frac{\sin(x)}{x} \tag{80b}$$

---

[2]What are Legendre polynomials? One definition starts with the orthogonality condition $\int_{-1}^{1} P_m(x) P_n(x) dx = \frac{2}{2n+1} \delta_{mn}$. (The δ_{mn} is the important term here; $\frac{2}{2n+1}$ is a somewhat arbitrary convention.) Then if we apply the Gram-Schmidt procedure to $1, x, x^2, \ldots$, we obtain the Legendre polynomials $P_0 = 1, P_1 = 1, P_2 = \frac{1}{2}(3x^2 - 1), \ldots$. Thus any degree-$n$ polynomial can be written as a linear combination of $P_0, \ldots, P_n$ and vice-versa.

The reason for (76) is that if $\psi(r, \theta)$ is independent of ϕ then we can write it as a power series in r and z, or equivalently r and $\frac{z}{r} = \cos(\theta)$. These power series can always be written in the form of (76).

One can check (by evaluating the derivatives in (79) repeatedly and keeping only the lowest power of $1/x$) that as $x \to \infty$ we have the asymptotic behavior

$$j_l(x) \to \frac{1}{x}\sin(x - l\pi/2) \qquad \text{and} \qquad n_l(x) \to -\frac{1}{x}\cos(x - l\pi/2). \qquad (81)$$

On the other hand, in the $x \to 0$ limit we can keep track of only the lowest power of x to find that as $x \to 0$ we have

$$j_l(x) \to \frac{x^l}{(2l+1)!!} \qquad \text{and} \qquad n_l(x) \to -\frac{(2l-1)!!}{x^{l+1}}, \qquad (82)$$

where $(2l+1)!! \equiv (2l+1)(2l-1)\cdots 3 \cdot 1$.

If j_l and n_l are sin-like and cos-like, it will be convenient to define functions that resemble incoming and outgoing waves. These are the spherical Hankel functions of the first and second kind:

$$h_l^{(1)} = j_l + i n_l \qquad \text{and} \qquad h_l^{(2)} = j_l - i n_l \qquad (83)$$

For large r, $h_l^{(1)}(kr) \to (-i)^{l+1}\frac{e^{ikr}}{kr}$, and so our scattered wave should be proportional to $h_l^{(1)}$. More precisely, $R_l(r)$ (i.e. the angular-momentum-l component) should be proportional to $h_l^{(1)}(kr)$. Putting this together we get

$$\psi(r,\theta) \stackrel{\text{when } V \approx 0}{=} e^{ikz} + \underbrace{\sum_{l\geq 0} c_l h_l^{(1)}(kr) P_l(\cos(\theta))}_{\psi_{\text{scat}}}, \qquad (84)$$

for some coefficients c_l. It will be convenient to write the c_l in terms of new coefficients a_l as $c_l = k i^{l+1}(2l+1)a_l$, so that we have

$$\psi_{\text{scat}}(r,\theta) \stackrel{\text{when } V \approx 0}{=} k\sum_{l\geq 0} i^{l+1}(2l+1)a_l h_l^{(1)}(kr) P_l(\cos(\theta)) \qquad (85)$$

$$\stackrel{\text{when } r \to \infty}{=} \underbrace{\sum_{l\geq 0}(2l+1)a_l P_l(\cos(\theta))}_{f(\theta)} \frac{e^{ikr}}{r} \qquad (86)$$

We can then compute the differential cross-section in terms of the a_l

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2 = \left|\sum_l (2l+1)a_l P_l(\cos(\theta))\right|^2. \qquad (87)$$

Recall that the Legendre polynomials satify the orthogonality relation

$$\int_{-1}^1 dz\, P_l(z) P_{l'}(z) = \delta_{l,l'} \frac{2}{2l+1}. \qquad (88)$$

Thus we can calculate

$$\sigma = \int d\Omega \frac{d\sigma}{d\Omega} = 4\pi \sum_l (2l+1)|a_l|^2. \qquad (89)$$

While the differential cross section involves interference between different values of l, the total cross section is simply given by an incoherent sum over l. Intuitively this is because we can think of l and θ as conjugate observables, analogous to momentum and position. If we measure the probability

of observing something at a particular position, we will see interference effects between different momenta, but if we integrate over all positions, these will go away.

The beauty of the partial-wave approach is that it reduces to a series of 1-d scattering problems, one for each value of l. We have written down the form of the scattered wave. For the incoming wave, we need to express e^{ikz} in the form of (76). From the arguments above we can see that e^{ikz} should have form $\sum_l (A_l j_l(kr) + B_l n_l(kr)) P_l(\cos\theta)$. The specific solution is given by *Rayleigh's formula* which we state without proof:

$$e^{ikz} = e^{ikr\cos\theta} = \sum_{l\geq 0} i^l (2l+1) j_l(kr) P_l(\cos\theta) \tag{90}$$

Plugging this in, we find that in the $V \approx 0$ region we have

$$\psi(r,\theta) \overset{\text{when } V\approx 0}{=} \sum_{l\geq 0} i^l (2l+1) P_l(\cos\theta) \left[\underbrace{j_l(kr)}_{\text{plane wave}} + \underbrace{ik a_l h_l^{(1)}(kr)}_{\text{scattered}} \right] \tag{91}$$

$$= \frac{1}{2} \sum_{l\geq 0} i^l (2l+1) P_l(\cos\theta) \left[\underbrace{h_l^{(1)}(kr)}_{\text{outgoing}} (1 + 2ik a_l) + \underbrace{h_l^{(2)}(kr)}_{\text{incoming}} \right] \tag{92}$$

Now we have really expressed this as a series of 1-d scattering problems. Here comes the crucial simplifying move. Because we are assuming that the collision is elastic, probability and angular-momentum conservation means that "what goes in must come out"; i.e.

$$|1 + 2ik a_l| = 1. \tag{93}$$

The outgoing wave's amplitude must have the same absolute value as the incoming wave's amplitude. We can rewrite (93) by introducing the *phase shift* δ_l defined by

$$1 + 2ik a_l = e^{2i\delta_l}. \tag{94}$$

The factor of 2 is conventional. We can rewrite (94) to solve for a_l as

$$a_l = \frac{2e^{i\delta_l} - 1}{2ik} = \frac{e^{i\delta_l} \sin(\delta_l)}{k} = \frac{1}{k} \frac{1}{\cot(\delta_l) - i}. \tag{95}$$

Many equivalent expressions are also possible. In terms of the phase shifts, we can write

$$f(\theta) = \frac{1}{k} \sum_l (2l+1) P_l(\cos\theta) e^{i\delta_l} \sin(\delta_l) \tag{96}$$

$$\sigma = \frac{4\pi}{k^2} \sum_l (2l+1) \sin^2(\delta_l). \tag{97}$$

As an application we can verify that this satisfies the optical theorem. Observe that $P_l(1) = 1$ for all l. Thus

$$\frac{4\pi}{k} \operatorname{Im} f(0) = \frac{4\pi}{k} \frac{1}{k} \sum_l (2l+1) P_l(1) \sin^2(\delta_l) = \sigma. \tag{98}$$

Another easy application is *partial-wave unitarity*, which bounds the total amount of scattered wave with angular momentum l. Define $\sigma_l = \frac{4\pi}{k^2}(2l+1)\sin^2(\delta_l)$, so that $\sigma = \sum_l \sigma_l$. Then using $\sin^2(\delta_l) \leq 1$ we have

$$\sigma_l \leq \frac{4\pi}{k^2}(2l+1). \tag{99}$$

33

This bound is called "partial-wave unitarity."

How to compute phase shifts. Let us look again at the $r \to \infty$ solution. We can use the fact that overall phase and normalization don't matter to obtain

$$R_l(r) = h_l^{(1)}(kr)e^{2i\delta_l} + h_l^{(2)}(kr) \tag{100a}$$

$$= (1 + e^{2i\delta_l})j_l(kr) + i(e^{2i\delta_l} - 1)n_l(kr) \tag{100b}$$

$$\propto \cos(\delta_l)j_l(kr) - \sin(\delta_l)n_l(kr) \tag{100c}$$

$$\propto j_l(kr) - \tan(\delta_l)n_l(kr) \tag{100d}$$

Suppose that we know the interior solution $R_l(r)$ for $r \le b$ and that $V(r) \approx 0$ for $r > b$. Then we can compute the phase shift by matching $\frac{R_l'(r)}{R_l(r)}$ for $r = b \pm \epsilon$.

Here is a simple example. Consider a hard sphere of radius b. Then $R_l(r) = 0$ for $r \le b$ and we have

$$j_l(kb) - \tan(\delta_l)n_l(kb) = 0 \implies \delta_l = \tan^{-1}\left(\frac{j_l(kb)}{n_l(kb)}\right). \tag{101}$$

One particularly simple case is when $l = 0$. Then

$$\delta_0 = \tan^{-1}\left(\frac{j_0(kb)}{n_0(kb)}\right) = \tan^{-1}\left(\frac{\frac{\sin(kb)}{kb}}{-\frac{\cos(kb)}{kb}}\right) = -\tan^{-1}(\tan(kb)) = -kb. \tag{102}$$

It turns out that in general repulsive potentials yield negative phase shifts.

What about larger values of l? Suppose that $kb \ll 1$. Then using the $x \to 0$ approximations for $j_l(x), n_l(x)$, we obtain

$$\delta_l \xrightarrow{k \to 0} \tan^{-1}\left(\frac{(kb)^{2l+1}}{(2l+1)!!(2l-1)!!}\right) \approx \frac{(kb)^{2l+1}}{2l+1!!(2l-1)!!} \sim (kb)^{2l+1}. \tag{103}$$

Thus the $l = 0$ scattering dominates. In terms of cross-sections, we have

$$\sigma_l = \frac{4\pi}{k^2}(2l+1)\sin^2(\delta_l) \approx \frac{4\pi}{(2l+1)((2l-1)!!)^4}(kb)^{4l}b^2. \tag{104}$$

For $l = 0$ this is $4\pi b^2$ which four times the classical value of πb^2, and for higher value of l this drops exponentially (assuming $kb \ll 1$). Even if $kb > 1$ this drops exponentially once $l \gg kb$, which confirms our intuition that the angular momentum should be on the order of $\hbar kb$.

Another way to think about this reason to favor low values of l is because the $l(l+1)/r^2$ term in V_{eff} forms an angular momentum "barrier" that prevents low-energy incoming waves from penetrating to small enough r to see the potential $V(r)$.

The high-energy limit. The partial wave approximation is easiest to use in the low-k limit because then we can restrict our attention to a few values of l, or even just $l = 0$. But for the hard sphere we can also evaluate the $kb \gg 1$ limit. In this case we expect to find angular momenta up to $l_{\text{max}} \equiv kb$. Thus we approximate the total cross section by

$$\sigma \approx \frac{4\pi}{k^2}\sum_{l=0}^{l_{\text{max}}}(2l+1)\sin^2(\delta_l). \tag{105}$$

The phases δ_l will vary over the entire range from 0 to 2π so we simply approximate $\sin^2(\delta_l)$ by its average value of $1/2$. Thus we obtain

$$\sigma \approx \frac{2\pi}{k^2} \sum_{l=0}^{kb} (2l+1) = 2\pi b^2. \tag{106}$$

This is now *twice* the classical result. Even though the particles are moving quickly they still diffract like waves. One surprising consequence is that even though a hard sphere leaves a shadow, there is a small bright spot at the center of the shadow. Indeed the optical theorem predicts that $\text{Im}\, f(0) = \frac{k}{4\pi}\sigma \approx kb^2/2$. Thus $|f(0)|^2 \geq (kb)^2 b^2/4$. For a further discussion of this bright spot and the role it played in early 19th-century debates about whether light is a particle and/or a wave, look up "Arago spot" on wikipedia.

Phase shifts. As we have seen, scattering can be understood in terms of phase shifts. Now we describe a simple physical way of seeing this. If $V = 0$ then a plane wave has $u_0(r) = \sin(kr)$, due to the $u_0(0) = 0$ boundary condition. When there is scattering, the phase shift δ_0 will become nonzero and we will have

$$u_0(r) = \sin(kr + \delta_0).$$

If the potential is attractive then the phase will oscillate more rapidly in the scattering region and so we will have $\delta_0 > 0$ while if it is repulsive then the phase will oscillate more slowly and we will have $\delta_0 < 0$. See Fig. 2 for an illustration.
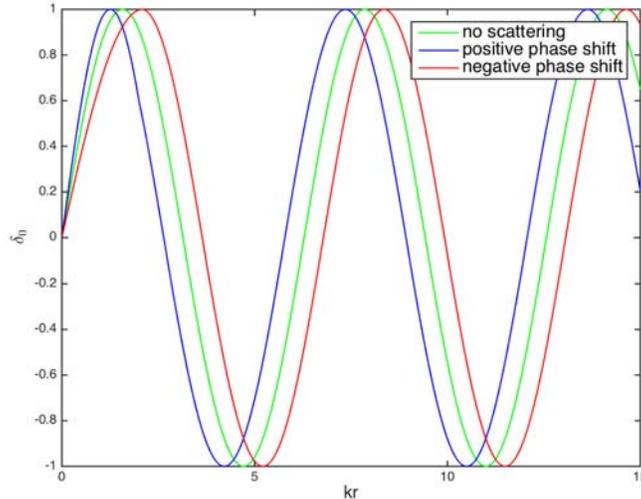

Figure 2: The phase shift δ_0 is positive for attractive potentials and negative for repulsive potentials.

Scattering length. In the regime of low k it turns out that many potentials behave qualitatively like what we have seen with the hard sphere, with a characteristic length scale called a "scattering length." To derive this suppose there is some b such that $V(r) \approx 0$ for $r \geq b$. In this region we have $u_0(r) \approx \sin(kb + \delta_0)$ (neglecting normalization). In the vicinity of b we have

$$u_0(r) \approx u_0(b) + u_0'(b)(r - b).$$

If we extrapolate to smaller values of r our approximation hits 0 at $r = a$ where

$$a = b - \frac{u_0(b)}{u_0'(b)} = b - \frac{\tan(kb + \delta_0)}{k}.$$

Using the tan addition formula and taking the limit $k \to 0$ we find

$$a = b - \frac{1}{k}\frac{\tan(kb) + \tan(\delta_0)}{1 - \tan(kb)\tan(\delta_0)} \approx= b - \frac{kb + \tan(\delta_0)}{k} = -\frac{\tan(\delta_0)}{k}. \tag{107}$$

Rearranging we have $\tan(\delta_0) = -ka$, and in the $ka \ll 1$ limit this yields

$$\sigma_0 = \frac{4\pi}{k^2}\sin^2(\delta_0) \approx 4\pi a^2,$$

which is again the hard sphere result. Similar results hold for larger value of l. Thus the scattering length can be thought of as an effective size of a scattering target.

8.06 Quantum Physics III
Spring 2016