

Lecture Notes for Quantum Mechanics II

Spring 2011

Lecturer: Professor *Boris Altshuler*
Transcriber: *Alexander Chen*

May 10, 2011

Contents

| | | |
|----------|--|-----------|
| 1 | Lecture 1 | 4 |
| 1.1 | Symmetries | 4 |
| 1.2 | Continuous Symmetries | 4 |
| 1.3 | Discrete Symmetries | 6 |
| 2 | Lecture 2 | 8 |
| 2.1 | Parity Continued | 8 |
| 2.2 | Examples | 8 |
| 3 | Lecture 3 | 11 |
| 3.1 | Parity Selection Rule | 11 |
| 3.2 | Lattice Translations | 12 |
| 3.3 | Time Reversal | 13 |
| 4 | Lecture 4 | 15 |
| 4.1 | Time Reversal Continued | 15 |
| 5 | Lecture 5 | 18 |
| 5.1 | Time Reversal Symmetry Continued | 18 |
| 5.2 | Approximation Methods | 19 |
| 5.3 | Perturbation Theory | 19 |
| 6 | Lecture 6 | 22 |
| 6.1 | Time-independent Perturbation Theory Continued | 22 |
| 6.2 | Normalization of States | 23 |
| 6.3 | Example | 23 |
| 6.4 | Degenerate Case | 24 |
| 7 | Lecture 7 | 25 |
| 7.1 | Examples: Band Structure | 25 |
| 7.2 | Examples: Stark Effect | 27 |

| | | |
|-----------|--|-----------|
| 8 | Lecture 8 | 29 |
| 8.1 | Spin-Orbit Coupling | 29 |
| 8.2 | The Zeeman Effect | 30 |
| 9 | Lecture 9 | 32 |
| 9.1 | Van der Waals Interaction | 32 |
| 9.2 | Polarization | 33 |
| 9.3 | Summarizing Perturbation Theory | 34 |
| 10 | Lecture 10 | 36 |
| 10.1 | Variational Method | 36 |
| 10.2 | Example | 36 |
| 10.3 | Generalization | 37 |
| 11 | Lecture 11 | 40 |
| 11.1 | Time-dependent Perturbation | 40 |
| 11.2 | Examples | 41 |
| 12 | Lecture 12 | 43 |
| 12.1 | Time-dependent Perturbation Theory Continued | 43 |
| 12.2 | Formal Development | 43 |
| 13 | Lecture 13 | 47 |
| 13.1 | Time-dependent Perturbation Continued | 47 |
| 14 | Lecture 14 | 51 |
| 15 | Lecture 15 | 55 |
| 15.1 | Identical Particles | 55 |
| 16 | Lecture 16 | 58 |
| 16.1 | Identical Particles Continued | 58 |
| 17 | Lecture 17 | 63 |
| 17.1 | Second Quantization | 63 |
| 18 | Lecture 18 | 68 |
| 18.1 | Second Quantization Reviewed | 68 |
| 19 | Lecture 19 | 72 |
| 19.1 | Scattering Theory | 72 |
| 20 | Lecture 20 | 75 |
| 20.1 | Scattering Continued | 75 |
| 20.2 | Born Approximation | 75 |
| 20.3 | Beyond Born Approximation | 77 |

| | |
|---|------------|
| 21 Lecture 21 | 79 |
| 21.1 Scattering Theory Reviewed | 79 |
| 21.2 Optical Theorem | 80 |
| 21.3 Spherically Symmetric Scattering | 81 |
| 22 Lecture 22 | 83 |
| 22.1 Spherical Symmetric Scattering Reviewed | 83 |
| 22.2 Examples | 85 |
| 23 Lecture 23 | 87 |
| 23.1 Scattering off a Central Potential, Review | 87 |
| 23.2 Finite Potential at Finite Distance | 88 |
| 24 Lecture 24 | 91 |
| 25 Lecture 25 | 94 |
| 25.1 Coulomb Scattering | 95 |
| 26 Lecture 26 | 98 |
| 26.1 Quantum Computation | 98 |
| 27 Lecture 27 | 101 |
| 27.1 Quantum Information Continued | 101 |

1 Lecture 1

Textbook will be Sakurai, but will insert things from other places because that book is a bit outdated. This course is also about tools, because QM will be used in condense matter physics and high energy physics as a tool.

1.1 Symmetries

Symmetry is a very useful concept. Sometimes we can get useful things which is otherwise obscure out of a problem using the symmetry in the problem itself. Let's consider an example of classical mechanics first. Suppose the Lagrangian $\mathcal{L}(q, \dot{q})$ depends on the coordinates and its derivative. If we have

$$\frac{\partial \mathcal{L}}{\partial q_i} = 0 \quad (1.1)$$

for some q_i then we have the Lagrange equations of this coordinate

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) = \frac{\partial \mathcal{L}}{\partial q} = 0 \quad (1.2)$$

So the quantity $\partial \mathcal{L} / \partial \dot{q}$ will be a constant which doesn't change with time. In classical mechanics we know if $\partial \mathcal{L} / \partial x = 0$ then the corresponding momentum is conserved. This is an explicit example of this symmetry.

Equivalently we can work in the Hamiltonian picture. The Hamiltonian depends on the coordinates and momenta $\mathcal{H}(q_i, p_i)$. Because of the Hamiltonian equations we know that

$$\frac{\partial \mathcal{H}}{\partial q_i} = 0 \quad \implies \quad \dot{p}_i = 0 \quad (1.3)$$

This applies to translation or rotation, etc. However this symmetry may not be apparent, and we need to change into a new set of generalized coordinates to discover the symmetry of the above form. Say azimuthal symmetry in spherical coordinates.

1.2 Continuous Symmetries

Now we come to QM. Let \hat{S} be a symmetry operator. It should be a unitary operator satisfying

$$\hat{S}^\dagger \hat{S} = 1 \quad (1.4)$$

This is because we want the symmetry operation to preserve norm of states. This is the basic assumption, because otherwise we can hardly call it a symmetry operation. We know that we can always write an infinitesimal unitary operator as

$$\hat{S} = 1 - \frac{i\epsilon}{\hbar} \hat{G} \quad (1.5)$$

where $\hat{G}^\dagger = \hat{G}$ is a hermitian operator. We call \hat{G} the generator of the symmetry transformation.

The above construction applies to any unitary operator. The fact that \hat{S} is a symmetry operator means that we have the following relation

$$\hat{S}^\dagger \mathcal{H} \hat{S} = \mathcal{H} \quad (1.6)$$

It follows that the Hamiltonian \mathcal{H} commutes with the symmetry generator \hat{G} , $[\hat{G}, \hat{\mathcal{H}}] = 0$. This implies that

$$\frac{d\hat{G}}{dt} = 0 \quad (1.7)$$

which is a result of the Heisenberg equation of motion.

Let's consider some examples. In the case of translation, the symmetry operator is just translation $T(\mathbf{l})$ by a fixed vector \mathbf{l} . The generator is just the momentum \mathbf{p} . If we consider 3D rotation then the operator is just a rotation by a fixed angle around a vector, and the generators are the angular momentum operators. Now suppose we have some initial state which is an eigenstate of the symmetry operator $\hat{G}|g\rangle = g|g\rangle$. We want to ask after time t what is the state. We can just do the usual time evolution

$$|g, t\rangle = \hat{U}(t, t_0)|g, t_0\rangle \quad (1.8)$$

The time evolution operator depends purely on the Hamiltonian, so we know that \hat{G} commutes with the time evolution operator, so we know

$$\hat{G}|g, t\rangle = \hat{U}\hat{G}|g, t_0\rangle = g|g, t\rangle \quad (1.9)$$

So the eigenket remains an eigenket with the same eigenvalue in all time.

One of the most important consequences of symmetries are degeneracies of the quantum states. Degeneracy means that there is more than one states that correspond to the same energy. It is important to study degeneracies because in many physics problems we are interested in the energy spectrum of the system. We know that

$$[\hat{\mathcal{H}}, \hat{G}] = 0 \quad (1.10)$$

so if $|\alpha, E\rangle$ is an eigenstate of the Hamiltonian with eigenvalue E , then if we apply the symmetry operator to the state we will get an energy eigenstate with the same eigenvalue E

$$\hat{\mathcal{H}}(\hat{S}|\alpha, E\rangle) = E\hat{S}|\alpha, E\rangle \quad (1.11)$$

Now if we label the symmetry operator by a continuous variable, for example θ in rotation, then in this way we can generate a whole class of energy eigenstates all with the same energy eigenvalue.

Now let's be specific and consider rotations only. We know that $[\hat{\mathbf{J}}, \hat{\mathcal{H}}] = 0$ and also $[\hat{J}^2, \hat{\mathcal{H}}] = 0$. If we label the states using $|n, m, j\rangle$, then in principle energy can depend on all three quantum numbers. But in this case we know that $\mathcal{D}|n, j, m\rangle$ is an eigenstate of the Hamiltonian with the same energy as $|n, j, m\rangle$. From rotation theory we know the rotated state is a linear combination of $|n, j, m'\rangle$. So we have to demand that energy do not depend on the quantum number m . This introduces a $(2j + 1)$ -fold degeneracy. Note that this is not a result of the Schrödinger equation. We haven't specified the nature of the angular momentum J . But this is a general result which applies to all kinds of angular momenta. In hydrogen atom we have additional accidental degeneracy so that the energy doesn't depend on l , but in general this is the minimal number of degeneracy in a spherically symmetric system.

Consider spin-orbit coupling

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{\mathcal{H}}_{SO}, \quad \hat{\mathcal{H}}_{SO} = \lambda_{SO}\mathbf{S} \cdot \mathbf{L} \quad (1.12)$$

In the case when $\lambda_{SO} = 0$ there is a $(2l + 1)(2s + 1)$ degeneracy. When we turn on spin-orbit coupling we have a $(2j + 1)$ degeneracy. For example when $l = 1, s = 1/2$, we originally have 6-fold degeneracy and the spin-orbit coupling lifts it for different j value. There is a 2-fold degeneracy for $j = 1/2$ and a 4-fold degeneracy for $j = 3/2$.

1.3 Discrete Symmetries

The above consideration about the symmetry operator assumes that the symmetry operator depends on a continuous variable, thus a continuous symmetry. We can also consider discrete symmetries. One important discrete symmetry is parity $\hat{\Pi}$, which is just space inversion. If our coordinate system consists of (x, y, z) coordinates, then the effect of a parity transformation is to invert all of the three coordinates $(x, y, z) \rightarrow (-x, -y, -z)$. This is not a rotation because there is no way to rotate all the axes by π at the same time.

If we apply the parity operation to a quantum state, then trivially we have $|\alpha\rangle \rightarrow \hat{\Pi}|\alpha\rangle$. The expectation value of the position operator becomes

$$\langle \mathbf{x} \rangle_\alpha = \langle \alpha | \mathbf{x} | \alpha \rangle \longrightarrow \langle \alpha | \hat{\Pi}^\dagger \mathbf{x} \hat{\Pi} | \alpha \rangle = - \langle \alpha | \mathbf{x} | \alpha \rangle \quad (1.13)$$

As with a continuous symmetry the parity operator is unitary $\hat{\Pi}^\dagger \hat{\Pi} = 1$. From what we have above we know that

$$\mathbf{x} \hat{\Pi} = - \hat{\Pi} \mathbf{x} \quad (1.14)$$

So the position operator anticommutes with the parity operator. Now suppose we operate the parity operator on a position eigenstate, then we should have

$$\hat{\Pi} |x\rangle = e^{i\delta} |-x\rangle \quad (1.15)$$

There is no apparent way to decide the phase factor δ , so we use for simplicity that $\delta = 0$. If we apply the operator twice then we will get back the same state. So we know that $\hat{\Pi}^2 = 1$. Combined with unitarity we know the parity operator is both hermitian and unitary

$$\hat{\Pi} = \hat{\Pi}^\dagger = \hat{\Pi}^{-1} \quad (1.16)$$

This implies that the only possible eigenvalues are ± 1 . Obviously $|x\rangle$ is not an eigenstate because it changes into $|-x\rangle$ but we can construct linear combinations of the position eigenkets to be the eigenkets of the parity operator.

For an arbitrary operator \hat{A} if $\hat{A}\hat{\Pi} = \hat{\Pi}\hat{A}$ then we say \hat{A} is even, whereas if $\hat{A}\hat{\Pi} = -\hat{\Pi}\hat{A}$ then we say \hat{A} is odd. Obviously \mathbf{x} is odd, and the momentum operator, defined to be $\mathbf{p} = m d\mathbf{x}/dt$, is also odd. We can investigate this more deeply by considering an infinitesimal translation

$$T(d\mathbf{x}) = 1 - \frac{i}{\hbar} \mathbf{p} \cdot d\mathbf{x} \quad (1.17)$$

Then we should have

$$\hat{\Pi} T(d\mathbf{x}) = T(-d\mathbf{x}) \hat{\Pi} \quad (1.18)$$

This means that if we translate first and then do the inversion, it is equivalent to do the parity first and then translate to the other direction. So if we expand the equation using \mathbf{p} then we can find $\hat{\mathbf{p}} \hat{\Pi} = -\hat{\Pi} \hat{\mathbf{p}}$.

What about angular momentum operator? For example, the orbital angular momentum is defined as

$$\mathbf{L} = \mathbf{x} \times \mathbf{p} \quad (1.19)$$

Since \mathbf{x} is odd and \mathbf{p} is odd, we should expect \mathbf{L} to be even. So we have $[\hat{\mathbf{L}}, \hat{\Pi}] = 0$. In fact we can go further and prove that

$$[\hat{\Pi}, \hat{\mathbf{J}}] = 0 \quad (1.20)$$

for all kinds of angular momenta. So angular momentum does not change sign under parity inversion. Note that under rotation, $\mathbf{x}, \mathbf{p}, \mathbf{L}$, all transform the same way as a vector. But from the point of view of reflections, \mathbf{x} and \mathbf{p} are odd and they are called *polar vectors* or just *vectors* whereas \mathbf{J} which is even under reflection is called a *pseudovector*, or *axial vector*. This is also true for all cross products of two polar vectors. From the above we will expect that $\mathbf{x} \cdot \mathbf{p}$ and \mathbf{J}^2 are scalars under both rotation and reflection, but the product $\mathbf{x} \cdot \mathbf{J}$ is invariant under rotation but odd under reflection, so we call it a *pseudoscalar*.

Let's now consider a wavefunction $\psi(x) = \langle x | \alpha \rangle$. The effect of the parity operation on this wavefunction is

$$\psi(x) = \langle x | \alpha \rangle \longrightarrow \langle x | \hat{\Pi} | \alpha \rangle = \langle -x | \alpha \rangle = \psi(-x) \quad (1.21)$$

If initially we know the wavefunction to be even or odd, $\psi(-x) = \pm\psi(x)$, then the above equation means that this is an eigenfunction of the parity operator. However, the state $|\mathbf{p}\rangle$ does not have a definite parity because in general it will not be proportional to itself after parity transformation. Suppose we are considering a eigenstate of the orbital angular momentum operator labelled by $|n, l, m\rangle$ then we have

$$\hat{\Pi} |n, l, m\rangle = (-1)^l |n, l, m\rangle \quad (1.22)$$

This is because \mathbf{L} and $\hat{\Pi}$ commute, so they can have simultaneous eigenstates. The above relation can be checked directly using the spherical harmonic function representation of the angular momentum eigenstates, and the effect of the parity operation is

$$r \rightarrow r, \quad \theta \rightarrow \pi - \theta, \quad \varphi \rightarrow \varphi + \pi \quad (1.23)$$

And the fact that L_+ and L_- commute with $\hat{\Pi}$ means that the parity does not depend on the quantum number m .

Next time we will show that if \mathcal{H} and $\hat{\Pi}$ commute and if $|n\rangle$ is non-degenerate then $|n\rangle$ has definite parity.

2 Lecture 2

2.1 Parity Continued

Remember symmetry means that there is some generator that commutes with the Hamiltonian

$$[\mathcal{H}, S] = 0 \quad (2.1)$$

We also mentioned that symmetries will introduce degeneracies of quantum states. It is not necessary that all symmetries give degeneracy, for example in harmonic oscillator potential, there is always the parity symmetry, but there is no degenerate odd and even states.

Remember we have continuous and discrete symmetries. Last time we started a somewhat detailed discussion of parity symmetry. Remember it is just the operation of inverting all of the spatial axes. $\mathbf{x} \rightarrow -\mathbf{x}$. We denote the quantum mechanical operator as $\hat{\Pi}$. It was shown that this operator is both unitary and hermitian

$$\hat{\Pi}^\dagger = \hat{\Pi}^{-1} = \hat{\Pi} \quad (2.2)$$

This constraints the eigenvalues of the operator to be ± 1 . By its definition we saw that parity operator anticommutes with the position and momentum operators and commutes with the angular momentum operator

$$\{\hat{\Pi}, \mathbf{x}\} = 0, \quad \{\hat{\Pi}, \mathbf{p}\} = 0, \quad [\hat{\Pi}, \mathbf{J}] = 0 \quad (2.3)$$

We call the vectors that anticommute with parity polar vectors, and those commuting with parity pseudovectors. Similarly we call operators that commute with $\hat{\Pi}$ scalars and operators anticommute with $\hat{\Pi}$ pseudoscalars.

We call quantum states odd or even if their wavefunctions satisfy

$$\psi(-x) = \pm \psi(x) \quad (2.4)$$

And remember Y_ℓ^m has parity $(-1)^\ell$. Suppose we have a Hamiltonian that commutes with the parity operator, and the energy eigenstate $|n\rangle$ is nondegenerate, then $|n\rangle$ is also an eigenstate of $\hat{\Pi}$. The proof is quite straightforward. Because $\hat{\Pi}^2 = 1$, we can see obviously that $\frac{1}{2}(1 \pm \hat{\Pi})|n\rangle$ is an eigenstate of the parity operator

$$\hat{\Pi} \frac{1}{2} (1 \pm \hat{\Pi}) |n\rangle = \frac{1}{2} (\hat{\Pi} \pm \hat{\Pi}^2) |n\rangle = \pm \frac{1}{2} (1 \pm \hat{\Pi}) |n\rangle \quad (2.5)$$

Because the parity operator commutes with the Hamiltonian, the above two states, by construction are also energy eigenstates with energy E_n . But by assumption we know that the energy spectrum is non-degenerate, so one of the above states must be zero, the other proportional to $|n\rangle$. This proves that the state $|n\rangle$ is an eigenstate of the parity operator.

2.2 Examples

Let's see some examples. Consider simple harmonic oscillator, where we know that $[\mathcal{H}, \hat{\Pi}] = 0$, since x^2 and p^2 are scalars. We know the energy eigenstates to be

$$|0\rangle, \quad |n\rangle = \frac{1}{\sqrt{n!}} (\hat{a}^\dagger)^n |0\rangle \quad (2.6)$$

We know that the ground state is a gaussian, therefore even. We can also get this fact by a simple argument. Remember the ground state is defined by $\hat{a}|0\rangle = 0$, and that both \hat{a} and \hat{a}^\dagger are linear combinations of \hat{x} and \hat{p} , so both of them should anticommute with the parity operator. Now if we operate \hat{a} on the wave function $\psi_0(x) = \langle x|0\rangle$ then we will get a differential equation which tells us that the wavefunction is even. For the excited states we can use the anticommutation relation

$$\hat{\Pi}|1\rangle = \hat{\Pi}\hat{a}^\dagger|0\rangle = -\hat{a}^\dagger\hat{\Pi}|0\rangle = -|1\rangle \quad (2.7)$$

Now let's consider a symmetric double well potential. For simplicity we consider finite square wells. Because it is a 1D problem we know that the energy eigenstates all have even or odd parity. If we call the even state $|S\rangle$ and the odd state $|A\rangle$, then we have

$$\hat{\Pi}|S\rangle = |S\rangle, \quad \hat{\Pi}|A\rangle = -|A\rangle \quad (2.8)$$

and that $E_S < E_A$. This can be seen when we take the barrier in the middle to zero, then the well becomes a infinite square well, where the antisymmetric state is the excited state. If we take the barrier to be infinitely large, then the two states are degenerate. We can define the left and right states by taking the symmetric and antisymmetric combinations of the two states, and these are not parity eigenstates

$$|L\rangle = \frac{1}{\sqrt{2}}(|S\rangle - |A\rangle), \quad |R\rangle = \frac{1}{\sqrt{2}}(|S\rangle + |A\rangle) \quad (2.9)$$

Now let's consider time evolution and take the initial state to be $|\alpha, t=0\rangle = |L\rangle$, then the subsequent states will become

$$|\alpha, t > 0\rangle = \frac{1}{\sqrt{2}} \left(e^{-iE_S t/\hbar} |S\rangle - e^{-iE_A t/\hbar} |A\rangle \right) \propto |S\rangle - e^{-i\omega t} |A\rangle \quad (2.10)$$

So if $E_S \neq E_A$ then we have tunnelling between the two states and the oscillation frequency between the two is $\omega = (E_A - E_S)/\hbar$.

A concrete example of this comes from organic molecules, for example NH_3 . If we take the hydrogen atoms to be in the same plane, then the nitrogen atom can be either above or below the plane. Because there is some energy difference between them, these will oscillate with frequency $\omega = 24 \text{ MHz}$.

Recall our discussion of tunnelling using the WKB approximation. When the well is symmetric and barrier is very large, then the probability for tunnelling will be exponentially small, and the energy difference $E_A - E_S$ will also be exponentially small.

Suppose now we have an asymmetric double well. If we neglect tunnelling, we can use $|L\rangle$ and $|R\rangle$ as a basis with energies ϵ_L and ϵ_R . Any state can be written as

$$|\alpha\rangle = c_L |L\rangle + c_R |R\rangle \quad (2.11)$$

Now we want to account for tunnelling effects. We modify our Hamiltonian to have a slight off-diagonal perturbation $\langle L|\mathcal{H}|R\rangle = I$. It is easy to diagonalize this 2×2 Hamiltonian

$$E^2 - (\epsilon_L + \epsilon_R)E + (\epsilon_L\epsilon_R - |I|^2) = 0 \quad (2.12)$$

with energies

$$E = \frac{\epsilon_L + \epsilon_R}{2} \pm \sqrt{\frac{(\epsilon_L - \epsilon_R)^2}{4} + |I|^2} \quad (2.13)$$

Now that we have the expression, let's consider two limits. First consider $|\epsilon_L - \epsilon_R| \gg I$ then the energies are approximately

$$E = \epsilon_{L,R} + O\left(\frac{I^2}{|\epsilon_L - \epsilon_R|^2}\right) \quad (2.14)$$

On the other hand when $|\epsilon_L - \epsilon_R| \ll I$ then the energies become

$$E = \frac{\epsilon_L + \epsilon_R}{2} \pm |I| + O\left(\frac{|\epsilon_L - \epsilon_R|^2}{I^2}\right) \quad (2.15)$$

In both cases we always have $E_1 - E_0 \geq 2|I|$. So even if the energies ϵ_L and ϵ_R crosses at some point, the above two states will never cross, and always separated by energy at least $2|I|$. This is called the Wigner-von Neumann noncrossing rule.

Suppose we have a spin 1/2 particle in a magnetic field $\mathbf{B} = (B_x, B_z)$. The Hamiltonian is

$$\mathcal{H} = 2\mu_B \mathbf{B} \cdot \mathbf{S} = 2\mu_B (B_x S_x + B_z S_z) = \hbar\mu_B \begin{pmatrix} B_z & B_x \\ B_x & -B_z \end{pmatrix} \quad (2.16)$$

If we compare this Hamiltonian with the one above for asymmetric double well, we find them to be the same form. The correspondence is

$$|\epsilon_L - \epsilon_R| \longleftrightarrow 2\hbar\mu_B B_z, \quad I \longleftrightarrow \hbar\mu_B B_x \quad (2.17)$$

The states are in 1-1 correspondence as the above case. We know that this corresponds to spin precession around the magnetic field, which is in analogue with the oscillation between the two potential wells. These two phenomena are mathematically the same.

3 Lecture 3

Remember last time we were considering two problems. One is a double potential well, and the other is a spin 1/2 particle in a magnetic field. The Hamiltonian for the later is

$$\mathcal{H} = 2\mu_B \mathbf{B} \cdot \mathbf{S} = \hbar\mu_B \begin{pmatrix} B_z & B_x \\ B_x & -B_z \end{pmatrix} \quad (3.1)$$

where $\mathbf{B} = (B_x, B_z)$. Written in matrix form, the former Hamiltonian is

$$\mathcal{H} = \begin{pmatrix} \epsilon_L & I \\ I & \epsilon_R \end{pmatrix} \quad (3.2)$$

The correspondence between the states of these two problems can be summarized in table 1. It is obvious

| Double Well | Spin 1/2 |
|--|---|
| $ L\rangle, R\rangle$ | $ z, +\rangle, z, -\rangle$ |
| $ S\rangle, A\rangle$ | $ x, +\rangle, x, -\rangle$ |
| $\epsilon_{L,R}$ | $\pm \hbar\mu_B B_z$ |
| $\Delta = \epsilon_L - \epsilon_R, I$ | $2\hbar\mu_B B_z, \hbar\mu_B B_x$ |
| $\Delta \gg I, 0\rangle \approx R\rangle, 1\rangle \approx L\rangle$ | $B_z \gg B_x, 0\rangle \approx z, +\rangle, 1\rangle \approx z, -\rangle$ |
| $\hbar\omega = \sqrt{\Delta^2 + 4I^2}$ | $\hbar\omega = 2\hbar\mu_B \mathbf{B} $ |

Table 1: Comparison between tunnelling and spin precession

that these two models share a lot of similarities.

3.1 Parity Selection Rule

In atomic physics, parity can be used to determine the selection rule which tells what transitions are allowed and what are forbidden. Suppose there is a collection of states $|\alpha\rangle, |\beta\rangle, \dots$ which are eigenstates of the parity operator $\hat{\Pi}$, with eigenvalues

$$\hat{\Pi}|\alpha\rangle = \eta_\alpha |\alpha\rangle, \quad \hat{\Pi}|\beta\rangle = \eta_\beta |\beta\rangle \quad (3.3)$$

and we know that the eigenvalues must be ± 1 .

Now suppose we are studying radiative transitions. We are often interested in the matrix element $\langle \alpha | \mathbf{x} | \beta \rangle$. We have the condition

$$\langle \alpha | \mathbf{x} | \beta \rangle = 0, \quad \text{unless } \eta_\alpha = -\eta_\beta \quad (3.4)$$

The proof of this condition is simple. We can insert two identities in the bracket and get

$$\langle \alpha | \mathbf{x} | \beta \rangle = \langle \alpha | \hat{\Pi}^{-1} \hat{\Pi} \mathbf{x} \hat{\Pi}^{-1} \hat{\Pi} | \beta \rangle = -\eta_\beta \eta_\alpha \langle \alpha | \mathbf{x} | \beta \rangle \quad (3.5)$$

So this matrix element vanishes when $\eta_\alpha \eta_\beta = 1$.

We can also prove this using the language of wave functions:

$$\langle \alpha | \mathbf{x} | \beta \rangle = \int \psi_\alpha^*(x) \mathbf{x} \psi_\beta(x) d^3x \quad (3.6)$$

And if we apply parity operation on everything we will get the same result as above. This parity selection rule was known as Laporte's rule historically for radiative transitions. This generalizes to the matrix element of any operator

$$\langle \alpha | A | \beta \rangle = 0 \quad \text{unless} \quad \begin{cases} \eta_\alpha = \eta_\beta, & \text{if } [A, \hat{\Pi}] = 0 \\ \eta_\alpha = -\eta_\beta, & \text{if } \{A, \hat{\Pi}\} = 0 \end{cases} \quad (3.7)$$

Another special case applies to the expectation value of the operator \mathbf{x} . We have

$$\langle \alpha | \mathbf{x} | \alpha \rangle = 0 \quad (3.8)$$

when $|\alpha\rangle$ is an eigenstate of the parity operator. This occurs if $|\alpha\rangle$ is nondegenerate and \mathcal{H} commutes with parity. So under these conditions there is no induced dipole moment.

3.2 Lattice Translations

Now let's consider the symmetry of a lattice structure. Recall when we have a free particle, we have total translation invariance. We can describe this translation invariance by commutation relations

$$[\mathcal{H}, \hat{T}(\mathbf{l})] = 0 \quad (3.9)$$

for any translation vector \mathbf{l} . The solution is nothing but plane waves

$$\psi(\mathbf{x}) = \psi_{\mathbf{p}}(\mathbf{x}) \propto \exp(i\mathbf{p} \cdot \mathbf{x}/\hbar) \quad (3.10)$$

Now suppose we are not living in free space, but in a crystal lattice. This means that we have a periodic potential. If we consider 3D crystal then we have three vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ which characterizes the periodic structure of the 3D lattice. Our potential does not change under a translation of the following vector

$$\mathbf{R}_{n_1, n_2, n_3} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad (3.11)$$

for any set of integers n_1, n_2, n_3 . So the Hamiltonian has the following commutation

$$[\mathcal{H}, \hat{T}(\mathbf{R}_{n_1, n_2, n_3})] = 0 \quad (3.12)$$

We now introduce the Bloch theorem which states that the wavefunctions in the crystal should have the following form

$$\psi(\mathbf{x}) = \psi_{\mathbf{k}, m}(\mathbf{x}) = u_{\mathbf{k}, m}(\mathbf{x}) e^{i\mathbf{k} \cdot \mathbf{x}} \quad (3.13)$$

where $u(\mathbf{x})$ is a periodic function symmetric under the above translations $\mathbf{R}_{n_1, n_2, n_3}$. Now because the u function is periodic, we can expand it using Fourier series. But this will give rise to some ambiguities because the lattice is symmetric, so when we add the momenta we will oversum. We can introduce an inverse lattice which is periodic in the momentum space with basis vectors \mathbf{b}_j satisfying

$$\mathbf{b}_j \cdot \mathbf{a}_i = 2\pi \delta_{ij} \quad (3.14)$$

We can write \mathbf{b} explicitly as

$$\mathbf{b}_1 = 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} \quad (3.15)$$

and similar for the other two vectors. Then we require the momentum \mathbf{k} must lie in a Brillouin zone. The Brillouin zone is a unit cell of the inverse lattice. We call this momentum the quasimomentum, and the energy of m -th eigenstate depends on the quasimomentum

$$E_{m, \mathbf{k}} = E_m(\mathbf{k}) \quad (3.16)$$

which is called the dispersion relation.

3.3 Time Reversal

Now we investigate another important discrete symmetry operation, which is time reversal. Intuitively it is easy to understand: just change the direction of time

$$t \longrightarrow -t \quad (3.17)$$

In classical physics we have Newton's equation of motion

$$m\ddot{\mathbf{x}} = -\nabla V(\mathbf{x}) \quad (3.18)$$

What happens if we change the direction of time? The answer is

$$\mathbf{x} \rightarrow \mathbf{x}, \quad \mathbf{v} \rightarrow -\mathbf{v} \quad \ddot{\mathbf{x}} \rightarrow \ddot{\mathbf{x}} \quad (3.19)$$

So the above equation of motion doesn't change. If $\mathbf{x}(t)$ is a classical trajectory, then $\mathbf{x}(-t)$ is also a trajectory. But now if the force depends on velocity, for example the Lorentz force, then $\mathbf{x}(-t)$ is not a trajectory, unless we also change the sign of the magnetic field.

Now in quantum mechanics the equation of motion is the Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V\psi \quad (3.20)$$

This equation is first order in time, so if $\psi(\mathbf{x}, t)$ is a solution, then in general $\psi(\mathbf{x}, -t)$ is not a solution. However, if we take complex conjugation on the whole equation, we can see that $\psi^*(\mathbf{x}, -t)$ is indeed a solution of the conjugated equation. This can be easily understood for stationary states, where

$$\psi(\mathbf{x}, t) = \psi_0(\mathbf{x}) e^{-iE_n t/\hbar} \quad (3.21)$$

Our guess is that this applies to all states under time-reversal transformation.

Let's define the time-reversal operator Θ . For a state $|\alpha\rangle$, the time-reversed state will be $|\tilde{\alpha}\rangle = \Theta |\alpha\rangle$. Now if we make an infinitesimal time evolution on $|\alpha\rangle$, then we have

$$|\alpha, dt\rangle = \left(1 - \frac{i}{\hbar} \mathcal{H} dt\right) |\alpha\rangle \quad (3.22)$$

and the time-reversed state will be

$$|\tilde{\alpha}, dt\rangle = \left(1 - \frac{i}{\hbar} \mathcal{H} dt\right) \Theta |\alpha\rangle \quad (3.23)$$

We want to have $|\tilde{\alpha}, dt\rangle = \Theta |\alpha, dt\rangle$. So the condition we would like to have is

$$-i\mathcal{H}\Theta |\alpha\rangle = i\Theta\mathcal{H} |\alpha\rangle, \quad \{\mathcal{H}, \Theta\} = 0 \quad (3.24)$$

The appearance of anticommutation is not very nice in our formalism. However, there is one problem with the time-reversal operator if we define it this way. It is not linear. Instead we have

$$\Theta(c|\alpha\rangle) = c^* \Theta |\alpha\rangle \quad (3.25)$$

We call this an *antilinear* operator.

Let's see what is the effect on wave functions. For the following states we have

$$|\alpha\rangle = \int d\mathbf{x} \psi_\alpha(\mathbf{x}) |\mathbf{x}\rangle, \quad |\beta\rangle = \int d\mathbf{x} \psi_\beta(\mathbf{x}) |\mathbf{x}\rangle \quad (3.26)$$

Now because of antilinearity we expect

$$|\tilde{\alpha}\rangle = \int d\mathbf{x} \psi_\alpha^* |\mathbf{x}\rangle \quad (3.27)$$

So as a result the matrix element will change as

$$\langle\alpha|\beta\rangle = \int d\mathbf{x} \psi_\alpha^* \psi_\beta = \langle\tilde{\alpha}|\tilde{\beta}\rangle^* = \langle\tilde{\beta}|\tilde{\alpha}\rangle \quad (3.28)$$

4 Lecture 4

4.1 Time Reversal Continued

Remember last time we were discussing the time reversal symmetry, i.e. the symmetry operation which reverses the direction of time

$$t \longrightarrow -t, \quad \psi(x, t) \longrightarrow \psi^*(x, -t) \quad (4.1)$$

We introduced the operator Θ to invert the time and change a state into a time-reversed state

$$|\alpha\rangle \longrightarrow \Theta |\alpha\rangle = |\tilde{\alpha}\rangle \quad (4.2)$$

Recall that the time-reversal operator differs from usual symmetry operators in a subtle way. Usually a symmetry operator is a unitary operator

$$\hat{S}^\dagger \hat{S} = 1 \quad (4.3)$$

which is necessary to preserve inner products between states. However we showed last time that under time reversal we have instead

$$\langle \alpha | \beta \rangle \longrightarrow \langle \alpha | \Theta^\dagger \Theta | \beta \rangle = \langle \beta | \alpha \rangle \quad (4.4)$$

An operator satisfying this relation for any states $|\alpha\rangle$ and $|\beta\rangle$ is called an *antiunitary* operator. The Θ operator is an example of an antiunitary operator. It has the property that

$$\Theta (c_1 |\alpha_1\rangle + c_2 |\alpha_2\rangle) = c_1^* \Theta |\alpha_1\rangle + c_2^* \Theta |\alpha_2\rangle \quad (4.5)$$

We also call this property antilinearity.

We showed last time that $\{\mathcal{H}, \Theta\} = 0$. But that was not quite correct. Now that taking into the effect of antilinearity, we should write

$$-i\mathcal{H}\Theta = \Theta i\mathcal{H} \quad (4.6)$$

But because of antilinearity we can't cancel the i 's from both sides and conclude the anticommutator of the two operators is zero. Instead we should take the i on the right hand side and put it in front of the Θ operator with a minus sign and get

$$[\Theta, \mathcal{H}] = 0 \quad (4.7)$$

Now that we have this relation, we can study the time evolution operator which is just the exponential of the Hamiltonian

$$|\alpha, t\rangle = U(t) |\alpha\rangle, \quad |\tilde{\alpha}, t\rangle = U(t) \Theta |\alpha\rangle = \Theta U(-t) |\alpha\rangle \quad (4.8)$$

However the wave function will become

$$\psi_{\tilde{\alpha}}(x, t) = \langle x | \tilde{\alpha}, t \rangle = \langle x | \Theta U(-t) | \alpha \rangle = \langle \alpha, -t | x \rangle = \psi_{\alpha}^*(x, -t) \quad (4.9)$$

Consider arbitrary linear operator O . Any matrix element will have the following

$$\langle \beta | O | \alpha \rangle = \langle \gamma | \alpha \rangle = \langle \tilde{\alpha} | \tilde{\gamma} \rangle = \langle \tilde{\alpha} | \Theta O^\dagger | \beta \rangle = \langle \tilde{\alpha} | \Theta O^\dagger \Theta^\dagger | \tilde{\beta} \rangle \quad (4.10)$$

So for a Hermitian operator the O^\dagger in the bracket will become just O . Let's define for any operator

$$A \quad \begin{cases} \text{is even if } \Theta A \Theta^\dagger = A \\ \text{is odd if } \Theta A \Theta^\dagger = -A \end{cases} \quad (4.11)$$

And any operator can be written as a linear combination of even and odd operators.

Let's consider some examples. First consider the momentum operator \mathbf{p} . We have

$$\langle \mathbf{p} \rangle_{\alpha} = \langle \alpha | \mathbf{p} | \alpha \rangle = - \langle \mathbf{p} \rangle_{\bar{\alpha}} \quad (4.12)$$

So \mathbf{p} is an odd operator. The eigenstates of the momentum operator transform as

$$\mathbf{p} |p_1\rangle = p_1 |p_1\rangle \implies \mathbf{p}\Theta |p_1\rangle = -\Theta\mathbf{p}\Theta^\dagger |p_1\rangle = -p_1\Theta |p_1\rangle \quad (4.13)$$

So we have $|\tilde{p}_1\rangle = |-p_1\rangle$. In contrast we know the position operator is invariant under time reversal, so \mathbf{x} is even and $|x\rangle = |\tilde{x}\rangle$. Now for a consistency check we can work out the transformation for the commutator

$$\Theta[x_i, p_j]\Theta^\dagger = -[x_i, p_j] = -i\hbar\delta_{ij} \quad (4.14)$$

The minus sign comes from the fact that p is odd. But this also makes sense because the result is purely imaginary.

Now let's consider angular momentum. Because of the definition $\mathbf{L} = \mathbf{x} \times \mathbf{p}$ we expect \mathbf{L} to be odd. This generalizes to all angular momenta. It can be also seen from the commutation relation between angular momenta

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k \quad (4.15)$$

So we should expect that J transforms just as i does, which changes sign under time reversal. For angular wave functions we have the spherical harmonics Y_l^m . Because $Y_l^m = e^{im\varphi}P_l^m$, and that $P_l^m = (-1)^m P_l^{-m}$, so we expect

$$(Y_l^m)^* = Y_l^{-m}(-1)^m \quad (4.16)$$

Now what about spin? Because S is odd, if \mathbf{B} were even then $\mathbf{B} \cdot \mathbf{S}$ is odd, but this is the energy of a spin coupled to a magnetic field, so we have a contradiction. So we expect \mathbf{B} to be odd. We also have this result in classical physics, as discussed in the last lecture. Now let's consider the operator $\mathbf{S} \cdot \hat{\mathbf{n}}$. It has eigenstates

$$|\hat{\mathbf{n}}, +\rangle = e^{-iS_z\varphi/\hbar}e^{-iS_y\theta/\hbar}|z, +\rangle, \quad |\hat{\mathbf{n}}, -\rangle = e^{-iS_z\varphi/\hbar}e^{-iS_y(\theta+\pi)/\hbar}|z, +\rangle \quad (4.17)$$

Because we know \mathbf{S} is odd, we expect that

$$\Theta |\hat{\mathbf{n}}, +\rangle = e^{i\eta} |\hat{\mathbf{n}}, -\rangle \quad (4.18)$$

where η is an arbitrary phase factor. In light of the above, we can write the time reversal operator as

$$\Theta = e^{i\eta}e^{-iS_y\pi/\hbar}K \quad (4.19)$$

where K is the complex conjugation operator. Now let's define this operator. It must be antilinear so $K(c|\alpha\rangle) = c^*K|\alpha\rangle$. If we expand the state in $|\beta_i\rangle$ basis, then we will get

$$K|\alpha\rangle = K\sum_i \langle b_i|\alpha\rangle |b_i\rangle = \sum_i \langle \alpha|b_i\rangle K|b_i\rangle \quad (4.20)$$

We choose a basis that $K|b_i\rangle = |b_i\rangle$, then we can define the complex conjugation operator using this expression

$$K|\alpha\rangle = \sum_i \langle \alpha|b_i\rangle |b_i\rangle \quad (4.21)$$

Note that this definition is basis-dependent. But note that our definition of the time reversal operator is also basis-dependent, so the dependency actually cancels and there is no problem. In general we can always write an antiunitary operator into the product of a unitary operator with K :

$$A = UK \quad (4.22)$$

Remember that

$$e^{-iS_y\pi/\hbar} = -i\sigma_y \quad (4.23)$$

So the above representation for the time-reversal operator is actually

$$\Theta = e^{i\eta} (-i\sigma_y) K \quad (4.24)$$

This is the correct representation of time reversal operator in the spin-1/2 basis. Now it is interesting to evaluate the square of the time reversal operator

$$\Theta^2 = (-ie^{i\eta}\sigma_y K) (-ie^{i\eta}\sigma_y K) = \sigma_y K \sigma_y K = -\sigma_y \sigma_y K^2 = -1 \quad (4.25)$$

In the second last equality we have used the fact that $\sigma_y K = -K \sigma_y$, because σ_y is a purely imaginary matrix. Note that this says that for spin 1/2 particle if we reverse the time twice, we don't come back to the same state, but pick up a minus sign. This is a profound and counter-intuitive result. The generalization of this result is that for arbitrary j the action of Θ^2 on the eigenstate will be

$$\Theta^2 = (-1)^{2j} \quad (4.26)$$

This can be seen if we consider the spin as consisting of $2j$ spin-1/2 particles.

There is the so-called Kramers degeneracy which states the follows. Assume that \mathcal{H} commutes with Θ , which means that the system is time-reversal invariant. Let $|n\rangle$ be an eigenstate of the Hamiltonian, then $\Theta|n\rangle$ is an eigenstate with the same energy. This is the analogue of the similar result in classical mechanics where time-reversed trajectory is also a solution of the equation of motion. Now the question is: can $\Theta|n\rangle = e^{i\delta}|n\rangle$? The answer is no. Assume this statement is true, then we should have

$$\Theta^2|n\rangle = \Theta(e^{i\delta}|n\rangle) = e^{-i\delta}\Theta|n\rangle = |n\rangle \quad (4.27)$$

But we know this is not true for spin-1/2 particles. So this is a contradiction. For spin-1/2 particles these two states are different states, so there is a 2-fold degeneracy.

Let's consider the spin-orbit coupling

$$\mathcal{H}_{SO} = \lambda_{SO} \mathbf{L} \cdot \mathbf{S} \quad (4.28)$$

Because both angular momenta are odd, the Hamiltonian is even and symmetric under time-reversal. When there is no spin-orbit coupling, we know that the hydrogen atom has $2s + 1 = 2$ degeneracy, but with spin-orbit coupling there is still 2-fold degeneracy because of the Kramer's degeneracy.

5 Lecture 5

5.1 Time Reversal Symmetry Continued

Remember that we were discussing the time-reversal symmetry. Because the Schrödinger equation is essentially a diffusion equation in imaginary time, a sign change of time can be compensated by complex conjugation. We defined the time-reversal operator $\hat{\Theta}$ and under its operation the wave function transforms as

$$\psi = \langle x|\alpha\rangle \longrightarrow \langle x|\tilde{\alpha}\rangle = \langle \alpha|x\rangle = \psi^*(x) \quad (5.1)$$

Remember that the time-reversal operator is antilinear so we have the above property. We classified the observables as even and odd according to their transformation property under time reversal. We found that \mathbf{x} is even whereas \mathbf{p} and \mathbf{J} are odd. We would have expected that $\hat{\Theta} = 1$, but that turned out to only be true for spinless or integer spin particles, and instead we have $\hat{\Theta} = -1$ for half-integer particles. In general

$$\hat{\Theta}^2 = (-1)^{2j} \quad (5.2)$$

For spin 1/2 particles we can represent the time-reversal operator as

$$\hat{\Theta} = e^{i\eta} \mathcal{D}_y(\pi) K \quad (5.3)$$

The K is the complex conjugation operator which we defined last time and \mathcal{D} is the rotation operator. We considered spin-1/2 states using the $|z, \pm\rangle$ basis and expanded arbitrary spin states in terms of these vectors. Remember we have

$$|x, \pm\rangle = \frac{1}{\sqrt{2}} (|z, +\rangle \pm |z, -\rangle), \quad |y, \pm\rangle = \frac{1}{\sqrt{2}} (|z, +\rangle \pm i|z, -\rangle) \quad (5.4)$$

The rotation by π operator operates on these states according to

$$\mathcal{D}_y(\pi) |z, \pm\rangle = |z, \mp\rangle, \quad \mathcal{D}_y(\pi) |x, \pm\rangle = |x, \mp\rangle, \quad \mathcal{D}_y(\pi) |y, \pm\rangle = |y, \pm\rangle \quad (5.5)$$

But the complex conjugation operates non-trivially on the y eigenstates

$$K |z, \pm\rangle = |z, \pm\rangle, \quad K |x, \pm\rangle = |x, \pm\rangle, \quad K |y, \pm\rangle = |y, \mp\rangle \quad (5.6)$$

So for an arbitrary spinor we have

$$\chi = \begin{pmatrix} \psi_{\uparrow} \\ \psi_{\downarrow} \end{pmatrix}, \quad \hat{\Theta}\chi = -e^{i\eta} \begin{pmatrix} \psi_{\downarrow}^* \\ -\psi_{\uparrow}^* \end{pmatrix}, \quad \hat{\Theta}^2\chi = \begin{pmatrix} -\psi_{\uparrow} \\ -\psi_{\downarrow} \end{pmatrix} \quad (5.7)$$

So we can see explicitly that $\hat{\Theta}^2 = -1$ for a spin-1/2 particle. This leads to the Kramers degeneracy when $[\mathcal{H}, \hat{\Theta}] = 0$ that there is always a twofold degeneracy associated with time-reversal.

Let's spell out explicitly one example where space and spin are decoupled:

$$\chi = \psi(\mathbf{x}) \begin{pmatrix} a \\ b \end{pmatrix} \quad (5.8)$$

where $|a|^2 + |b|^2 = 1$. We know that there is a twofold degeneracy because $E_{\uparrow} = E_{\downarrow}$. However if we add a magnetic field, there will be a splitting

$$E_{\uparrow} - E_{\downarrow} = \hbar\mu B \quad (5.9)$$

This is because \mathbf{B} field violates the time-reversal symmetry. However if we only add in spin-orbit coupling, then time-reversal is not violated, so there will be Kramers degeneracy. However now S_z is no longer a good quantum number, so we can't say up or down states. The only thing we know is that there are two states

$$|1\rangle \propto \hat{\Theta} |2\rangle \quad (5.10)$$

When can this spin-orbit coupling arise? Let's say the particle is moving in a potential $V(\mathbf{x})$ for example the Coulomb potential $V(r) = Ze^2/r$. Then there will be an electric field $\mathbf{E} = -(1/e)\nabla V(\mathbf{x})$. In the particle's rest frame it sees a time-dependent electric field, and this time varying electric field will induce a magnetic field. In this magnetic field the spin will rotate, and this is really a relativistic effect. However this magnetic field does not violate time reversal because it is reversed when we reverse the time. Now if $\mathbf{L} = 0$ then the spin rotation will cancel because the net magnetic field averaged over the orbit will add up to zero. However this is not the case for $\mathbf{L} \neq 0$ so this looks like a spin-orbit coupling.

5.2 Approximation Methods

There are only a few problems in quantum mechanics that we can exactly solve. So the approximation methods are invaluable in situations where we can't solve for the exact energy spectrum and eigenfunctions. First we will discuss perturbation theory.

5.3 Perturbation Theory

Suppose we have a Hamiltonian which can be separated into two parts

$$\mathcal{H} = \mathcal{H}_0 + \lambda V \quad (5.11)$$

where \mathcal{H}_0 is a Hamiltonian which we can exactly solve and know everything about it. In particular we know the eigenstates to be $|n\rangle_0$ and eigenvalues $E_n^{(0)}$. We can evaluate the matrix element of any observable in this unperturbed basis. The parameter λ is a parameter of the perturbation theory and in the limit $\lambda \rightarrow 0$ we recover the unperturbed theory. We need λ to be small for perturbation theory to work, because we want to use the unperturbed eigenstates and eigenvalues to approximate those of the full Hamiltonian. We want to ask what are the eigenstates and eigenenergies such that

$$\mathcal{H} |n\rangle = E_n |n\rangle \quad (5.12)$$

We want to make λ as small as possible so that we can do Taylor expansions around $\lambda = 0$. In order to do the expansion, let's rewrite the above Schrödinger equation a little bit

$$(\mathcal{H}_0 + \lambda V) |n\rangle = E_n |n\rangle \quad (5.13)$$

Let's introduce $\Delta_n = E_n - E_n^{(0)}$ which vanish when $\lambda \rightarrow 0$. We can write the above equation as

$$(\lambda V - \Delta_n) |n\rangle = (E_n^{(0)} - \mathcal{H}_0) |n\rangle \quad (5.14)$$

If we only want the first order approximation we can just take $|n\rangle = |n\rangle_0$ on the left hand side, and invert the operator on the right hand side to get an explicit form of the state $|n\rangle$. Now let's define first the inverse operator. Note that the inverse is actually ill defined for $|n\rangle_0$

$$(E_n^{(0)} - \mathcal{H}_0)^{-1} |n\rangle_0 =? \quad (5.15)$$

To resolve this we introduce the projection operator

$$\Phi_n = 1 - |n\rangle_0 \langle n| = \sum_{m \neq n} |m\rangle_0 \langle m| \quad (5.16)$$

Now we want to operate this operator onto the above equation. First we see that

$${}_0\langle n | (\lambda V - \Delta_n) | n \rangle = 0 \quad (5.17)$$

So we actually have

$$(\lambda V - \Delta_n) | n \rangle = \Phi_n (\lambda V - \Delta_n) | n \rangle \quad (5.18)$$

It is now safe to define the inverse operator so we can formally write the result

$$\frac{1}{E_n^{(0)} - \mathcal{H}_0} \Phi_n (\lambda V - \Delta_n) | n \rangle = | n \rangle \quad (5.19)$$

We can call the whole operator in front by

$$\frac{\Phi_n}{E_n^{(0)} - \mathcal{H}_0} = G_n \quad (5.20)$$

Now there is a problem for the above equation. In the limit $\lambda \rightarrow 0$ the above equation gives us $|n\rangle = 0$ instead of $|n\rangle = |n\rangle_0$. This is because the above is just a particular solution. We need to add to that the solution to the homogeneous equation which are the unperturbed states, so we should have

$$|n\rangle = G_n (\lambda V - \Delta_n) |n\rangle + C_n(\lambda) |n\rangle_0 \quad (5.21)$$

We require $C_n(\lambda) = 1$ for any λ . Note that from this equation $\langle n|n\rangle \neq 1$, but this is not a big problem as we can always renormalize the states. We always have

$$\langle n|n\rangle_0 = 1 \quad (5.22)$$

Combined with an equation above we have

$$\Delta_n = \lambda {}_0\langle n | V | n \rangle \quad (5.23)$$

And the other equation we have is

$$|n\rangle = |n\rangle_0 + G_n (\lambda V - \Delta_n) |n\rangle \quad (5.24)$$

We look for solutions of the above equations of the following form

$$|n\rangle = |n\rangle_0 + \lambda |n\rangle_1 + \lambda^2 |n\rangle_2 + \dots \quad (5.25)$$

and

$$\Delta_n = \lambda \Delta_n^{(1)} + \lambda^2 \Delta_n^{(2)} + \dots \quad (5.26)$$

We do this by iterations. The first order correction is just

$$\Delta_n^{(1)} = {}_0\langle n | V | n \rangle_0 = V_{nn} \quad (5.27)$$

which is just the expectation value of the perturbation with respect to the unperturbed states. In general the higher order corrections are

$$\Delta_n^{(k)} = {}_0\langle n | V | n \rangle_{k-1} \quad (5.28)$$

This means that we can't just solve for the corrections to the eigenvalues but also the corrections to the eigenstates. So that is what we will do in the next step:

$$|n\rangle_1 = G_n V |n\rangle_0 \quad (5.29)$$

Therefore we can write the form of Δ_2 explicitly

$$\begin{aligned} \Delta_2 &= {}_0\langle n | V G_n V | n \rangle_0 = \sum_{m,m'} {}_0\langle n | V | m \rangle_0 {}_0\langle m | G_n | m' \rangle_0 {}_0\langle m' | V | n \rangle_0 \\ &= \sum_{m \neq n} \frac{|V_{mn}|^2}{E_n^{(0)} - E_m^{(0)}} \end{aligned} \quad (5.30)$$

6 Lecture 6

6.1 Time-independent Perturbation Theory Continued

What we have been considering is time independent perturbation theory. We have only been considering nondegenerate case, where the energy spectrum has been assumed to be nondegenerate. Remember we separated the Hamiltonian into two parts

$$\mathcal{H} = \mathcal{H}_0 + \lambda V \quad (6.1)$$

Where \mathcal{H}_0 is a Hamiltonian which we know the full spectrum and V is a small perturbation. λ is a parameter which we will take to 1 in the end of calculation.

Remember we assumed the eigenstates and eigenvalues of the unperturbed Hamiltonian are

$$\mathcal{H}_0 |n\rangle_0 = E_n^{(0)} |n\rangle_0 \quad (6.2)$$

and we want to know the eigenstates and eigenvalues of the full Hamiltonian

$$\mathcal{H} |n\rangle = E_n |n\rangle \quad (6.3)$$

The way we do this is to expand everything in power series of the parameter λ and collect terms with the same power of λ . We obtained the result

$$\Delta_n = E_n - E_n^{(0)} = \lambda \langle n | V | n \rangle \quad (6.4)$$

which is the energy correction due to the perturbation. Note this result is exact, and the state satisfy the following equation

$$|n\rangle = |n\rangle_0 + \frac{\Phi_n (\lambda V - \Delta_n)}{E_n^{(0)} - \mathcal{H}_0} |n\rangle \quad (6.5)$$

And this equation is exact, too. Now approximation comes in when we try to solve this equation iteratively. Plugging in the power series expansion in terms of λ , we obtained last time the following result

$$\Delta_n^{(1)} = \lambda \langle n | V | n \rangle_0 = V_{nn} \quad (6.6)$$

$$\Delta_n^{(2)} = \lambda^2 \sum_{m \neq n} \frac{|V_{mn}|^2}{E_n^{(0)} - E_m^{(0)}} \quad (6.7)$$

Note the form of the second order perturbation. Because the ground state has minimal energy, so we can see from the form that the second order correction to the ground state energy is negative in general, independent of the specific form of the perturbation itself.

The perturbed energy eigenstates were obtained as

$$|n\rangle = |n\rangle_0 + \lambda \sum_{m \neq n} \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} |m\rangle_0 + \lambda^2 \left[\sum_{m \neq n} |m\rangle_0 \left(-\frac{V_{nn}V_{mn}}{(E_n^{(0)} - E_m^{(0)})^2} + \sum_{l \neq n} \frac{V_{ml}V_{ln}}{(E_n^{(0)} - E_l^{(0)})(E_n^{(0)} - E_m^{(0)})} \right) \right] + \dots \quad (6.8)$$

Note that the energy denominators never vanish for nondegenerate unperturbed energy spectrum. In general these results are valid when

$$|V_{mn}| \ll |E_n - E_m| \quad (6.9)$$

Although a bit rough, this can be taken as a criterion for the validity of perturbation theory, and it has a nice interpretation that the perturbation is small relative to the energy level difference.

6.2 Normalization of States

Recall that in our adopted notation we have

$$\langle n|n\rangle \neq 1, \quad \langle n|n\rangle_0 = 1 \quad (6.10)$$

We want to normalize the state $|n\rangle$. We define the following state

$$\left(|n\rangle_0 + \lambda |n\rangle_1 + \dots + \lambda^k |n\rangle_k \right) = \frac{1}{\sqrt{Z_{n,k}}} |n\rangle^{(k)} \quad (6.11)$$

and we require ${}^{(k)}\langle n|n\rangle^{(k)} = 1$. Under this condition we can work out the renormalization coefficients

$$\begin{aligned} \frac{1}{Z_{n,k}} &= ({}_0\langle n| + \lambda {}_1\langle n| + \dots) (|n\rangle_0 + \lambda |n\rangle_1 + \dots) \\ &= 1 + \lambda^2 \sum_{m \neq n} \frac{|V_{mn}|^2}{(E_n^{(0)} - E_m^{(0)})^2} + O(\lambda^3) \end{aligned} \quad (6.12)$$

So we can invert it and get

$$Z_n = 1 - \lambda^2 \sum_{m \neq n} \frac{|V_{mn}|^2}{(E_n^{(0)} - E_m^{(0)})^2} + O(\lambda^3) = \frac{\partial E_n}{\partial E_n^{(0)}} \quad (6.13)$$

6.3 Example

Let's consider an explicit example where we know everything. We consider a two-level system where

$$\mathcal{H} = \begin{pmatrix} \epsilon_1 & \lambda I \\ \lambda I & \epsilon_2 \end{pmatrix} = \begin{pmatrix} \epsilon_1 & 0 \\ 0 & \epsilon_2 \end{pmatrix} + \lambda \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \quad (6.14)$$

We assume $\epsilon_1 < \epsilon_2$ without loss and generality. The two energy eigenstates of the unperturbed Hamiltonian are

$$\mathcal{H}_0 |0\rangle_0 = \epsilon_1 |0\rangle_0, \quad \mathcal{H}_0 |1\rangle_0 = \epsilon_2 |1\rangle_0 \quad (6.15)$$

Plug in the above general formulae, we can get the perturbed energies

$$E_0 = E_0^{(0)} + \lambda^2 \frac{|V_{01}|^2}{E_0^{(0)} - E_1^{(0)}} = \epsilon_1 + \frac{\lambda^2 I^2}{\epsilon_1 - \epsilon_2}, \quad E_1 = \epsilon_2 + \frac{\lambda^2 I^2}{\epsilon_2 - \epsilon_1} \quad (6.16)$$

Remember the exact solution for this two-level system by diagonalizing the Hamiltonian is

$$E_{0,1} = \frac{\epsilon_1 + \epsilon_2}{2} \mp \sqrt{\frac{(\epsilon_1 - \epsilon_2)^2}{4} + \lambda^2 I^2} \quad (6.17)$$

It can be easily checked that in the limit $\lambda I \ll |\epsilon_1 - \epsilon_2|/2$ the exact solution will become approximately our perturbative solution as written above.

Using perturbation theory we can also get the expression for the states

$$|0\rangle = |0\rangle_0 + \frac{\lambda I}{\epsilon_1 - \epsilon_2} |1\rangle_0, \quad |1\rangle = |1\rangle_0 + \frac{\lambda I}{\epsilon_2 - \epsilon_1} |0\rangle_0 \quad (6.18)$$

We can also get the result from the exact result by diagonalizing the Hamiltonian. Again we can see that in the perturbative regime the exact solution reduces to the above solution. The formula also tells us about the normalization coefficient

$$Z_0 = \left(1 + \frac{\lambda^2 I^2}{(\epsilon_1 - \epsilon_2)^2} \right)^{-1} \quad (6.19)$$

6.4 Degenerate Case

Up to now we are considering perturbation theory under the assumption that the unperturbed energy spectrum is nondegenerate. The problem with degeneracy is that the formulae we derived are no longer well-defined, because even for state other than n we can have some state $|n'\rangle$ such that $E_n^{(0)} - E_{n'}^{(0)} = 0$ and this will lead to some terms in the perturbation expansion blowing up.

How do we address this problem? Let's consider a family of states $\{|n_i\rangle\}$ which are degenerate

$$\mathcal{H}_0 |n_i\rangle = E_n |n_i\rangle \quad (6.20)$$

These kets span a degenerate subspace of the whole Hilbert space of the quantum states. Obviously this space is more than 1D, otherwise it would not be degenerate. We can form any eigenstate with energy E_n by linear combination $\sum c_i |n_i\rangle$. Let's try to understand what we can do on this degenerate eigenspace with perturbation theory. Usually under perturbation the spectrum will split and not remain degenerate. Say they split into energy levels $|\ell_i\rangle$, and we require they go back to original energy eigenstates $|\ell_i\rangle_0$ when $\lambda \rightarrow 0$. We can expand the original states in the degenerate eigenspace of the original Hamiltonian

$$|\ell_i\rangle_0 = \sum_j |n_j\rangle_0 \langle n_j | \ell_i \rangle_0 \quad (6.21)$$

So the idea to attack the problem is to choose a set of coefficients such that the corresponding original states also diagonalize the perturbation part of the Hamiltonian, which means

$${}_0\langle \ell_j | V | \ell_{j'} \rangle_0 = 0, \quad \text{when } j \neq j' \quad (6.22)$$

So we need to diagonalize the matrix $V_{jj'}$ and solve the secular equation

$$\text{Det} (V_{jj'} - V \delta_{jj'}) = 0 \quad (6.23)$$

Whenever we have distinct eigenvalues of this matrix, then we have splitting in the perturbed energy spectrum. We can always do this because the unperturbed Hamiltonian is just a multiple of the identity in the degenerate eigenspace \mathcal{D} . The corrected eigenvector should be

$$|\ell_j\rangle = |\ell_j\rangle_0 + \lambda \sum_{m \notin \mathcal{D}} \frac{V_{\ell_j m}}{E_{\ell_j}^{(0)} - E_m^{(0)}} |m\rangle_0 + \dots \quad (6.24)$$

and corrected energy will be

$$E_{\ell_j} = E_{\mathcal{D}}^{(0)} + \lambda V_j + \lambda^2 \sum_{m \notin \mathcal{D}} \frac{|V_{\ell_j m}|^2}{E_{\mathcal{D}}^{(0)} - E_m^{(0)}} + \dots \quad (6.25)$$

Note that the sum is over $m \notin \mathcal{D}$ because for $m \in \mathcal{D}$ we have already diagonalized the matrix $V_{jj'}$ and the contribution is already in the lower order term.

7 Lecture 7

Remember last time we were considering degenerate perturbative theory, where there is a degenerate subspace \mathcal{D} of the unperturbed Hamiltonian with the same energy $E_{\mathcal{D}}^{(0)}$. The recipe we developed last time was to diagonalize the matrix $V_{ij} = \langle n_i | V | n_j \rangle_0$. The diagonal entries, i.e. the eigenvalues of this perturbation matrix give the first order energy shifts. If the eigenvalues are distinct, then there will be splitting of magnitude $\Delta E = \lambda (V_j - V_{j'})$.

7.1 Examples: Band Structure

We want to consider some examples where these rules of perturbation can be applied. First we want to see the band structure in a periodic potential. We want to consider the so-called weak coupling model, which means that the potential $V(x)$ is weak. The Hamiltonian of the problem is

$$\mathcal{H} = \frac{p^2}{2m} + \lambda V(x) \quad (7.1)$$

In order to simplify the discussion we only consider 1 dimensional case. The condition of periodicity means that

$$V(x+a) = V(x) \quad (7.2)$$

for some constant a . First we need to identify the unperturbed state. For free particle we know them very well: they are just momentum eigenstates and the wave function is

$$\psi_p(x) = \langle x | p \rangle = C e^{ipx/\hbar} \quad (7.3)$$

Now we need the matrix elements of the perturbation. They are just

$$V_{pp'} = \int dx |C|^2 e^{ix(p-p')/\hbar} V(x) \quad (7.4)$$

We observe that the matrix elements are just functions of the difference of momenta $p - p'$. Note also that this is just the Fourier transform of the potential $V(x)$. But we know that $V(x)$ is periodic, and the Fourier transform reduces to a Fourier series. So we can write the above in the following form

$$V_q = \sum_{k=-\infty}^{\infty} V_k \delta \left(q - \frac{2\pi\hbar}{a} k \right) \quad (7.5)$$

Remember the unperturbed energy is just the energy of a free particle

$$E_p^{(0)} = \frac{p^2}{2m}, \quad E_{-p}^{(0)} = E_p^{(0)} \quad (7.6)$$

We can calculate the first order shift in energy easily

$$\Delta_p^{(1)} = \tilde{V}_{q=0} = \frac{1}{a} \int_0^a V(x) dx \quad (7.7)$$

which is just the average value of $V(x)$ in one period. This is a constant shift in energy, so it is uninteresting. We are more interested when there is energy splitting. Note that $\tilde{V}_{2p} = 0$ unless $p = \pi\hbar k/a$, so in most cases we have non-degenerate perturbation. In this case we have

$$\Delta_p^{(2)} = \sum_{k \neq 0} 2m \frac{|V_k|^2}{p^2 - (p + 2\pi\hbar k/a)^2} \quad (7.8)$$

However we also want to know the behavior near or at $p = \pi\hbar k/a$. In this case we have nontrivial correction due to degenerate perturbation at first order, and it dominates the second order perturbations. In this case we need to solve the secular equation

$$\text{Det}(V - V_{\pm}\delta) = 0 \quad (7.9)$$

This is just a 2×2 matrix and what we get from diagonalization is just $V_{\pm} = \pm V_k$.

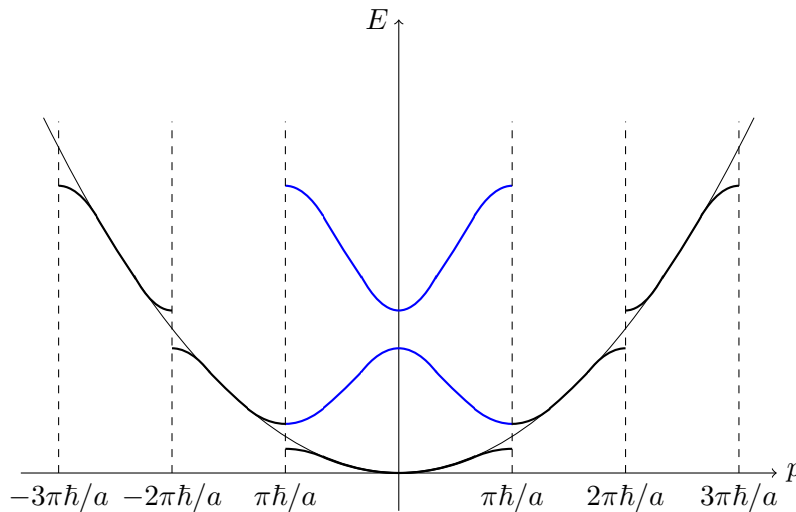


Figure 7.1: Band Structure

Now all the intervals of the momentum can be mapped into the central interval by the quasimomentum and the band number, just as shown in figure 7.1:

$$p = q + \frac{2\pi\hbar}{a}b \quad (7.10)$$

We call the central interval the first Brillouin zone, which is from $-\pi\hbar/a$ to $\pi\hbar/a$. The energy now depends both on the quasimomentum, which characterizes the position in the Brillouin zone, and the band number, which tells which band are we on. The bands are separated by energy gaps, which are bands forbidden by the periodic structure. So the perturbed state can be labelled now as $|n\rangle = |q, b\rangle$ and that

$$|q, b\rangle = \int dp' C(q, b, p') |p'\rangle \quad (7.11)$$

where

$$C(q, b, p') = V_k \delta \left(q + \frac{2\pi\hbar}{a} b - p' - \frac{2\pi\hbar}{a} k \right) \frac{1}{E_p^{(0)} - E_{p'}^{(0)}} \quad (7.12)$$

$$|q, b\rangle = \sum_k \left| p' = q + \frac{2\pi\hbar(b+k)}{a} \right\rangle \frac{2mV_k}{(q + 2\pi\hbar b/a)^2 - (q + 2\pi\hbar(b+k)/a)^2} \quad (7.13)$$

7.2 Examples: Stark Effect

We consider a hydrogen atom in an external electric field. The Hamiltonian is written as

$$\mathcal{H} = \mathcal{H}_0 - ezE \quad (7.14)$$

Now strictly speaking this is not a good potential because towards $z \rightarrow \infty$ the potential drops to $-\infty$ and any state can tunnel to infinity after a sufficient long time, so the system is inherently unstable. However we can estimate the atomic electric field

$$E_a \sim \frac{e^2}{a_B^2} \sim 10^9 \text{ eV/cm} \quad (7.15)$$

So if our external field is on the order of 10 eV/cm, then it literally takes forever for the electron to tunnel out, and it makes sense to talk about perturbation theory on the bound states.

Now we recall our recipe for perturbation theory. We call $V = z$ and $\lambda = -eE$. The eigenstates of the unperturbed Hamiltonian are just hydrogen states $|n, l, m\rangle$ where l and m label orbital angular momentum. In hydrogen we have the nice accidental symmetry which allows the energy to only depend on the principal quantum number $E^{(0)} = E_n^{(0)}$.

Consider first the 1s state $|1, 0, 0\rangle$ which is not degenerate. The energy shift is just

$$E_1 = E_1^{(0)} - eE \langle z \rangle_{11} + e^2 E^2 \sum_{n' > 1} \sum_{l, m} \frac{\langle n', l, m | z | 1, 0, 0 \rangle}{E_1^{(0)} - E_{n'}^{(0)}} + \dots \quad (7.16)$$

Note that by parity symmetry the first order term actually vanishes, $\langle z \rangle_{11} = 0$. So we have the so-called quadratic Stark effect because the first energy shift comes from the quadratic term in E . Note in the quadratic term we also have the selection rule that the matrix element vanishes whenever l is even.

Now the more interesting case is for $n = 2$. We have three different states in $l = 1$ where $m = 1, 0, -1$; and we have one state in $l = 0$ where $m = 0$. So in general there is 4-fold degeneracy. We need to evaluate the matrix element $\langle n, l, m | z | n', l', m' \rangle$ between these four states. Note that this matrix element vanishes unless $m = m'$ and $l = l' \pm 1$. This is a direct consequence of the Wigner-Eckart theorem, but we can also see it if we write out the integral out explicitly in terms of spherical harmonic functions. This result largely simplifies our calculation: only the element $\langle 200 | z | 210 \rangle$ and its conjugate are nonzero. Its value can be calculated directly by means of spherical harmonic integrals

$$\langle 200 | z | 210 \rangle = 3eEa_B \quad (7.17)$$

The eigenstates are just

$$|\pm\rangle = \frac{1}{\sqrt{2}} (|200\rangle \pm |210\rangle) \quad (7.18)$$

So the first order energy shifts are just $\pm 3eEa_B$. Thus the 4 states are split into 3 different energy levels by the external electric field.

How about when we also consider electronic fine structure? The fine structure perturbation splits the accidental symmetry. There are two major contributions for fine structure, one is from relativistic effect and the other is spin-orbit coupling. Spin-orbit coupling splits states with different total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ and states are labelled by $|n, j, m_j\rangle$ where $j = l \pm 1/2$. Now in two cases it is easy, i.e. E is either very weak or very strong. In both limits we can treat one kind of states as original states and apply perturbation theory again on those states. The hardest case is when Stark effect is comparable with the fine structure effect. Then we need to work harder and taking into account all kinds of perturbation at the same time.

8 Lecture 8

We have been doing time-independent perturbation theory last week, including nondegenerate and degenerate cases. We built an iterative prescription to evaluate the perturbed eigenstates and energy eigenvalues from the unperturbed quantities. The way to do it was to expand the respective quantities in powers of the parameter λ , and comparing terms with the same power of λ . Recall if we have degeneracies, then we need to diagonalize the matrix V_{nm} in the degenerate subspace \mathcal{D} , and different eigenvalues will give us splitting between the originally degenerate states.

We discussed the Stark effect in hydrogen atom where $\lambda V = eEz$. The condition for perturbation theory to be applicable is when $E \ll E_0$ where E_0 is the characteristic electric field in the atomic scale, $E_0 = e^2/a_B^2$. We found that the ground state only has quadratic Stark effect. The first excited state has 4-fold degeneracy and we found that they are split into 3 different levels. The energy-shifted states are linear combinations of the $|200\rangle$ and $|210\rangle$ states

$$|\pm\rangle = \frac{1}{\sqrt{2}} (|200\rangle \pm |210\rangle) \quad (8.1)$$

and the first order energy shift is $3eEa_B$.

8.1 Spin-Orbit Coupling

We then went into discussing fine structure of the hydrogen atoms. Remember we have the spin-orbit coupling $\lambda V = \lambda_{SO} \mathbf{L} \cdot \mathbf{S}$, under which the good quantum numbers are changed

$$l, s \longrightarrow j = l \pm \frac{1}{2} \quad (8.2)$$

When $\lambda = 0$ there is only one possible value of j and we can effectively ignore the effect of spin-orbit coupling. But when $l = 1$ we can have $j = 1/2, 3/2$, then there will be a splitting of the energy levels. The energy eigenstates are split into two multiplets $|2P_{1/2}\rangle$ and $|2P_{3/2}\rangle$. There is also another splitting between $|2S_{1/2}\rangle$ and $|2P_{1/2}\rangle$, but this is due to Lamb shift, and one needs to go beyond quantum mechanics to describe it properly.

This can be generalized to hydrogenic atoms like K, Na, Li, ... The Hamiltonian is written as

$$\mathcal{H} = \frac{\mathbf{p}^2}{2m} + V_c(r) + \mathcal{H}_{SO}, \quad \mathcal{H}_{SO} \propto \mathbf{L} \cdot \mathbf{S} \quad (8.3)$$

and the perturbation term may have a prefactor which depends on the radius r . We can in fact approximately calculate the prefactor using techniques of atomic physics:

$$\mathbf{E}_a = -\frac{1}{e} \nabla V_c(r), \quad \mathbf{B}_{\text{eff}} = -\frac{\dot{\mathbf{x}}}{c} \times \mathbf{E}_a, \quad \boldsymbol{\mu} = \frac{e}{m_e c} \mathbf{S} \quad (8.4)$$

And plugging into $\mathcal{H}_{SO} = -\boldsymbol{\mu} \cdot \mathbf{B}_{\text{eff}}$, we have the full expression for the spin-orbit coupling is

$$\mathcal{H}_{SO} = \frac{1}{2m_e^2 c^2} \frac{1}{r} \frac{\partial V_c}{\partial r} \mathbf{L} \cdot \mathbf{S} \quad (8.5)$$

However, what we will consider in the subsequent part of the lecture will only concern the $\mathbf{L} \cdot \mathbf{S}$ part of the perturbation, but the prefactor is not very important.

When there is no spin-orbit coupling, we can write equivalently in two different basis for the eigenstates, i.e. the states of definite $l, s, |l, m, s = 1/2, m_s = \pm 1/2\rangle$, or the states of definite $j, |l, j, s = 1/2, m_j\rangle$. They are connected by Clebsch-Gordan coefficients. Remember we have

$$\mathbf{J} = \mathbf{L} + \mathbf{S}, \quad \mathbf{L} \cdot \mathbf{S} = \frac{1}{2}(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2) \quad (8.6)$$

It is apparent that \mathbf{S}^2 is always going to give $3/4\hbar^2$, and the other terms will give a linear splitting according to the corresponding quantum numbers. The first order energy shifts are just

$$\Delta_{SO}^{(1)} = \frac{1}{2} \langle V_{SO}(r)[\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2] \rangle_{n,l,j} = \frac{1}{2} \left[j(j+1) - l(l+1) - \frac{3}{4} \right] \int_0^\infty r^2 V_{SO}(r) R_{nl}^2(r) dr \quad (8.7)$$

Note that the integral on r is not very important in the energy level splitting qualitatively. It can be checked that the term in the square bracket is

$$\left[j(j+1) - l(l+1) - \frac{3}{4} \right] = \begin{cases} l, & \text{if } j = l + 1/2 \\ -l - 1, & \text{if } j = l - 1/2 \end{cases} \quad (8.8)$$

So the energy splitting between these levels is

$$\Delta E = \frac{2l+1}{2} \int r^2 R_{nl}^2(r) V_{SO}(r) dr \quad (8.9)$$

Consider the Na atom where the outer electron is on $3s$ orbit. Remember $|3s\rangle$ and $|3p\rangle$ states are split due to core effects. The $|3s\rangle$ has no spin-orbit coupling effect, but the 6 states in $|3p\rangle$ multiplet is split according to $j = 1/2$ or $j = 3/2$, which have 2 and 4 states respectively.

Now let's estimate the magnitude of splitting. Now we need to estimate

$$\left\langle \frac{1}{r} \frac{\partial V_c(r)}{\partial r} \right\rangle \sim \frac{e^2}{a_B} \frac{1}{a_B^2} \quad (8.10)$$

So the splitting will become

$$\Delta E \sim \frac{e^2}{a_B^3} \left(\frac{\hbar}{m_e c} \right)^2 = \frac{e^2}{a_B} \left(\frac{e^2}{\hbar c} \right)^2 \quad (8.11)$$

The constant in the last bracket is called the fine structure constant $\alpha = e^2/\hbar c$. This quantity is central to the study of quantum electrodynamics, as this is the dimensionless coupling parameter and measures the strength of interaction.

8.2 The Zeeman Effect

In contrast to the Stark effect, let's consider what will happen in an external magnetic field $\mathbf{B} = B\hat{\mathbf{z}}$. We need to choose a gauge to introduce a vector potential to represent this field. A convenient choice is just

$$\mathbf{A} = \frac{1}{2} \mathbf{B} \times \mathbf{x} \quad (8.12)$$

The next step is just replace \mathbf{p}^2 in the Hamiltonian by $(\mathbf{p} - e\mathbf{A}/c)^2$. The perturbation part of the Hamiltonian can be written as

$$\mathcal{H}_B = -\frac{e}{2m_e c} B L_z + \frac{e^2}{8m_e c^2} B^2 (x^2 + y^2) - \frac{eB}{m_e c} S_z \quad (8.13)$$

We will assume that the magnetic field is small and ignore the second term which is second order in the magnetic field. Now the full Hamiltonian will contain 3 parts:

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{SO} + \mathcal{H}_B, \quad \mathcal{H}_B = -\frac{eB}{2m_e c} (J_z + S_z) \quad (8.14)$$

We are interested in the effect of \mathcal{H}_B as a perturbation over the spin-orbit coupling Hamiltonian. Note that the states $|n, j, l, m\rangle$ are not eigenstates of the spin operator S_z , so we need to use Clebsch-Gordan expansion

$$|n, j, l, m\rangle = \sum C_{j,l,m}^{l,m_l,m_s} |n, l, m_l, m_s\rangle \quad (8.15)$$

And we can then evaluate

$$\langle S_z \rangle_{n,j=l\pm 1/2} = \pm \frac{m_j \hbar}{2l+1} \quad (8.16)$$

We can also just use the projection theorem to reduce the above expectation value to the expectation value of J_z . The shift in energy levels is

$$E_{n,j,l,m}(B) = E_n^{(0)} - B \frac{e\hbar}{2m_e c} m_j \left[1 \pm \frac{1}{2l \pm 1} \right] \quad (8.17)$$

This is also called the Landé shift.

9 Lecture 9

9.1 Van der Waals Interaction

The problem we are going to discuss is as following. We have two hydrogen atoms at a distance R which is much larger than the Bohr radius a_B . By the distance of the atoms we mean the distance between the nuclei, i.e. protons. We measure the positions of electrons from the respective protons with vectors \mathbf{x}_1 and \mathbf{x}_2 .

Let's first assume they are infinitely far away, $R \rightarrow \infty$. In this case our Hamiltonian will just be the sum of two hydrogenic Hamiltonians

$$\mathcal{H}_0 = -\frac{\hbar^2}{2m} (\nabla_1^2 + \nabla_2^2) - \frac{e^2}{r_1} - \frac{e^2}{r_2} \quad (9.1)$$

Now if the distance R is finite we also need to take into account the interaction part of the Hamiltonian. We treat it as a perturbation. The potential is

$$V = \frac{e^2}{R} + \frac{e^2}{|\mathbf{R} + \mathbf{x}_2 - \mathbf{x}_1|} - \frac{e^2}{|\mathbf{R} - \mathbf{x}_1|} - \frac{e^2}{|\mathbf{R} + \mathbf{x}_2|} \quad (9.2)$$

The unperturbed ground state is just the product of two hydrogen wavefunctions

$$\Phi_0^{(0)}(\mathbf{x}_1, \mathbf{x}_2) = \psi_{100}(\mathbf{x}_1)\psi_{100}(\mathbf{x}_2) \quad (9.3)$$

where ψ_{nlm} is the hydrogen wave function. We can now use perturbation expansion in powers of a_B/R . It is apparent that the first order terms in the perturbation potential cancel with each other, so we have, as the first nonvanishing contribution

$$\begin{aligned} V &= \frac{e^2}{R^3} (x_1x_2 + y_1y_2 - 2z_1z_2) + O\left(\frac{a_B^4}{r^4}\right) \\ &= \frac{e^2}{R^3} \left(\mathbf{x}_1 \cdot \mathbf{x}_2 - 3 \frac{(\mathbf{x}_1 \cdot \mathbf{R})(\mathbf{x}_2 \cdot \mathbf{R})}{R^2} \right) + O\left(\frac{a_B^4}{r^4}\right) \end{aligned} \quad (9.4)$$

where we have chosen $\mathbf{R} = R\hat{\mathbf{z}}$ and the terms in the second line are obviously dipole interactions. This is just the first nonvanishing interaction in the multipole expansion of the interaction potential. Note that the first term in the expansion is just the total charge and the two atoms are neutral, therefore that interaction vanishes. The multipole expansion of the potential due to a localized charge distribution is

$$V(\mathbf{r}) = \frac{Q}{r} + \frac{\mathbf{d} \cdot \mathbf{r}}{r^3} + \sum_{\alpha, \beta} \frac{Q_{\alpha\beta} (r_\alpha r_\beta - \delta_{\alpha\beta} r^2)}{r^5} + \dots \quad (9.5)$$

which should be familiar from classical electrodynamics.

In the ground state of the hydrogen atom, all of the multipole moments vanish due to the symmetry of the configuration. If we take the hydrogen wave function to be a charge distribution, then the total charge of the atom falls exponentially with r , so does the electric field. Therefore, the first order correction to the energy vanishes

$$\langle V \rangle_0 = 0 \quad (9.6)$$

because there is no dipole interaction between two hydrogen atoms in the $1s$ configuration. So we must go to second order in perturbation theory

$$E_0(R) = E_0^{(0)} + \frac{e^4}{r^6} \sum_{\alpha} \frac{|\langle 0 | \mathbf{x}_1 \mathbf{x}_2 - 3(\mathbf{x}_1 \cdot \mathbf{R})(\mathbf{x}_2 \cdot \mathbf{R})/R^2 | \alpha \rangle_0|^2}{E_0^{(0)} - E_{\alpha}^{(0)}} \quad (9.7)$$

The index α in the above equation is a collective index to denote the configurations of the two hydrogen atoms. The fact that the second order energy shift is negative means that there is attraction between the two atoms. Therefore, even though the electric field due to individual hydrogen atom in the ground state falls exponentially, the attractive potential between the two atoms goes like $1/r^6$. This interaction is called the Van der Waals force.

Now let's estimate the magnitude of this force. Note that the leading contribution in the summation is from the first excited state because $E_{n>1} > E_1$. So we can get an upper bound

$$|V(R)| < \frac{e^4}{r^6 (E_0^{(0)} - E_1^{(0)})} \sum_{\alpha} |\langle 0 | \mathbf{x}_1 \mathbf{x}_2 - 3(\mathbf{x}_1 \cdot \mathbf{R})(\mathbf{x}_2 \cdot \mathbf{R})/R^2 | \alpha \rangle_0|^2 \quad (9.8)$$

And the sum of the square of the expectation values is just a number.

9.2 Polarization

If we are naïve and take the hydrogen atom as a point-like central proton surrounded by a spherically symmetric electron gas. If we measure the dipole moment of this system it must vanish. However, if we introduce an external electric field, then the potential will cease to be spherically symmetric and the electron gas will be distorted to create a dipole moment \mathbf{d} . This phenomenon is called polarization and we have

$$\mathbf{d} = P\mathbf{E} \quad (9.9)$$

where the coefficient P is called the polarizability of the atom. The energy of a dipole in an external electric field is just

$$E = -\mathbf{E} \cdot \mathbf{d} + \frac{\mathbf{d}^2}{2P} \quad (9.10)$$

The first term is just the dipole interacting with the electric field, and the second term is a measurement of the difficulty of creating the dipole moment. The minimum of the energy is at

$$E_{\min} = -\frac{1}{2} \frac{\mathbf{d}^2}{P} = -\frac{1}{2} \mathbf{E} \cdot \mathbf{d} \quad (9.11)$$

Now we consider it in quantum mechanics. For nondegenerate states, this problem is just the quadratic Stark effect. We can get the expectation for the dipole moment by

$$\langle \mathbf{d}_n \rangle = \frac{\partial E_n}{\partial \mathbf{E}} = \mathbf{E} \cdot 2e^2 \sum_{m \neq n} \frac{|\langle n | z | m \rangle|^2}{E_n^{(0)} - E_m^{(0)}} \quad (9.12)$$

And the term multiplying \mathbf{E} is just the polarizability of the atom. We know from parity selection rule that the matrix element z_{mn} vanishes unless $|n\rangle$ and $|m\rangle$ have different parities.

However, we also have the linear Stark effect. We would find

$$\langle \mathbf{d}_n \rangle = \frac{\partial E_n}{\partial \mathbf{E}} \neq 0 \quad \text{at } \mathbf{E} = 0 \quad (9.13)$$

This seems to contradict with our intuition that the hydrogen atom is spherically symmetric. Let's recall when do we have linear Stark effect: we have it when there is degeneracy on the same m . If we have two states with the same energy and different parity, then the linear combination of them will have no definite parity, and this is what creates the nonvanishing dipole moment.

To illustrate this effect, we consider again the double potential well. Remember we have the symmetric and antisymmetric states

$$|S\rangle = \frac{1}{\sqrt{2}} (|L\rangle + |R\rangle), \quad |A\rangle = \frac{1}{\sqrt{2}} (|L\rangle - |R\rangle) \quad (9.14)$$

There is no dipole moment in either of these two states. However, if we neglect this splitting and take arbitrary linear combinations of $|L\rangle$ and $|R\rangle$ states, then we can get nonvanishing dipole moment. In particular we have

$$\langle d \rangle_L = -\langle d \rangle_R \neq 0 \quad (9.15)$$

So now there are two competing effects. One is the tunnelling between the left and right states, and the other is the interaction between dipole moment and the electric field. When the electric field is relatively small we should take the symmetric and antisymmetric states as the unperturbed states, whereas when the tunnelling parameter is small we should have the left and right states as unperturbed states.

Now let's come back to linear Stark effect. Remember the degenerate states were the $|200\rangle$ and $|210\rangle$ and the linear combinations are just

$$|\pm\rangle = \frac{1}{\sqrt{2}} (|200\rangle \pm |210\rangle) \quad (9.16)$$

The above discussion is based on the existence of an external electric field. Same thing can happen for two interacting atoms, because one atom will induce a dipole moment on the other atom, as shown in the Van der Waals force model. In fact, when we have quantum fluctuations we should have

$$\langle \mathbf{d} \rangle_0 = 0, \quad \text{but} \quad \langle \mathbf{d}^2 \rangle_0 \neq 0 \quad (9.17)$$

In the interacting atoms case we have

$$\langle \mathbf{d}_{1,2} \rangle_0 = 0, \quad \text{but} \quad \langle \mathbf{d}_1 \cdot \mathbf{d}_2 \rangle_0 \neq 0 \quad (9.18)$$

This is because the interaction brings in a correlation between the two dipole moments.

9.3 Summarizing Perturbation Theory

Now that we have finished the discussion of perturbation theory, let's summarize the up and down sides of the theory. The first good thing about it is that it is broadly applicable. We can apply it to virtually any system where we can write the interaction as a small perturbation. It is also straightforward and we developed a systematic procedure to apply it to any system at hand. We can also extract some important qualitative features from a complicate system using perturbation theory, just like we did in the periodic

lattice potential case. It is also easy to estimate the accuracy of our approximation, as we are always in control of the power series expansion.

How about the downsides of the theory? People usually just go to the first few orders of perturbation theory, and assume that these are sufficient to approximate the whole series. But how do we know the series itself even converges? The higher order terms will be extremely messy and hard to calculate, and they can even diverge, as in the case of loop diagrams in quantum field theory. Remember in QED we expand the perturbation series in powers of the fine structure constant α . Dyson argued that if the power series is behaving nicely, then it will be an analytic function in α , so it has no singularities at $\alpha \rightarrow 0$. Moreover it should be nice even at $\alpha \rightarrow -\alpha$. But then if we have $-\alpha$ then the electrons will attract each other instead of repelling, and there will be serious problems occurring in high order interactions.

The resolution of this issue is that we say the perturbation series is an asymptotic series instead of a converging series, and it is valid only up to some order in the parameter. However we also have the nonperturbative corrections like $e^{1/\lambda}$ which does not show up to any order of perturbation.

10 Lecture 10

10.1 Variational Method

Recall that in Classical Mechanics we use the variational method to minimize the Lagrangian and obtain the equation of motion. The same idea can be transformed in QM, and give us an upper bound in the ground state energy E_0 . We can prove the following theorem

Theorem 1. *For any state $|\alpha\rangle$, the expectation value $\langle\alpha|H|\alpha\rangle/\langle\alpha|\alpha\rangle$ can't be smaller than the ground state energy E_0 . Moreover, if $\langle H\rangle_\alpha = E_0$, then $|\alpha\rangle \propto |0\rangle$.*

Proof. We can expand $|\alpha\rangle$ using an eigenbasis of the Hamiltonian

$$|\alpha\rangle = \sum_k |k\rangle \langle k|\alpha\rangle \quad (10.1)$$

Then we can write the expectation value of H with respect to $|\alpha\rangle$ in the following way

$$\begin{aligned} \langle H\rangle_\alpha &= \frac{1}{\langle\alpha|\alpha\rangle} \sum_k E_k |\langle\alpha|k\rangle|^2 \\ &= E_0 - \frac{E_0}{\langle\alpha|\alpha\rangle} \sum_k |\langle\alpha|k\rangle| + \frac{1}{\langle\alpha|\alpha\rangle} \sum_k E_k |\langle\alpha|k\rangle|^2 \\ &= E_0 + \frac{1}{\langle\alpha|\alpha\rangle} \sum_k (E_k - E_0) |\langle\alpha|k\rangle|^2 \end{aligned} \quad (10.2)$$

Obviously the second term is non-negative, because E_0 is the ground state and the other quantities are the norms and squares. Therefore the expectation must be larger than or equal to E_0 . The equality only holds when $\langle\alpha|k\rangle = 0$ for all $k \neq 0$. \square

Although this is a theorem that is easy to prove, it can give us important information about the ground state. It allows us to make an approximation in the following way. Suppose we can't or don't want to find the exact ground state energy, but only want to approximate it. We can find the minimum of $\langle H\rangle_\alpha$ among a class of states which we call trial states. This will be a good approximation of the ground state energy if we have chosen the trial functions cleverly.

We can choose our trial states characterized by a finite set of parameters $\lambda_1, \lambda_2, \dots$ and evaluate the expectation value $\langle H\rangle = \tilde{E}(\lambda_i)$ as a function of these parameters. It is then easy to find the minimum of the expectation value using conventional methods, i.e. the derivative

$$\frac{\partial \tilde{E}}{\partial \lambda_i} = 0 \quad \implies \quad \lambda_i = \lambda_i^* \quad (10.3)$$

Then our approximate ground state is just $|\alpha\rangle = |\lambda_i^*\rangle$, and the value $\tilde{E}(\lambda_i^*)$ is an approximation of E_0 .

10.2 Example

Let's consider the familiar example of simple harmonic oscillator

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2} x^2 \quad (10.4)$$

We know the exact solution of the ground state which is $E_0 = \hbar\omega/2$. Let's try to calculate this using variational method and compare it with the exact solution. Let's consider trial functions of the form of a Gaussian

$$\psi_\lambda(x) = \langle x|\lambda\rangle = \left(\frac{\lambda}{\pi}\right)^{1/4} \exp\left(-\frac{\lambda x^2}{2}\right) \quad (10.5)$$

The wavefunction is normalized to simplify calculations. We want to find the λ which gives the smallest value of the expectation. Because the expectation value integral is just a Gaussian integral, we can just write down the result

$$\tilde{E}(\lambda) = \frac{\lambda\hbar^2}{4m} + \frac{m\omega^2}{4\lambda} \quad (10.6)$$

Taking derivative we can get

$$\frac{\partial \tilde{E}}{\partial \lambda} = \frac{\hbar^2}{4m} - \frac{m\omega^2}{4\lambda^2} = 0 \implies \lambda^* = \frac{m\omega}{\hbar} \quad (10.7)$$

And the energy at this minimum is $\tilde{E}(\lambda^*) = \hbar\omega/2$, which is the same as the exact value. So we have success using the variational method. This success is based on our clever choice of trial function, which coincides with the true ground state wavefunction.

Let's now choose another set of trial functions

$$\psi_\lambda(x) = \sqrt{\frac{2\lambda^3}{\pi}} \frac{1}{x^2 + \lambda^2} \quad (10.8)$$

The expectation value can be evaluated as

$$\tilde{E}(\lambda) = \frac{\hbar^2}{4m\lambda^2} + \frac{1}{2}m\omega^2\lambda^2 \quad (10.9)$$

and using the same prescription we get

$$\lambda^{*2} = \frac{\hbar}{m\omega\sqrt{2}}, \quad \tilde{E}(\lambda^*) = \frac{\hbar\omega}{\sqrt{2}} > \frac{\hbar\omega}{2} \quad (10.10)$$

Note that instead of the exact value, we get something bigger. This is because we have greatly over estimated the behavior of the wavefunction at large x , and the price to pay is getting a larger energy estimation.

10.3 Generalization

Let's say we have a state $|\alpha\rangle$ and define $|\alpha'\rangle = |\alpha\rangle + \epsilon|\alpha_1\rangle$ where ϵ is small. In general we get

$$\langle H \rangle_{\alpha'} = \langle H \rangle_\alpha + O(\epsilon) \quad (10.11)$$

This tells us to define an energy functional $\mathcal{F}[|\alpha\rangle]$ which is defined on the Hilbert space of interest. We generalize our previous problem into the minimization of this functional, which corresponds to finding the state where first order variation vanish

$$|\alpha\rangle \rightarrow |\alpha\rangle + \epsilon|\alpha_1\rangle \implies \mathcal{F}[|\alpha'\rangle] = \mathcal{F}[|\alpha\rangle] + O(\epsilon^2) \quad (10.12)$$

Now we have the following generalized theorem

Theorem 2. *The energy functional $\langle H \rangle_\alpha$ is stationary with respect to variation of $|\alpha\rangle$ if and only if $|\alpha\rangle = |n\rangle$.*

Proof. We expand the functional $\langle H \rangle_{\alpha'}$ to first order in ϵ and we have the following

$$\langle \alpha | \alpha \rangle [\langle H \rangle_{\alpha'} - \langle H \rangle_\alpha] = \epsilon [\langle \alpha_1 | H - E_\alpha | \alpha \rangle + \langle \alpha | H - E_\alpha | \alpha_1 \rangle] \quad (10.13)$$

where $E_\alpha = \langle H \rangle_\alpha$. So the first order variation vanishes if $|\alpha\rangle$ is an eigenket. We have assumed $|\alpha\rangle$ orthogonal to $|\alpha_1\rangle$, which can always be chosen to be true. Now if the above expression is zero for any $|\alpha_1\rangle$ orthogonal to $|\alpha\rangle$, then we can choose $|\alpha_1\rangle = (H - E_\alpha)|\alpha\rangle$ then we have

$$2\epsilon \langle \alpha | (H - \langle H \rangle_\alpha)^2 | \alpha \rangle = \langle \alpha_1 | \alpha_1 \rangle = 0 \quad (10.14)$$

so $|\alpha_1\rangle = 0$. Therefore $|\alpha\rangle$ is an eigenket. \square

Therefore we can set out to find all the extrema of the energy functional to get the energy spectrum, as well as any degenerate ground state.

We can find connections between the variational method and perturbation theory. Because the Hamiltonian can be written as $H = H_0 + V$, so the energy expectation value of the Hamiltonian with respect to the unperturbed ground state is just

$$E_0 = E_0^{(0)} + \langle V \rangle_0 \quad (10.15)$$

which is just the result from the first order perturbation theory. This is the result of choosing the unperturbed states as our trial wavefunctions. The question is can we do better?

Without loss of generality we can assume $\langle V \rangle_0 = 0$ which can always be done by shifting the energy. Let's choose the trial functions to be

$$|\lambda\rangle = |0\rangle_0 + \lambda V |0\rangle_0 \quad (10.16)$$

Because the expectation value of V with respect to $|0\rangle_0$ is zero, the second term is always orthogonal to the first term. So the energy expectation value is

$$\begin{aligned} \langle H \rangle_\lambda &= \frac{\langle 0 | (1 + \lambda V)(H_0 + V)(1 + \lambda V) | 0 \rangle}{\langle 0 | (1 + \lambda V)^2 | 0 \rangle} \\ &= \frac{E_0^{(0)} + 2\lambda \langle V^2 \rangle_0 + \lambda^2 \langle 0 | V H_0 V | 0 \rangle}{1 + \lambda^2 \langle V^2 \rangle_0} \end{aligned} \quad (10.17)$$

For small λ we can expand the denominator and get

$$E_0 \leq E_0^{(0)} + 2\lambda \langle V^2 \rangle_0 + \lambda^2 [\langle V H_0 V \rangle_0 - E_0^{(0)} \langle V^2 \rangle_0] \quad (10.18)$$

Therefore we can straightforwardly work out

$$\lambda^* = \frac{\langle V^2 \rangle_0}{E_0^{(0)} \langle V^2 \rangle_0 - \langle V H_0 V \rangle_0} \quad (10.19)$$

Note than we can write $\langle V^2 \rangle_0 = \sum |V_{0k}|^2$ and we get

$$E_0 \leq E_0^{(0)} - \frac{1}{E_1^{(0)} - E_0^{(0)}} \sum_{k \neq 0} |V_{0k}|^2 \quad (10.20)$$

The E_1 is there to get a simpler form of the upper bound. And we recover a result similar to second order perturbation theory.

From the variational theorem, in simple harmonic oscillator we have

$$\frac{1}{2m} \langle p^2 \rangle_\alpha + \frac{m\omega^2}{2} \langle x^2 \rangle_\alpha \geq \frac{\hbar\omega}{2} \quad (10.21)$$

This is true for any state $|\alpha\rangle$ and any ω . The inequality can be written as

$$\langle p^2 \rangle + m^2\omega^2 \langle x^2 \rangle - \hbar m\omega \geq 0 \quad (10.22)$$

The frequency where the above expression takes the minimum is at

$$\omega^* = \frac{\hbar}{2m \langle x^2 \rangle_\alpha} \quad (10.23)$$

Plugging this into the above inequality it should still be larger or equal to zero, and we get

$$\langle p^2 \rangle + \frac{\hbar^2}{4 \langle x^2 \rangle_\alpha} - \frac{\hbar^2}{2 \langle x^2 \rangle_\alpha} \geq 0 \quad \implies \quad \langle x^2 \rangle \langle p^2 \rangle \geq \frac{\hbar^2}{4} \quad (10.24)$$

So from the variational principle we derived the Heisenberg uncertainty principle.

Similar considerations can be applied to the hydrogen atom, and we have

$$\frac{\langle p^2 \rangle_\alpha}{2m} - e^2 \left\langle \frac{1}{r} \right\rangle_\alpha \geq -\frac{me^2}{2\hbar^2} \quad (10.25)$$

and this is correct for all $|\alpha\rangle$ and all parameters me^2 , and we have

$$\langle p^2 \rangle \geq \hbar^2 \left\langle \frac{1}{r} \right\rangle^2 \quad (10.26)$$

11 Lecture 11

11.1 Time-dependent Perturbation

When the Hamiltonian is independent of time, we know that the energy eigenstates $|n\rangle$ with energy n will evolve in the simple relation in Schrödinger picture

$$|n, t\rangle = e^{-iE_n t/\hbar} |n\rangle \quad (11.1)$$

We also have the Heisenberg picture, where our states are time-independent and the operators carry all the time-dependence, and we have the Heisenberg equation of motion instead of the Schrödinger equation

$$i\hbar \frac{dA}{dt} = [A, H] \quad (11.2)$$

The motivation of considering time-dependent perturbation theory lies in atomic physics, where people use lasers to excite atoms to study atomic spectra. The Hamiltonian now looks like

$$\mathcal{H} = \mathcal{H}_0 + \lambda V(t) \quad (11.3)$$

In principle we can work this out in either Heisenberg or Schrödinger picture. But instead we will follow the practice in the field and use the interaction picture, which is more convenient in this situation. Let's start by considering the Schrödinger picture kets $|\alpha, t\rangle_S$. The initial state can be expanded by the unperturbed energy eigenstates

$$|\alpha, 0\rangle_S = |\alpha\rangle = \sum |n\rangle \langle n|\alpha\rangle \quad (11.4)$$

Now drawing analogy with time-independent theory we will expect the time evolution of the state to be

$$|\alpha, t\rangle_S = \sum c_n(t) |n\rangle e^{-E_n t/\hbar} \quad (11.5)$$

The coefficient $c_n(t)$ is time-dependent because the perturbation is time-dependent. This is just expanding the time evolved state into the time evolved energy eigenstates, and the probability of a transition to state $|n\rangle$ is $P(n) = |c_n(t)|^2$. But in general it is hard to determine the coefficients $c_n(t)$. To do this we introduce the interaction picture as follows

$$|\alpha, t\rangle_S \longrightarrow |\alpha, t\rangle_I = e^{i\mathcal{H}_0 t/\hbar} |\alpha, t\rangle_S \quad (11.6)$$

The time derivative of this term is

$$\begin{aligned} i\hbar \frac{\partial |\alpha, t\rangle_I}{\partial t} &= -\mathcal{H}_0 |\alpha, t\rangle_I + e^{i\mathcal{H}_0 t/\hbar} (\mathcal{H}_0 + V) |\alpha, t\rangle_S \\ &= e^{i\mathcal{H}_0 t/\hbar} V(t) e^{-i\mathcal{H}_0 t/\hbar} |\alpha, t\rangle_I \\ &= V_I(t) |\alpha, t\rangle_I \end{aligned} \quad (11.7)$$

where we define the interaction picture operator as

$$O_I(t) = e^{i\mathcal{H}_0 t/\hbar} O(t) e^{-i\mathcal{H}_0 t/\hbar} \quad (11.8)$$

And the time evolution equation for any interaction picture operator is

$$i\hbar \frac{dO_I}{dt} = [O_I, \mathcal{H}_0] \quad (11.9)$$

Again we can expand the interaction picture states into the energy eigenstates at different times

$$|\alpha, t\rangle_I = \sum c_n(t) |n\rangle, \quad c_n(t) = \langle n|\alpha, t\rangle_I \quad (11.10)$$

Then the above equation of evolution of state $|\alpha, t\rangle_I$ becomes

$$\begin{aligned} i\hbar \frac{\partial c_n}{\partial t} &= \sum_m \langle n|V_I|m\rangle \langle m|\alpha, t\rangle_I \\ &= \sum_m c_m(t) \langle n|e^{iE_n t/\hbar} V_S(t) e^{-iE_m t/\hbar}|m\rangle \\ &= \sum_m e^{i\omega_{nm}t} V_{nm} c_m(t) \end{aligned} \quad (11.11)$$

where we have defined $\omega_{nm} = (E_n - E_m)/\hbar$, and the matrix element V_{nm} is the usual matrix element with respect to the unperturbed energy eigenstates. This is a linear system of first order partial differential equations and in principle can be solved to get $c_n(t)$. One thing we know without solving the equation is that the norm is conserved

$$\frac{\partial}{\partial t} \langle \alpha, t|\alpha, t\rangle_I = 0 \quad (11.12)$$

This can be checked directly using the above equations.

11.2 Examples

Let's consider a driven 2-level system. The basis for the original Hamiltonian is $|1\rangle$ and $|2\rangle$ and the original Hamiltonian in this basis is

$$\mathcal{H}_0 = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} = E_1 |1\rangle\langle 1| + E_2 |2\rangle\langle 2| \quad (11.13)$$

Now we perturb it with some oscillating harmonic perturbation

$$V(t) = \hbar\gamma [e^{i\omega t} |1\rangle\langle 2| + e^{-i\omega t} |2\rangle\langle 1|] \quad (11.14)$$

where γ and ω are real. This is not the most general case for harmonic perturbation, but it is enough for our purpose for the moment. This could be realized as the spin in a time-dependent magnetic field

$$\mathbf{B} = B_0 \hat{\mathbf{z}} + B_1 \cos \omega t \hat{\mathbf{x}} + B_1 \sin \omega t \hat{\mathbf{y}} \quad (11.15)$$

This is a field that is rotating around the z axis. In this case we have

$$\mathcal{H}_0 = \frac{2\mu_B}{\hbar} B_0 S_z, \quad V(t) = 2\mu_B B_1 [S_x \cos \omega t + S_y \sin \omega t] \quad (11.16)$$

This problem will be identical to our 2-level system above if we identify $|1\rangle = |z, +\rangle$ and $|2\rangle = |z, -\rangle$, and use the usual Pauli matrix representation of the spin matrices. So let's go back to the toy problem. The perturbation matrix elements are just

$$V_{11} = V_{22} = 0, \quad V_{12} = V_{21}^* = \hbar\gamma e^{i\omega t} \quad (11.17)$$

Now we can plug into the equation (11.11) and get a set of equations

$$\begin{cases} i\hbar\dot{c}_1 = \hbar\gamma e^{i(\omega-\omega_{21})t} c_2 \\ i\hbar\dot{c}_2 = \hbar\gamma e^{-i(\omega-\omega_{21})t} c_1 \end{cases} \quad (11.18)$$

This is just a system of two linear first order PDEs and we know that we can convert it to two independent second order PDEs, one of which is

$$\ddot{c}_1 = i\Omega_0\dot{c}_1 - \gamma^2 c_1 \quad (11.19)$$

where we have introduced $\Omega_0 = \omega - \omega_{21}$. We look for solutions of the type $c_1 = Ae^{i\Omega_*t}$ and plug this trial solution into the above equation

$$-\Omega_*^2 = -\Omega_0\Omega_* - \gamma^2 \quad (11.20)$$

This is a quadratic equation with solution

$$\Omega_* = \Omega_{\pm} = \frac{\Omega_0}{2} \pm \frac{1}{2}\sqrt{\Omega_0^2 + 4\gamma^2} = \frac{\Omega_0}{2} \pm \Omega \quad (11.21)$$

So the general solution is

$$c_1 = (A_1 \sin \Omega t + B_1 \cos \Omega t) e^{i\Omega_0 t/2} \quad (11.22)$$

$$c_2 = -i\gamma e^{-i\Omega_0 t} \dot{c}_1 = e^{-i\Omega_0 t/2} (A_2 \cos \Omega t + B_2 \sin \Omega t) \quad (11.23)$$

By normalization condition $|c_1|^2 + |c_2|^2 = 1$ and choose an initial condition $c_1(0) = 1$ and $c_2(0) = 0$, we have the explicit result

$$|c_1|^2 = 1 - \frac{\gamma^2}{\gamma^2 + \Omega_0^2/4} \sin^2 \Omega t, \quad |c_2|^2 = \frac{\gamma^2}{\gamma^2 + \Omega_0^2/4} \sin^2 \Omega t \quad (11.24)$$

The oscillation amplitude is

$$A^2 = \frac{\gamma^2}{\gamma^2 + (\omega - \omega_{21})^2/4} \quad (11.25)$$

so the oscillation is small when ω is away from ω_{21} , but it becomes large at $\omega = \omega_{21}$ and we have a resonance, where the state becomes oscillating back and forth between $|1\rangle$ and $|2\rangle$. If we make γ smaller, the amplitude at resonance is not changed, but the width of the resonance peak becomes narrower and narrower.

Now let's go back to our system of spin 1/2 considered above. If the magnetic is purely in the z -direction and time-independent, we know that we have spin precession

$$\langle S_z \rangle = \frac{\hbar}{2} (|c_1(0)|^2 - |c_2(0)|^2), \quad \langle S_x \rangle = \frac{\hbar}{2} 2\text{Re} [c_1^*(0)c_2(0)e^{-i\omega_{21}t}] \quad (11.26)$$

Now if $B_1 \neq 0$ but it is small, then we have

$$\langle S_z \rangle = \frac{\hbar}{2} (2|c_1(t)|^2 - 1) = \frac{\hbar}{2} \left(|A_1^{(+)}|^2 + |A_1^{(-)}|^2 - \frac{1}{2} \right) + \hbar\text{Re} (A_1^{(+)*} A_2 e^{-2i\Omega t}) \quad (11.27)$$

In resonance we have $\Omega = 2\gamma$, and $\langle S_z \rangle$ oscillates with frequency 4γ .

12 Lecture 12

12.1 Time-dependent Perturbation Theory Continued

Last time we considered the problem where the Hamiltonian is explicitly time-dependent, but the time dependence can be written as a perturbation. We developed a way to solve this kind of problems approximately. Last time we considered a two level system in an external field. The unperturbed Hamiltonian is diagonal

$$\mathcal{H}_0 = E_1 |1\rangle\langle 1| + E_2 |2\rangle\langle 2| \quad (12.1)$$

We introduced a perturbation of the form

$$V(t) = \hbar\gamma [e^{i\omega t} |1\rangle\langle 2| + e^{-i\omega t} |2\rangle\langle 1|] \quad (12.2)$$

which are the off-diagonal elements. We introduced the interaction picture and expanded the states in the two eigenstates of the unperturbed Hamiltonian

$$|\alpha, t\rangle_I = c_1(t) |1\rangle + c_2(t) |2\rangle, \quad |c_1|^2 + |c_2|^2 = 1 \quad (12.3)$$

If we take the initial condition that $c_1(0) = 1$, then we found the solution to this initial value problem to be

$$|c_1(t)|^2 = 1 - \frac{\gamma^2}{\Omega^2} \sin^2 \Omega t, \quad |c_2(t)|^2 = \frac{\gamma^2}{\Omega^2} \sin^2 \Omega t \quad (12.4)$$

where $\Omega = \sqrt{(\omega - \omega_{21})^2/\hbar + 4\gamma/2}$ is called the Rabi frequency. The amplitude of oscillation will go to 1 when $\omega = \omega_{21}$, which is the expected resonance behavior. If we consider the very short time behavior, then

$$|c_2(t)|^2 \xrightarrow{t \rightarrow 0} \gamma^2 t^2 \quad (12.5)$$

Note the probability grows like t^2 instead of t , which is a little counter-intuitive.

12.2 Formal Development

Now we want to develop a systematic way of treating time-dependent perturbation theory. Let's consider a Hamiltonian of the form

$$\mathcal{H} = \mathcal{H}_0 + V(t), \quad \mathcal{H}_0 |n\rangle = E_n |n\rangle \quad (12.6)$$

We expand any state in the interaction picture as above, and the coefficients satisfy the following differential equation, as shown last time

$$\frac{\partial c_n}{\partial t} = \sum_m \frac{V_{nm}}{i\hbar} e^{i\omega_{nm}t} c_m(t) \quad (12.7)$$

The above result is exact. Now for perturbation theory, we can expand the coefficients and try to solve the equation iteratively

$$c_n(t) = c_n^{(0)} + c_n^{(1)}(t) + c_n^{(2)}(t) + \dots \quad (12.8)$$

Then we can write the above differential equation as

$$\frac{\partial c_n^{(j+1)}}{\partial t} = \sum_m \frac{V_{nm}}{i\hbar} e^{i\omega_{nm}t} c_m^{(j)}(t) \quad (12.9)$$

So in the perturbation theory every order only depends on the previous order, so we have a nice solvable theory. However, this set of equation, like demonstrated above, will need a set of initial conditions to have a unique solution. For example we can have $c_n(0) = \delta_{nm}$. So it would be nicer to have an alternative set of equations which do not depend on the initial conditions.

Let's start with the time-evolution operator $U(t, t_0)$. In the interaction picture its definition is

$$|\alpha, t\rangle = U_I(t, t_0) |\alpha, t_0\rangle \quad (12.10)$$

The equation that corresponds to the Schrödinger equation in the interaction picture is

$$i\hbar \frac{\partial |\alpha, t\rangle_I}{\partial t} = V_I(t) |\alpha, t\rangle_I \quad (12.11)$$

and it is true if we replace $|\alpha, t\rangle_I$ with $U_I(t, t_0)$

$$i\hbar \frac{\partial U_I(t, t_0)}{\partial t} = V_I(t) U_I(t, t_0) \quad (12.12)$$

with the initial condition $U_I(t_0, t_0) = 1$. We can integrate on both sides to get an integral equation

$$i\hbar [U_I(t, t_0) - 1] = \int_{t_0}^t V_I(t') U_I(t', t_0) dt' \quad (12.13)$$

In general integral equations are much easier to iterate, because in order to get the first order solution we just substitute the zero order result into the integral, and so on. So we have

$$U_I(t, t_0) = 1 + \frac{1}{i\hbar} \int_{t_0}^t dt_1 V_I(t_1) + \frac{1}{(i\hbar)^2} \int_{t_0}^t dt_1 V_I(t_1) \int_{t_0}^{t_1} dt_2 V_I(t_2) + \dots \quad (12.14)$$

It can happen that the perturbation $V_I(t)$ fail to commute at different times, i.e. $[V_I(t), V_I(t')] \neq 0$. This can happen, for example, when the perturbation is spin coupled with magnetic field and the field changes direction at different times, and we know that the spin of different directions don't commute. We need to keep track of this non-commutativity in our perturbation expansion (12.14). Let's consider the second term, we write it as

$$\begin{aligned} \text{Second term} &= \frac{1}{2(i\hbar)^2} \left[\int_{t_0}^t dt_1 V_I(t_1) \int_{t_0}^{t_1} dt_2 V_I(t_2) + \int_{t_0}^t dt_2 V_I(t_2) \int_{t_0}^{t_2} dt_1 V_I(t_1) \right] \\ &= \frac{1}{2(i\hbar)^2} \left[\int_{t_0}^t dt_1 V_I(t_1) \int_{t_0}^{t_1} dt_2 V_I(t_2) + \int_{t_0}^t dt_1 \int_{t_1}^t dt_2 V_I(t_2) V_I(t_1) \right] \end{aligned} \quad (12.15)$$

We introduce the time-ordering operator

$$T \{A(t_1)B(t_2)\} = \begin{cases} A(t_1)B(t_2), & \text{if } t_1 > t_2 \\ B(t_2)A(t_1), & \text{if } t_2 > t_1 \end{cases} \quad (12.16)$$

and it is straight forward to generalize the definition to products of n operators at different times. Then the second term can be written now as

$$\text{Second term} = \frac{1}{2(i\hbar)^2} \int_{t_0}^t dt_1 dt_2 T \{V_I(t_1)V_I(t_2)\} \quad (12.17)$$

And the perturbation expansion (12.14) will be generalized to the following form

$$U_I(t, t_0) = \sum_k \frac{1}{(i\hbar)^k} \frac{1}{k!} \int_{t_0}^t dt_1 \dots dt_k T \{V_I(t_1) \dots V_I(t_k)\} \quad (12.18)$$

It can be realized that apart from a time-ordering symbol, the summand in the above equation matches exactly that in an exponential. If the perturbation at different times commute, then we can just write the above equation as

$$U_I(t, t_0) = \exp \left[\frac{1}{i\hbar} \int_{t_0}^t dt' V_I(t') \right] \quad (12.19)$$

In general we have

$$U_I(t, t_0) = T \left\{ \exp \left[\frac{1}{i\hbar} \int_{t_0}^t dt' V_I(t') \right] \right\} \quad (12.20)$$

where the time-ordering is understood to be carried out for each order. Directly from this expression we can easily derive the following identity

$$U(t, t_1)U(t_1, t_0) = U(t, t_0), \quad \text{for any } t_0 \leq t_1 \leq t \quad (12.21)$$

It is also easy to write down the time evolution of a quantum state

$$|\alpha, t\rangle_I = \sum_n |n\rangle \langle n | U_I(t, t_0) | \alpha, t_0 \rangle_I = \sum_n c_n(t) |n\rangle \quad (12.22)$$

So the initial conditions only appear in the state $|\alpha, t_0\rangle_I$, but the time evolution is all contained in the above exponential. If the initial condition is chosen such that $|\alpha, t_0\rangle_I = |m\rangle$, then we have the following

$$c_n(t) = \delta_{nm} + \frac{1}{i\hbar} \int_{t_0}^t dt_1 \langle m | V_I(t_1) | n \rangle + \frac{1}{2(i\hbar)^2} \int_{t_0}^t dt_1 dt_2 \langle m | T \{V_I(t_1)V_I(t_2)\} | n \rangle + \dots \quad (12.23)$$

One of the most important application of the above formal theory is to consider the transition probability of a state to another under the time-dependent perturbation. The probability of transition is

$$\mathcal{P}_{nm}(t, t_0) = |\langle n | U_I(t, t_0) | m \rangle|^2 \quad (12.24)$$

As a demonstration let's consider harmonic perturbation, and only consider the first order perturbation. Let's consider

$$V_I(t) = \nu e^{i\omega t} + \nu^\dagger e^{-i\omega t} \quad (12.25)$$

The first order in our coefficient expansion is just

$$c_n^{(1)}(t) = -\frac{i}{\hbar} \int dt V_{mn} e^{i\omega_{mn}t} \quad (12.26)$$

For this particular potential we have

$$V_{mn}(t) e^{i\omega_{mn}t} = \nu_{mn} e^{i(\omega_{mn} + \omega)t} + \nu_{nm}^* e^{i(\omega_{mn} - \omega)t} \quad (12.27)$$

so it is straightforward to integrate this expression to get

$$c_n^{(1)} = -\frac{\nu_{mn} \exp[i(\omega_{mn} + \omega)t]}{\hbar(\omega_{mn} + \omega)} - \frac{\nu_{nm}^* \exp[i(\omega_{mn} - \omega)t]}{\hbar(\omega_{mn} - \omega)} \quad (12.28)$$

The above expression is correct as long as $\hbar|\omega_{mn} \pm \omega| \gg |\nu_{mn}|$, and in particular away from resonance. If we are dealing with resonance, perturbation is not quite correct, but then we can just approximate the system as a two-level system and use our result in the beginning of this lecture.

What if we have a continuous spectrum? For example we have continuous spectrum above E_{\min} but under which we have discrete spectrum. We want to study the situation when $\hbar\omega > E_{\min} - E_m$. Let's take $|m\rangle$ to be the initial state and $|f\rangle$ to be the final state, which can be in the continuous spectrum. The transition probability is

$$\mathcal{P}_{i \rightarrow f} = |c_{fi}|^2 = |c_m(t)|^2 \quad (12.29)$$

Our harmonic perturbation result above will be generalized to

$$\mathcal{P}_{i \rightarrow f} = \frac{|\nu_{fi}|^2}{\hbar^2} \frac{4}{(\omega_{fi} - \omega)^2} \sin^2 \left(\frac{\omega_{fi} - \omega}{2} t \right) \quad (12.30)$$

A mathematical identity tells us that

$$\lim_{t \rightarrow \infty} \frac{\sin^2 \alpha t}{\pi \alpha^2 t} = \delta(\alpha) \quad (12.31)$$

This can be checked easily by checking the properties and doing an integration to fix the coefficient. This allows us to write down the Fermi's golden rule, which is valid for time much larger than the characteristic transition time $1/(\omega_{mn} - \omega)$

$$\mathcal{P}_{i \rightarrow f}(t) = \frac{|\nu_{fi}|^2}{\hbar^2} \pi t \delta \left(\frac{\omega_{fi} - \omega}{2} \right) = \frac{2\pi}{\hbar} |\nu_{fi}|^2 \delta(E_f - E_i - \hbar\omega) t \quad (12.32)$$

Originally we have the probability going like t^2 , but now as we introduce a delta function, the probability becomes proportional to t . So the transition rate, which is the transition probability per unit time, is a constant.

13 Lecture 13

13.1 Time-dependent Perturbation Continued

Last time we were talking about time-dependent perturbation theory, with a Hamiltonian of the form

$$\mathcal{H} = \mathcal{H}_0 + V(t) \quad (13.1)$$

Our strategy is to expand the states at different times using the eigenstates of the unperturbed Hamiltonian. We studied the harmonic perturbation which has the form

$$V(t) = \nu e^{i\omega t} + \nu^\dagger e^{-\omega t} \quad (13.2)$$

where we assume that ν_{mn} are small. If the initial state is an eigenstate of the original Hamiltonian, then we found last time the transition probability of the state to another eigenstate at time t

$$\mathcal{P}_{fi} = \frac{4}{\hbar^2} \left[\frac{|\nu_{fi}|^2}{(\omega_{fi} + \omega)^2} \sin^2 \left(\frac{\omega_{fi} + \omega}{2} t \right) + \frac{|\nu_{fi}^\dagger|^2}{(\omega_{fi} - \omega)^2} \sin^2 \left(\frac{\omega_{fi} - \omega}{2} t \right) \right] \quad (13.3)$$

We found that there is resonance when the frequency of the perturbation is equal to the frequency between the two states, and oscillation between the initial state and final state will have amplitude 1. What we can do and what people usually do, when considering $\omega > 0$, then when $\omega_{fi} \approx \omega$ we can ignore the first term, and when $\omega_{fi} \approx -\omega$ then we can ignore the second term. The first case is called absorption, and the second case is called stimulated emission. It is called “stimulated” because this transition from a lower energy state to a higher state is not possible without the perturbation.

Let’s consider the case of absorption. The probability is just

$$\mathcal{P}_{fi}^{(a)} = \frac{4}{\hbar^2} |\nu_{if}|^2 \frac{\sin^2 [(\omega_{fi} - \omega)/2]}{(\omega_{fi} - \omega)^2} = \frac{\pi}{\hbar^2} |\nu_{if}|^2 t \mathcal{F} \left(\frac{\omega_{fi} - \omega}{2}, t \right) \quad (13.4)$$

where we have defined the function

$$\mathcal{F}(\alpha, t) = \frac{\sin^2 \alpha t}{\pi \alpha^2 t} \xrightarrow{t \rightarrow \infty} \delta(\alpha) \quad (13.5)$$

This is because we have $\int \sin^2 z / z^2 dz = \pi$ by contour integration. The plot of the function \mathcal{F} is

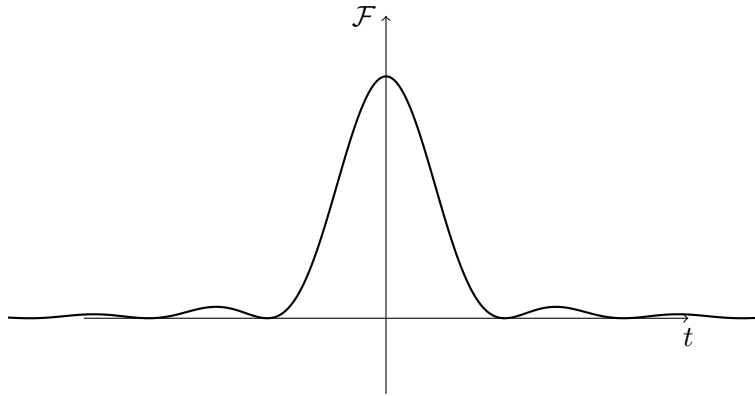
In the limit where t is much larger than the transition characteristic time, we have

$$\mathcal{P}_{fi}^{(a)}(t) = \frac{\pi}{\hbar^2} |\nu_{if}|^2 t \delta \left(\frac{\omega_{fi} - \omega}{2} \right) = \frac{2\pi}{\hbar} |\nu_{if}|^2 t \delta(E_f - E_i - \hbar\omega) \quad (13.6)$$

Where the probability, written in the final form, is called the Fermi Golden Rule. Notice the linear dependence on t which is different from the usual result from perturbation. We can define the *absorption rate*

$$W_{fi} = \frac{2\pi}{\hbar} |\nu_{if}|^2 \delta(E_f - E_i - \hbar\omega) \quad (13.7)$$

How to interpret the delta function in the absorption rate? Let’s first try to understand it formally. Let’s consider some initial state $|i\rangle$ going into a continuous spectrum. We can define the density of states

Figure 13.1: Plot of the Function \mathcal{F}

(DoS) $\rho(E)$ of the continuous spectrum. Its physical meaning is that if take the continuous spectrum as approximated by a very dense discrete spectrum, we can count the number of states in the energy interval ΔE . The density of state is just this number density under the limit where the spectrum is continuous

$$\rho(E) = \lim_{\Delta E \rightarrow 0} \frac{\Delta n}{\Delta E} \quad (13.8)$$

For example, consider the infinite square well, we have energy eigenstates

$$\langle x|n\rangle = \frac{1}{\sqrt{a}} \sin \frac{p_n x}{\hbar}, \quad p_n = \frac{n\pi\hbar}{a}, \quad E_n = \frac{n^2\pi^2\hbar^2}{2ma^2} \quad (13.9)$$

If we take the width of the well $a \rightarrow \infty$, then the spacing between two energies will go to zero and we will have an approximately continuous spectrum. We have

$$\Delta(n^2) = 2n\Delta n = \frac{2ma^2}{\pi^2\hbar^2} \Delta E \quad (13.10)$$

So the density of state is then

$$\rho(E) = \frac{\Delta n}{\Delta E} = \frac{ma^2}{\pi^2\hbar^2 n} = \sqrt{\frac{m}{2\pi^2\hbar^2}} \frac{a}{\sqrt{E}} \quad (13.11)$$

Notice that it depends on the volume of the system a , and has some dependency on E which in this case is $E^{-1/2}$. This is a generic behavior of any density of state function. It should have some energy dependence and also proportional to the total volume of the system.

If we consider not the transition to some state, but to some energy, we can modify our definition of transition probability to the following

$$\mathcal{P}_{fi}^{(a)} \rightarrow \overline{\mathcal{P}}^{(a)}(E) = \sum_f \mathcal{P}_{fi}^{(a)} \theta(-(E_f - E)^2 + (\Delta E)^2/2) \quad (13.12)$$

In the case of a continuous spectrum, our transition probability will now become an integral instead of a sum

$$\mathcal{P}(E) = \int \mathcal{P}(E_f) dn(E_f) = \int \mathcal{P}(E)\rho(E) dE = \frac{2\pi}{\hbar} |\overline{\nu}|^2 \rho(E_i + \hbar\omega) \quad (13.13)$$

This is essentially a weighted average over all the energy levels which are close to the final energy E . This obviously assumes that the energy level spacing much smaller than the energy change in the transition. This is the formal way to interpret the delta function in the transition rate.

Now let's take a physical interpretation. We need to understand that there is nothing perfect in the real world, and there is no such thing as a perfect harmonic signal. This means that the perturbation should always be written as a Fourier integral

$$V(t) = \int d\omega f(\omega) \left[\nu(\omega)e^{i\omega t} + \nu^\dagger(\omega)e^{-i\omega t} \right] \quad (13.14)$$

So the transition rate should also be integrated over the frequency

$$W_{fi} = \int W_{fi}(\omega) f^2(\omega) d\omega \quad (13.15)$$

If we substitute the Fermi golden rule into the above equation, we will get

$$W_{fi} = \frac{2\pi}{\hbar} |\nu_{if}(\omega_{fi})|^2 f^2(\omega_{fi}) \quad (13.16)$$

So now the delta function in the transition rate becomes a distribution function instead of a delta function. So we should view our delta function as a filtering function of a perfect harmonic signal, which is indeed a delta function. If we smooth it out using, say, a Gaussian profile, then the rate will be everywhere finite and well defined, just like shown above.

We can also take another physical interpretation. Our approach above assumes that the perturbation is always there for any time t . But in reality we need some mechanism to turn it on at some time $t = 0$ and then turn it off at some later time t' . Let's say the system is originally at state $|i\rangle$ for $t < 0$ when $V(t < 0) = 0$. The perturbation now should look like

$$V(t) = (\nu e^{i\omega t} + \nu^\dagger e^{-i\omega t})\theta(t) \quad (13.17)$$

Now because of the θ function, the Fourier transform of the perturbation is far from narrow. This is not what we want, because we still want to have a relatively narrow frequency range. What we need to do is to introduce an *adiabatic* switching of V . This means that we have $|\alpha\rangle = |i\rangle$ at time $t \rightarrow -\infty$ and $V(t \rightarrow -\infty) = 0$. We can achieve this by multiplying an adiabatic factor on the perturbation

$$V(t) = (\nu e^{i\omega t} + \nu^\dagger e^{-i\omega t})e^{\lambda t} \quad (13.18)$$

We want the characteristic time of turning on the perturbation to be large, which means taking $\lambda \rightarrow 0$. We can now repeat our previous procedures and solve this time-dependent perturbation problem, but there is a short cut. We can just substitute $\omega \rightarrow \omega + i\lambda$ for absorption or $\omega \rightarrow \omega - i\lambda$ for emission, then it becomes the same problem as before. In the end we will have a Lorentz distribution over Ω instead of a delta function distribution

$$\delta(\Omega) \longrightarrow \frac{\lambda}{\lambda^2 + \Omega^2} \quad (13.19)$$

This function has a well-defined peak and a good width which is proportional to the characteristic time of switching. This is what we expect experimentally.

Let's take yet another physical interpretation. If we have an energy spectrum without perturbation, and prepare our system at some energy eigenstate $|n\rangle$, it will stay there forever. Only when we introduce a

perturbation will it undergo a transition to another energy state. However if we consider an atomic system we have something called spontaneous emission, which means that once in a while the atom will emit a photon and an excited electron will go to a lower energy state. This process is not in our Hamiltonian, but we need to take it into account. We need to change our transition rate by

$$\frac{\partial \mathcal{P}}{\partial t} = \left(\frac{\partial \mathcal{P}}{\partial t} \right)_0 + \Gamma \mathcal{P} \quad (13.20)$$

The above expression is a phenomenological one in that we assume the transition rate for spontaneous emission will be proportional to the population. The energy should be replaced by $E_f \rightarrow E_f + i\hbar\Gamma/2$, which will introduce some behavior like what we had in adiabatic turning on of the perturbation

$$\delta(E_f - E_i - \hbar\omega) \longrightarrow \frac{(\hbar\Gamma/2)}{(E_f - E_i - \hbar\omega)^2 + (\hbar\Gamma/2)^2} \quad (13.21)$$

This is what spectroscopy in real life will give us, and which scenario it corresponds will depend on the physical situation we consider. Nowadays people doing spectroscopy are not only interested in the position of the peak, but also the peak width and profile, which could be non-Lorentzian if the frequency profile of the perturbation is a Gaussian or other distribution.

There is also something in Sakurai's book about constant perturbation of the type

$$V(t) = V\theta(t) \quad (13.22)$$

which is a constant perturbation suddenly turned on at some time instant. In modern days this is called quenching. We don't really need time-dependent perturbation theory to do this, because we can just separate the Hamiltonian into the $t < 0$ part and $t > 0$ part. When $t < 0$ the initial state is just $|\alpha\rangle = |n\rangle_0$. If we switch on the perturbation sufficiently quickly, the new state should be still approximately close to the initial state $|\alpha, t = 0_+\rangle = |n\rangle_0$. Now we can expand this state in terms of the new energy eigenstates of $t > 0$ Hamiltonian

$$|\alpha, t = 0_+\rangle = \sum |m\rangle \langle m|n\rangle_0 \quad (13.23)$$

and the transition probability will be just $\mathcal{P}_m = |\langle m|n\rangle_0|^2$. Then we can use time-independent perturbation theory to find the states $|m\rangle$ and work out this transition probability

$$\mathcal{P}_{mn} = \frac{|V_{nm}|^2}{(E_n - E_m)^2} \quad (13.24)$$

But first quenching is by far not so innocent as we describe here, and there is nothing really to do with the Fermi golden rule.

We can also consider the opposite limit where the perturbation is adiabatic and time derivative of the perturbation is very small

$$\hbar \langle m | (\partial V / \partial t) | n \rangle \ll (E_m - E_n)^2 \quad (13.25)$$

Then we can work out the energy eigenstate and eigenvalues at every time instant $|n, t\rangle$ and $E_n(t)$, just as happening in time-independent perturbation theory.

14 Lecture 14

Let's summarize time-dependent perturbation theory where we have a perturbation potential $V(t)$. The most common perturbation we take is the harmonic perturbation, or monochromatic perturbation

$$V(t) = \nu e^{i\omega t} + \nu^\dagger e^{-i\omega t} \quad (14.1)$$

We derived the great formula known as the Fermi golden rule which describes the transition rate

$$W_{fi} = \frac{2\pi}{\hbar} |\nu_{fi}|^2 \delta(E_f - E_i - \hbar\omega) \quad (14.2)$$

We discussed last time various regularizations to interpret this infinity at resonance and investigated how the line can will broaden in various situations.

Let's discuss quenching, which corresponds to abrupt switching of the time dependent potential. Remember we have the time-energy uncertainty relation due to Heisenberg

$$\Delta E \cdot \Delta t \geq \frac{\hbar}{2} \quad (14.3)$$

Recall we also have the uncertainty relation of p and x

$$\langle \Delta x^2 \rangle \langle \Delta p^2 \rangle \geq \frac{\hbar^2}{4} \quad (14.4)$$

Here the quantities make perfect sense because we can define $\Delta x = x - \langle x \rangle$ and the above is a perfect inequality about the expectation values of quantum observables. But note t is not an operator, so we can't interpret the energy-time uncertainty in the same way. Let's consider the time evolution of a stationary state

$$|n, t\rangle = e^{-iE_n t/\hbar} |n\rangle \quad (14.5)$$

and we know that $\langle A \rangle_n$ is constant for any observable A . Now if we consider a linear combination of stationary states $|\alpha\rangle = \sum c_n |n\rangle$ then we have

$$\langle A \rangle_\alpha = \sum_{n,m} c_m^* c_n A_{mn} \exp \left[\frac{i}{\hbar} (E_m - E_n) t \right] \quad (14.6)$$

Now by the generalized commutation relation we get

$$\langle (\Delta A)^2 \rangle_\alpha (\Delta E)^2 \geq \frac{1}{4} |\langle [A, \mathcal{H}] \rangle_\alpha|^2 \quad (14.7)$$

By Heisenberg equation of motion we have

$$\frac{d}{dt} \langle A \rangle_\alpha = \frac{1}{i\hbar} \langle [A, \mathcal{H}] \rangle_\alpha \quad (14.8)$$

so we can define a characteristic time

$$\tau^2 = \frac{\langle (\Delta A)^2 \rangle}{|\langle dA/dt \rangle|^2} \implies \tau \Delta E \geq \frac{\hbar}{2} \quad (14.9)$$

This is how we make sense of the energy-time uncertainty relation.

Consider \mathcal{H}_s as the system Hamiltonian and \mathcal{H}_d to be the detector Hamiltonian. At times $t < 0$ and $t > T$ let $\mathcal{H} = \mathcal{H}_s + \mathcal{H}_d$, so that the system and detector are decoupled. The detector is just another quantum system which we can describe as the state $|d\rangle$, and we know its energy ϵ_d precisely. If the total initial energy is $E_\alpha = E + \epsilon_d$ and the total final energy after time T is $E'_\alpha = E' + \epsilon'_d$ then we will know the difference in energy will be the order of

$$E_\alpha - E'_\alpha \sim \frac{\hbar}{T} \quad (14.10)$$

We can know this from the uncertainty relation, but also from the consideration of time-dependent perturbation theory. So if you don't make your measurements carefully enough, you will introduce a transition between states of this kind of energy separation. For example, if we make the measurements about the transition of a harmonic oscillator between ground state and a discrete state $|n\rangle$, and control the measurement time T such that $\hbar\Delta\omega \sim 1/T$ is much smaller than the energy difference between $|n\rangle$ and its neighboring states, then the probability of its final state to be something else than $|n\rangle$ will be exponentially small. However if we measure the transition to continuum, then the probability of finding neighboring states will be just given by the above uncertainty relation.

Let's consider a perturbation which acts on the state for time T , and define the variable $s = t/T$ which varies from 0 to 1. The adiabatic case is when $T \rightarrow \infty$ and antiadiabatic case is when $T \rightarrow 0$. Let's first consider the antiadiabatic case. The time-evolution operator is

$$U(T) = \exp\left[-\frac{i}{\hbar} \int_0^T \mathcal{H}(t) dt\right] = \exp\left[-\frac{iT}{\hbar} \int_0^1 \mathcal{H}(s) ds\right] \approx 1 - \frac{iT}{\hbar} \int_0^1 \mathcal{H}(s) ds \quad (14.11)$$

In the limit that $T \rightarrow 0$ we have $U(T) \rightarrow 1$. So we will expect that

$$|\alpha, t = 0_-\rangle = |\alpha, t = 0_+\rangle \quad (14.12)$$

This means that under infinitesimal time, the quantum state can't change much under perturbation. The probability of the state not to be in the initial state at time T is

$$\sum_{f \neq i} P_{fi}(T) = \sum_{f \neq i} |\langle f | U(T) | i \rangle|^2 = \langle i | U^\dagger(T) U(T) | i \rangle - |\langle i | U(T) | i \rangle|^2 \quad (14.13)$$

If we plug into the infinitesimal form of the time-evolution operator then we can get

$$\sum_{f \neq i} P_{fi}(T) = \frac{T^2}{\hbar^2} \left[\left\langle i \left| \left(\int_0^1 \mathcal{H}(s) ds \right)^2 \right| i \right\rangle - \left\langle i \left| \int_0^1 \mathcal{H}(s) ds \right| i \right\rangle^2 \right] = \frac{T^2}{\hbar^2} \langle (\Delta \tilde{\mathcal{H}})^2 \rangle \quad (14.14)$$

This is much smaller than 1 provided that

$$T \ll \frac{\hbar}{\sqrt{\langle (\Delta \tilde{\mathcal{H}})^2 \rangle}} \quad (14.15)$$

And this is the antiadiabatic condition.

Let's now consider the adiabatic case $T \rightarrow \infty$. Assume that the Hamiltonian $\mathcal{H}(q_1, \dots, q_p)$ depends on p parameters $\{q_i\}$. Now assume $q_i(s)$ depends on the parameter s and at different parameter configurations we can find the eigenstates of the Hamiltonian

$$\mathcal{H}(q_i) |n(q_i)\rangle = E_n(q_i) |n(q_i)\rangle \quad (14.16)$$

Let us now define the adiabatic states $|n, t\rangle_A$ under the slow time evolution of the parameters q_i , and define the time evolution of these states to be the same as the time-independent Hamiltonian

$$i\hbar \frac{\partial |n, t\rangle_A}{\partial t} = \mathcal{H}(q_i) |n, t\rangle_A = E_n(s) |n, t\rangle_A \quad (14.17)$$

and we can just integrate the equation

$$|n, t\rangle_A = \exp \left[-\frac{iT}{\hbar} \int_0^s E_n(q_i(s')) ds' \right] |n(q_i)\rangle = e^{-i\Omega_n(s)T} |n(q_i)\rangle \quad (14.18)$$

where we define the adiabatic phase

$$\Omega_n(s) = \frac{1}{\hbar} \int_0^s E_n(s') ds' \quad (14.19)$$

Now we can also work out the total time derivative of the adiabatic state as the partial derivative of the explicit time dependence t and that from q

$$\frac{d}{dt} |n, t\rangle_A = \frac{\partial}{\partial t} |n, t\rangle_A + \frac{1}{T} \frac{\partial q_i}{\partial s} \frac{\partial}{\partial q_i} |n, t\rangle_A = E_n(s) |n, t\rangle_A + \frac{1}{T} \frac{\partial q_i}{\partial s} \frac{\partial}{\partial q_i} |n, t\rangle_A \quad (14.20)$$

In the limit $T \rightarrow \infty$ the second term will be small and will be treated as perturbation. Now we can expand the time-evolution of any state in terms of these adiabatic states

$$|n, t\rangle = c_n(s) |n, t\rangle_A + \sum_{k \neq n} c_k(s) e^{-iT[\Omega_n(s) - \Omega_k(s)]} |k, t\rangle_A \quad (14.21)$$

We will find that if $c_n(0) = 1$, then afterwards we have

$$c_n(s) = e^{i\gamma(s)} + O(1/T), \quad c_k(s) = O(1/T) \quad (14.22)$$

so in the adiabatic limit $T \rightarrow \infty$ the state will not evolve, and we have $|n\rangle \rightarrow |n\rangle$ with probability 1. We need an expression for $\gamma(s)$. It satisfies the equation

$$-\frac{d\gamma}{ds} |n(q_i)\rangle = i \frac{\partial q_i}{\partial s} \frac{\partial}{\partial q_i} |n(q_i)\rangle \quad (14.23)$$

so we can integrate this equation to get

$$\gamma(s) = \int_{q(0)}^{q(s)} i \langle n(q_i) | \partial / \partial q_j | n(q_i) \rangle dq_j \quad (14.24)$$

This is the equation for the Berry phase. The integrand can be worked out as

$$\langle k(q_j) | \partial / \partial q_i | n(q_j) \rangle = \frac{\langle k(q_j) | \partial \mathcal{H} / \partial q_i | n(q_j) \rangle}{E_k - E_n} \quad (14.25)$$

And now we have an explicit equation for $c_k(s)$ which is

$$c_k(s) = -\frac{i\hbar}{T} \frac{c_n(s)}{[E_k(s) - E_n(s)]^2} \frac{\partial q_i}{\partial s} \langle k | \partial \mathcal{H} / \partial q_i | n \rangle \quad (14.26)$$

This equation doesn't give us $c_k(0) = 0$ unless the derivative vanishes at $s = 0$. But we are safe as long as we stay away from $s = 0$. The validity of this formula, or the adiabatic approximation depends on

$$\hbar \left(\frac{\partial \mathcal{H}}{\partial t} \right)_{kn} \ll (E_k - E_n)^2 \quad (14.27)$$

So the time derivative of the Hamiltonian should be smaller than the smallest gap between the energy levels.

If we consider the case of cyclic perturbation, which means $q(s = 1) = q(s = 0)$, then the Berry phase γ will give us a nontrivial geometric phase factor

$$\gamma_c = \oint i \langle n | \partial / \partial q_i | n \rangle dq_i \quad (14.28)$$

This looks very similar to the phase due to a gauge field. This phase factor can be detected in similar way as the Aharonov-Bohm effect. The integrand in the above expression is very much like a vector potential, but it lives in a fictitious configuration space.

15 Lecture 15

15.1 Identical Particles

Today we are going to start a new topic, which is about identical particles in quantum mechanics. In real world we know that it is impossible for two objects to occupy one place. This corresponds to the fact that in classical mechanics two particles can't share the same space point or phase space point. However we know that in quantum mechanics we have quantum states in stead of points in phase space, so the question we want to ask is that is it possible that two quantum particles occupy the same state?

The answer depends on if they are different particles. Now what do we mean different? We can distinguish a photon from an electron, and vice versa, but we can't truly distinguish two electrons from each other, and same for two photons. This is a very deep philosophical problem, because classical particles are always distinguishable, and the fact that indistinguishable objects exist is an amazing prediction of quantum physics. For classical particles, even if they look exactly the same, we could still in principle trace their trajectory and label them accordingly and keep track of the label consistently. However, this is not possible with quantum states, because we can't do the same trajectory-tracking.

Let's consider an example. Consider two independent simple harmonic oscillators

$$\mathcal{H} = \frac{1}{2m}p_1^2 + \frac{1}{2}m\omega^2x_1^2 + \frac{1}{2m}p_2^2 + \frac{1}{2}m\omega^2x_2^2 \quad (15.1)$$

We have seen this and remember that if we replace p_1, p_2 with p_x and p_y then we have a 2D harmonic oscillator. We know the energy eigenstates are just the tensor product states

$$|k, n\rangle = |k\rangle \otimes |n\rangle, \quad E_{k,n} = \left(k + \frac{1}{2}\right)\hbar\omega + \left(n + \frac{1}{2}\right)\hbar\omega \quad (15.2)$$

Let's for simplicity only consider two states. The ground state is just $|0, 0\rangle$ which is nondegenerate. The first excited states are $|0, 1\rangle$ and $|1, 0\rangle$ and they are doubly degenerate, so the first excited state can be described by any linear combination of these two states, with appropriate normalization. Suppose we have $|\alpha\rangle = \lambda|0, 1\rangle + \mu|1, 0\rangle$ then we can measure the dot product $\langle \mathbf{x}_1 \cdot \mathbf{x}_2 \rangle_\alpha$. This product is sensitive to the relative coefficients λ and μ because

$$\langle \mathbf{x}_1 \cdot \mathbf{x}_2 \rangle_\alpha = \frac{\hbar}{2m\omega}(\lambda^*\mu + \mu^*\lambda) = \begin{cases} 0, & \text{if } \lambda = 1, \mu = 0 \\ \frac{\hbar}{2m\omega}, & \text{if } \lambda = \mu = \frac{1}{\sqrt{2}} \end{cases} \quad (15.3)$$

Now let's consider the interesting case of exchanging two particles. Let's define an operator P_{12} which does the exchange as

$$P_{12}|k, n\rangle = |n, k\rangle, \quad P_{12}^2 = 1 \quad (15.4)$$

If the particles are indistinguishable we should expect that the Hamiltonian is symmetric under this exchange

$$[P_{12}, \mathcal{H}] = 0 \quad (15.5)$$

But we want more than this. We want to require that any two identical particles are truly indistinguishable. Therefore we need to guarantee a stronger statement that any physical state is an eigenstate of P_{12} . Note this stronger requirement depends on our previous condition. Remember for any symmetry operator A if $[A, \mathcal{H}] = 0$ and there is degeneracy, then the eigenstates of \mathcal{H} may not be the eigenstates of A because A can mix the states inside one degenerate eigenspace.

Now if we want to deal with the eigenstates of the permutation operator, then the eigenvalue could only be ± 1 because we need $P_{12}^2 = 1$. This can be generalized to N identical particles and we require each physical state is an eigenstate of each of the permutation operators, and the permutations form a group S_N .

Suppose we have an arbitrary state $|\alpha\rangle = \sum c_{k,n} |k, n\rangle$ then we have the following

$$P_{12} |\alpha\rangle = \sum_{k,n} c_{k,n} |n, k\rangle = \pm |\alpha\rangle \implies c_{k,n} = \pm c_{n,k} \quad (15.6)$$

Now this system is no longer the same as one particle in a 2D harmonic oscillator potential in that the coefficients of arbitrary physical states are severely restricted from a continuum to two discrete states. So essentially our requirement of indistinguishable particles is restricting our Hilbert space.

Now according to different eigenvalues of the permutation operator, we have two categories of particles, or two types of quantum statistics. First corresponds to $P_{12} = 1$ which we call bosons, and they obey Bose-Einstein statistics. The second corresponds to $P_{12} = -1$ which we call fermions, and they obey Fermi-Dirac statistics. From relativistic quantum theory it is found that the statistical properties of particles are closely related to their spin, where integer spins correspond to bosons and half-integer spins correspond to fermions. This connection between spin and statistics is highly nontrivial, and is one of the celebrated results of quantum field theory.

Let's recap. If we have two 1-particle states $|\alpha\rangle$ and $|\alpha'\rangle$. We can form 4 linearly independent states from these two states, namely

$$|\alpha, \alpha\rangle, \quad |\alpha', \alpha'\rangle, \quad \frac{1}{\sqrt{2}}(|\alpha, \alpha'\rangle + |\alpha', \alpha\rangle), \quad \frac{1}{\sqrt{2}}(|\alpha, \alpha'\rangle - |\alpha', \alpha\rangle) \quad (15.7)$$

For classical distinguishable particles all states are possible. For bosons only the first three states are possible whereas for fermions only the last state is possible. Also following from this is the celebrated Pauli exclusion principle which essentially says that

$$c_{n,n} = 0 \quad (15.8)$$

so that no two fermions can occupy the same quantum state.

Let's consider two spin-1/2 particles, or two electrons to be specific. If we don't consider spin-orbit coupling, so that the spin and orbital degrees of freedom are decoupled, then we can write the two-particle wave function as

$$\Psi = \varphi(\mathbf{x}_1, \mathbf{x}_2)\chi \quad (15.9)$$

where χ is the spin wave function. The χ function is not arbitrary. For symmetric χ we can have

$$|-, -\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad |+, +\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |+, -\rangle_+ = \frac{1}{\sqrt{2}} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] \quad (15.10)$$

These states correspond to total spin 1, and the three states correspond to $m_s = \mp 1$ and $m_s = 0$ respectively. And for antisymmetric χ we can only have

$$|+, -\rangle_- = \frac{1}{\sqrt{2}} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] \quad (15.11)$$

and this is the state of total spin zero. Now the permutation operator is the product of the permutation in orbital part and spin part

$$P_{12} = P_{12}^{\text{orb}} \otimes P_{12}^{\text{spin}} \quad (15.12)$$

So we need to split the orbital wave function into symmetric and antisymmetric parts and assume the two 1-particle wave functions are φ_A and φ_B we have

$$\varphi^{\pm}(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}} [\varphi_A(\mathbf{x}_1)\varphi_B(\mathbf{x}_2) \pm \varphi_A(\mathbf{x}_2)\varphi_B(\mathbf{x}_1)] \quad (15.13)$$

So the triplet state must be

$$\Psi_t = \chi_t \varphi^{-}(\mathbf{x}_1, \mathbf{x}_2) \quad (15.14)$$

and the singlet state must be the opposite.

Now let's calculate some quantities of interest. Suppose $d\mathbf{x}_1$ and $d\mathbf{x}_2$ are volume elements, the probability of finding the two particles at \mathbf{x}_1 and \mathbf{x}_2 will be

$$P = \frac{1}{2} \left\{ |\varphi_A(\mathbf{x}_1)|^2 |\varphi_B(\mathbf{x}_2)|^2 + |\varphi_B(\mathbf{x}_1)|^2 |\varphi_A(\mathbf{x}_2)|^2 \pm 2\text{Re} [\varphi_A(\mathbf{x}_1)\varphi_B(\mathbf{x}_2)\varphi_A^*(\mathbf{x}_2)\varphi_B^*(\mathbf{x}_1)] \right\} d\mathbf{x}_1 d\mathbf{x}_2 \quad (15.15)$$

The last term arises purely due to exchange effects and for $\mathbf{x}_1 = \mathbf{x}_2$ we have a probability twice as big as the classical one, and for triplet state we have 0. This gives rise to a kind of interference between the particle wave functions.

As another example we consider two hydrogen atoms with fixed nuclei and two 1-particle orbital states A and B with energies E_A and E_B respectively. Suppose the two atomic electrons interact with each other, and we know the interaction energy is

$$\mathcal{H}_{\text{int}} = \frac{e^2}{|\mathbf{x}_1 - \mathbf{x}_2|} \quad (15.16)$$

then the total energy would be

$$E = E_A + E_B + \Delta_{\pm} \quad (15.17)$$

Now the Δ_{\pm} term is sensitive to whether the electrons are in singlet or triplet state. Recall from first order perturbation theory we have

$$\Delta_{pm}^{(1)} = \langle \pm | \mathcal{H}_{\text{int}} | \pm \rangle = I \pm J - S(S+1)J \quad (15.18)$$

where S is the total spin of the state, and

$$I = \int d\mathbf{x}_1 d\mathbf{x}_2 \frac{e^2}{|\mathbf{x}_1 - \mathbf{x}_2|} |\varphi_A(\mathbf{x}_1)|^2 |\varphi_B(\mathbf{x}_2)|^2 \quad (15.19)$$

and

$$J = \int d\mathbf{x}_1 d\mathbf{x}_2 \frac{e^2}{|\mathbf{x}_1 - \mathbf{x}_2|} \varphi_A(\mathbf{x}_1)\varphi_B(\mathbf{x}_2)\varphi_A^*(\mathbf{x}_2)\varphi_B^*(\mathbf{x}_1) \quad (15.20)$$

Note that the triplet state has lower energy. This has nothing to do with magnetic or spin-orbit coupling, but totally comes from exchange interactions.

16 Lecture 16

16.1 Identical Particles Continued

We started a semi-philosophical discussion about what is distinguishable and indistinguishable. We converted the discussion into formal mathematical language, introducing the permutation operator P_{12} which permutes two particles. The important statement is that *all* physical states should be eigenstates of all permutation operators. This is a very strong statement, and the only way to satisfy this is to reduce the Hilbert space. This reduction has dramatic consequences. Remember we required that $P_{12}^2 = 1$ so that

$$P_{12} |\alpha\rangle = \pm |\alpha\rangle \quad (16.1)$$

From this we concluded that there are two types of statistics, one being Bose-Einstein and the other being Fermi-Dirac. For bosons we have $+$ and for fermions we have $-$. We also have the Pauli exclusion principle for fermions that basically tells that two fermions can't occupy the same quantum state. The last statement we made was due to Dirac, and he found that particles with integer spin are bosons and particles with half-integer spin are fermions. This theorem of spin and statistics is one of the triumphs of the relativistic quantum mechanics.

Let's consider the following problem. Suppose we have N particles, so we need N states ψ_1, \dots, ψ_N and N coordinates x_1, \dots, x_N to describe them. Then what we will tell is that if these particles are not identical, then we have to put one of them in state ψ_i and the other in ψ_j etc. The wave function for distinguishable particles is

$$\Psi(x_1, \dots, x_n) = \prod_{i=1}^N \psi_i(x_j) \quad (16.2)$$

Note we can permute the particles and we have $N!$ different states. However when we have indistinguishable particles we don't have the luxury of $N!$ different states. For bosons the wave function should be symmetrized, which means the bosonic wave function should be written as

$$\Psi_B = \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} \prod \psi_{\sigma(i)}(x_j) \quad (16.3)$$

For fermions the wave function should be antisymmetric, and we need to consider the set $\{\psi_i(x_j)\}$ as a matrix, and our wave function should be proportional to the determinant of this matrix, which we call the Slater determinant

$$\Psi_F \propto \text{Det}(\psi_i(x_j)) \quad (16.4)$$

We can consider a spin-1/2 system with two particles. The Hamiltonian is written as

$$\mathcal{H} = \mathcal{H}^{(1)} + \mathcal{H}_{\text{int}} \quad (16.5)$$

where the first term is the one-particle Hamiltonians

$$\mathcal{H}^{(1)} = \mathcal{H}_A^{(1)} + \mathcal{H}_B^{(1)}, \quad \mathcal{H}_{A,B}^{(1)} = -\frac{\hbar^2 \nabla^2}{2m} + V_{A,B} \quad (16.6)$$

Now if there is no interaction, we just solve two one-particle problems and we will have φ_A and φ_B which are two orbital states with energies E_A and E_B . The energy in the absence of interactions is just

$$E = E^{(0)} = E_A + E_B \quad (16.7)$$

and the wave function should be

$$\Psi^{(0)}(x_1, x_2) = \frac{1}{\sqrt{2}} \text{Det} \begin{pmatrix} \psi_A(x_1) & \psi_A(x_2) \\ \psi_B(x_1) & \psi_B(x_2) \end{pmatrix} \quad (16.8)$$

where $\psi_{A,B} = \varphi_{A,B} \chi_{A,B}$ includes the spin wave functions as well. Remember that if the spin states form the singlet state with $S = 0$ then the orbital wave function should be symmetric, and if the spin states form the triplet state with $S = 1$ then the orbital wave function should be antisymmetric. We found last time that the first order correction due to an interaction term like

$$\mathcal{H}_{\text{int}} = U(|\mathbf{x}_1 - \mathbf{x}_2|) \quad (16.9)$$

will be like

$$E = E^{(0)} + \Delta, \quad \Delta^{(1)} = I + (1 - S(S+1))J \quad (16.10)$$

where I and J are interaction integrals. Note that this energy correction is spin-dependent, even when the interaction does not involve spin at all. Note the triplet state has lower energy as $J > 0$. So usually the triplet state is the ground state. However this is not always the case, for example when the first order perturbation vanishes it often turns out that the singlet state has lower second order energy correction and therefore will become the ground state. Note the form of the interaction is like \mathbf{S}^2 so we can write it as

$$\mathbf{S}^2 = (\mathbf{S}_1 + \mathbf{S}_2)^2 = \frac{3}{2} + 2\mathbf{S}_1 \cdot \mathbf{S}_2 \quad (16.11)$$

So the exchange interaction is mainly proportional to $\mathbf{S}_1 \cdot \mathbf{S}_2$. However the magnetic dipole-dipole interaction is more complicated than this, which goes like

$$\mathbf{S}_1 \cdot \mathbf{S}_2 - \frac{2(\mathbf{S}_1 \cdot \mathbf{x}_{12})(\mathbf{S}_2 \cdot \mathbf{x}_{12})}{|\mathbf{x}_{12}|^2} \quad (16.12)$$

This is different from and usually lower than the exchange interaction. However if we want to consider ferromagnets then we need to consider the exchange interaction, otherwise there is no way we can have a compass, as the exchange interaction allows spin to couple with the spatial wave function of the electrons.

We can also look at the first term I in the first order correction to energy. It is called the Hartree term. Let's consider the density of electron B as $\rho_B(\mathbf{x}) = |\varphi_B(\mathbf{x})|^2$ and the potential of the electron A due to the interaction with electron B is

$$V_H(\mathbf{x}) = \int d\mathbf{x}' \rho_B(\mathbf{x}') U(|\mathbf{x} - \mathbf{x}'|) \quad (16.13)$$

and integrating this potential with φ_A will give exactly the integral I . This is like a mean field approximation, using the field of electron B as a background. The exchange term J comes from a potential not so familiar which is the Fock potential

$$\langle \mathbf{x} | V_F | A \rangle = \int d\mathbf{x}' V_F(\mathbf{x}, \mathbf{x}') \varphi_A(\mathbf{x}') \quad (16.14)$$

where the integration kernel is

$$V_F(\mathbf{x}, \mathbf{x}') = U(|\mathbf{x} - \mathbf{x}'|) \varphi_A^*(\mathbf{x}) \varphi_B(\mathbf{x}) \varphi_B^*(\mathbf{x}') \quad (16.15)$$

This is Hartree-Fock approximation to the zeroth order. If we want to do better we need to do variational principle on our Slater determinant and solve the equations iteratively, which usually yields very good agreement with experimental data.

Let's consider a Bose gas with no interaction. The Hamiltonian is just

$$\mathcal{H} = \sum_i \mathcal{H}_i^{(1)}, \quad \mathcal{H}_i^{(1)} = -\frac{\hbar^2 \nabla_i^2}{2m} + V(\mathbf{x}_i) \quad (16.16)$$

If we have N bosons then what is the ground state of the system? Suppose the one-particle energies are E_i where i is the energy level, then the total energy of the system is just

$$E = \sum_i E_i n_i \quad (16.17)$$

where n_i is the number of particles in state i , with constraint $\sum n_i = N$, and the sum runs over all the one-particle states. Now the ground state is obviously when $n_0 = N$, which means all the particles are in the one-particle ground state.

Let's now consider the transition from state $|i\rangle$ to a lower energy state $|j\rangle$. We know how to do this under weak perturbation $V(t)$ by the Fermi golden rule

$$P(i \rightarrow j) \propto |V_{ij}|^2 = |\langle i | V | j \rangle|^2 \quad (16.18)$$

Now what if the final state is not empty but occupied by N particles? How then do we calculate the transition probability? Now the initial state with one particle at $|i\rangle$ and N particles at $|j\rangle$ should be written as

$$|I\rangle = \frac{1}{\sqrt{N+1}} (|i, j, \dots, j\rangle + |j, i, j, \dots, j\rangle + \dots + |j, \dots, j, i\rangle) \quad (16.19)$$

The final state is just all the states being $|j\rangle$ so it is

$$|F\rangle = |j, \dots, j\rangle \quad (16.20)$$

where there are $N+1$ j 's in the ket. Now the transition probability is

$$P(I \rightarrow F) = |V_{FI}|^2 = |\langle F | V | I \rangle|^2 = (N+1) |V_{ij}|^2 \quad (16.21)$$

Now the probability is amplified by a factor equal to the population. This is important to the mechanism of the laser. Suppose we have N two level systems and all of them are in the excited state. If one atom undergoes a transition to the lower state then it will emit photon and the photon will run away. But if we put all of them in a cavity then the photon will not run away and stimulate the other atoms to also undergo the transition. By the above amplification, we know that when there are more atoms in the ground state the probability of transition gets larger, so this creates an avalanche. In addition the light emitted from the cavity is coherent, and this is the mechanism of the laser.

Let's now turn to fermions. Consider a Fermi gas of N particles and similar to above we have the total energy $E = \sum \epsilon_i n_i$. Now we want to ask again what is the ground state. Suppose all the one-particle spectra are the same, however we can only put two particles in the same state if the particle has spin 1/2. So we have a maximum energy state i_{\max} and we have

$$n_i = \begin{cases} 2s + 1, & \text{if } i < i_{\max} \\ 0, & \text{if } i > i_{\max} \end{cases} \quad (16.22)$$

We call the maximum energy corresponding to the state $|i_{\max}\rangle$ by μ and this quantity plays the role of chemical potential. If we bring a fermion from infinity into the system this is exactly the energy we need to pay to introduce the additional fermion.

We can define the density of states $\rho(\epsilon)$. Obviously we would expect that

$$\rho(\epsilon) = \sum_{i=0}^{\infty} \delta(\epsilon - \epsilon_i) \quad (16.23)$$

But when the energy spacing is small, we can approximate it using a smooth function. The total number of particles will then be

$$N = \int_0^{\infty} \rho(\epsilon) n(\epsilon) d\epsilon, \quad E = \int_0^{\infty} \rho(\epsilon) n(\epsilon) \epsilon d\epsilon \quad (16.24)$$

If we have a 3D cubic box with dimension $L \times L \times L$ then the momenta in each direction should be quantized according to $p_i = 2\pi\hbar k_i/L$. The energy will then be

$$\epsilon(\mathbf{p}) = \frac{p_x^2 + p_y^2 + p_z^2}{2m} \quad (16.25)$$

and the maximum energy $\epsilon(\mathbf{p}_F) = \mu$ where \mathbf{p}_F is called Fermi momentum. All the states with momentum smaller than \mathbf{p}_F are occupied and all states with larger momenta are empty. Then total number of energy states inside the momentum sphere with radius p_F will be

$$N = (2s + 1) \frac{L^3}{(2\pi\hbar)^3} \frac{4\pi}{3} p_F^3 = \frac{2s + 1}{6\pi^2} \frac{(Lp_F)^3}{\hbar^3} \quad (16.26)$$

so we can work out the Fermi momentum as

$$p_F = \hbar \left(\frac{6\pi^2}{2s + 1} \right)^{1/3} \left(\frac{N}{V} \right)^{1/3} \quad (16.27)$$

so the Fermi momentum is proportional to the number density to the 1/3 power.

Now suppose all my fermions are in the harmonic potential

$$V = \frac{m\omega^2}{2} |\mathbf{x}|^2 \quad (16.28)$$

then the energy states are

$$\epsilon_n = \left(n + \frac{3}{2} \right) \hbar\omega \quad (16.29)$$

These states are degenerate and the degeneracy is

$$g_n = \frac{(n + 2)(n + 1)}{2} \quad (16.30)$$

So the total number of particles and ground state energy are

$$N = (2s + 1) \sum_{n=0}^{n_{\max}} g_n, \quad E_0 = (2s + 1) \sum_{n=0}^{n_{\max}} \epsilon_n g_n \quad (16.31)$$

From the above expression of g_n we can calculate explicitly that

$$n_{\max} = \left(\frac{6N}{2s+1} \right)^{1/3}, \quad E_0 = \xi N^{4/3} \left(\frac{6}{2s+1} \right)^{1/3} \quad (16.32)$$

We can use a variational principle that $\langle H \rangle \geq E_0$ for any state and if we carry out this exercise we will get

$$\langle p^2 \rangle + m^2 \omega^2 \langle x^2 \rangle - 2\xi N^{4/3} m \hbar \omega \geq 0 \quad (16.33)$$

In the end we can get a relation of the type

$$\langle p^2 \rangle \langle x^2 \rangle \geq \xi^2 N^{2/3} \hbar^2 \quad (16.34)$$

17 Lecture 17

17.1 Second Quantization

Recall our results from simple harmonic oscillator. We know the Hamiltonian very well so no need to repeat here. The eigenstates are $|n\rangle$ and corresponding energies are

$$E_n = \left(n + \frac{1}{2}\right) \hbar\omega \quad (17.1)$$

We got a lot of information from the ladder operators a^\dagger and a . Remember we have

$$a|n\rangle = \sqrt{n}|n-1\rangle, \quad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \quad (17.2)$$

There are a lot of things we can do with the ladder operators. There is one interesting operator which is the number operator, and it diagonalizes the Hamiltonian

$$N = a^\dagger a, \quad \mathcal{H} = \hbar\omega \left(N + \frac{1}{2}\right) \quad (17.3)$$

We will see that these harmonic oscillators are like shelves, on to which we can put excitations. The operator a^\dagger will create an excitation on an empty “shelf”, and if we operate it n times we will get n excitations each with energy $\hbar\omega$. We can think of the excitations as “particles”, each with a definite amount of energy, which are created by the operator a^\dagger over the vacuum background $|0\rangle$.

Let’s try to make this intuition precise. We define an abstract vacuum state $|0\rangle$, and define a family of states $|\alpha\rangle$ to be one particle states, where

$$\langle x|\alpha\rangle = \varphi_\alpha(x) \quad (17.4)$$

is the wave function of the one-particle state. This is what we have done all the time, and this is called “first quantization”. Now let’s define an operator a_α^\dagger which operates on the vacuum states and creates a one-particle state

$$a_\alpha^\dagger|0\rangle = |\alpha\rangle \quad (17.5)$$

Note this is not field theory yet. The difference between field theory and quantum mechanics is that in field theory the number of particles is usually not fixed, whereas in quantum mechanics usually we have a given set of particles and we consider their evolution.

Now let’s consider the state of two particles and define it as

$$|\alpha, \beta\rangle = a_\alpha^\dagger|\beta\rangle = a_\alpha^\dagger a_\beta^\dagger|0\rangle \quad (17.6)$$

It is easily generalized to n particles, where we need to apply the operator n times. Now what if we have identical particles? Let’s do fermions first. We know from the properties of fermions that if we try to create two fermions at the same state, we will fail and get zero

$$a_\alpha^\dagger a_\alpha^\dagger|0\rangle = 0 \quad (17.7)$$

What we also know is that for fermions

$$|\alpha, \beta\rangle = -|\beta, \alpha\rangle, \quad a_\alpha^\dagger a_\beta^\dagger|0\rangle = -a_\beta^\dagger a_\alpha^\dagger|0\rangle \quad (17.8)$$

These two conditions are automatically satisfied if we require that $a_\alpha^\dagger a_\beta^\dagger = -a_\beta^\dagger a_\alpha^\dagger$, or in our notation

$$\{a_\alpha^\dagger, a_\beta^\dagger\} = 0 \quad (17.9)$$

So far so good. But remember that all these ladder operators come in pairs, so we need to consider annihilation operators. Note that

$$\langle \alpha | a_\alpha^\dagger | 0 \rangle = 1 \implies \langle 0 | a_\alpha | \alpha \rangle = 1, \quad a_\alpha | \alpha \rangle = 0 \quad (17.10)$$

So the hermitian conjugate of the creation operator is indeed the annihilation operator. From the above anticommutation relation of the creation operators, we can show that

$$\{a_\alpha, a_\beta\} = 0 \quad (17.11)$$

In order to get a full set of algebra, we need to get the anticommutator between a_β^\dagger and a_α . We will operate the anticommutator to the state $|\alpha\rangle$ and

$$\{a_\alpha, a_\beta^\dagger\} |\alpha\rangle = a_\alpha |\beta, \alpha\rangle + a_\beta^\dagger |0\rangle = -|\beta\rangle + |\beta\rangle = 0 \quad (17.12)$$

However is this the full truth? Not exactly because the above is correction only for $\alpha \neq \beta$. What happens if $\alpha = \beta$?

$$\{a_\alpha, a_\alpha^\dagger\} |\alpha\rangle = 0 + |\alpha\rangle = |\alpha\rangle \quad (17.13)$$

So this tells us that the anticommutator is not identically zero. The correct relation is

$$\{a_\alpha, a_\beta^\dagger\} = \delta_{\alpha\beta} \mathbf{1} \quad (17.14)$$

where $\mathbf{1}$ is the identity operator on the Hilbert space. We now have the complete anticommutation relations for fermions. Now what is the operator $a^\dagger a$? If the state is not in $|\alpha\rangle$ state then this operator will give zero, but if the state is $|\alpha\rangle$ then the operator will annihilate it and create it again to get back $|\alpha\rangle$. So this is just the occupation number

$$a_\alpha^\dagger a_\alpha = n_\alpha, \quad N = \sum_\alpha a_\alpha^\dagger a_\alpha = \sum_\alpha n_\alpha \quad (17.15)$$

and N is the operator for the total number of particles, and n_α is either 1 or 0.

We know that we can choose our basis for our Hilbert space, and in the above discussion we have not specified that $|\alpha\rangle$ depend on any basis, so the above works for any basis. Suppose $\{|\alpha\rangle\}$ where $\langle x|\alpha\rangle = \varphi_\alpha(x)$ is basis 1 and that $\{|j\rangle\}$ where $\langle x|j\rangle = \phi_j(x)$ is basis 2. We can always change basis

$$\phi_j(x) = \sum_\alpha \langle \alpha|j\rangle \varphi_\alpha(x) \quad (17.16)$$

Now in our new life we want to require that $|j\rangle = c_j^\dagger |0\rangle$ where c_j^\dagger creates a particle at state $|j\rangle$. Then we have a relation between the creation operators

$$c_j^\dagger = \sum_\alpha \langle \alpha|j\rangle a_\alpha^\dagger \quad (17.17)$$

However we need to check that this linear combination preserves all the anticommutation relations above. This is easy to check because this is just a linear transformation, and the anticommutator is bilinear. This is how the creation operators transform when we change basis.

Let's consider an example. Let $|j\rangle = |x\rangle$ where x is a collective index of the position and spin operators. Suppose there is some creation operator which acting on the vacuum state will give us this state

$$\Psi^\dagger(x) |0\rangle = |x\rangle, \quad \Psi^\dagger(x) = \sum_{\alpha} \langle \alpha | x \rangle a_{\alpha}^{\dagger} = \sum_{\alpha} \psi_{\alpha}^*(x) a_{\alpha}^{\dagger} \quad (17.18)$$

This is just the operator which creates a particle at point x , and similarly we can define the annihilation operator

$$\Psi(x) = \sum_{\alpha} \varphi_{\alpha}(x) a_{\alpha} \quad (17.19)$$

These operators are called *field operators*, and these easily generalize to quantum field theory. The operator $\Psi^\dagger(x)\Psi(x)$ is just the density of particles at point x and the total number operator is just

$$N = \int dx \Psi^\dagger(x)\Psi(x) \quad (17.20)$$

Similar construction can also be done in the momentum space and we can get the field operators in the momentum space.

Now that we have considered fermions, let's quickly go over bosons. The construction is almost identical, but there is some key differences. We know bosons can share the same state so we need to define many-particle states $|n_1, n_2, \dots, n_{\alpha}, \dots\rangle$. In particular the vacuum state is the state where all the excitations are zero

$$|0\rangle = |0, 0, \dots, 0, \dots\rangle \quad (17.21)$$

The creation operator will give us

$$a_{\alpha}^{\dagger} |0\rangle = |\dots, n_{\alpha} = 1, \dots\rangle, \quad a_{\alpha}^{\dagger} |\dots, n_{\alpha}, \dots\rangle \propto |\dots, n_{\alpha} + 1, \dots\rangle \quad (17.22)$$

And for annihilation operators

$$a_{\alpha} |0\rangle = 0, \quad a_{\alpha} |\dots, n_{\alpha}, \dots\rangle \propto |\dots, n_{\alpha} - 1, \dots\rangle \quad (17.23)$$

How can we calculate the proportionality constants? Similar to the harmonic oscillator case which is just one state, we can evaluate

$$\langle \dots, n_{\alpha}, \dots | a_{\alpha}^{\dagger} a_{\alpha} | \dots, n_{\alpha}, \dots \rangle = n_{\alpha} \implies a_{\alpha} |\dots, n_{\alpha}, \dots\rangle = \sqrt{n_{\alpha}} |\dots, n_{\alpha} - 1, \dots\rangle \quad (17.24)$$

Similarly we have

$$a_{\alpha}^{\dagger} |\dots, n_{\alpha}, \dots\rangle = \sqrt{n_{\alpha} + 1} |\dots, n_{\alpha} + 1, \dots\rangle \quad (17.25)$$

Everything is more or less identical to the simple harmonic oscillator. So indeed bosons can be think of excitations of simple harmonic oscillators. For bosons the algebraic relations between the creation and annihilation operators are replaced by commutators because we require our wave function to be symmetric rather than antisymmetric. So we have

$$[a_{\alpha}^{\dagger}, a_{\beta}^{\dagger}] = 0, \quad [a_{\alpha}, a_{\beta}] = 0, \quad [a_{\alpha}, a_{\beta}^{\dagger}] = \delta_{\alpha\beta} \mathbf{1} \quad (17.26)$$

What is this system good for? The formalism is good for dealing with many identical particles. There are many operators we are interested in when we are considering the physics of the system of particles, and among them the more interesting ones are the additive operators, for example the total momentum operator

$$\mathbf{P} = \sum_i \mathbf{p}_i \quad (17.27)$$

or the total kinetic energy operator

$$- \sum_i \frac{\hbar^2}{2m} \nabla_i^2 \quad (17.28)$$

or the total external potential

$$V = \sum_i V(\mathbf{x}_i) \quad (17.29)$$

In these examples the total operator is just the sum of the one particle operators, so it is sufficient to use one particle operators to describe the whole operator. Let's denote one of the total operator above by R and one particle operator by R_1 , and we want to consider this operator in our creation and annihilation operator basis. We have

$$R = \sum_{\alpha, \beta} \langle \alpha | R_1 | \beta \rangle a_\alpha^\dagger a_\beta \quad (17.30)$$

For example let $|\alpha\rangle = |p\rangle$ and $R_1 = \hat{p}$. Then we have

$$\langle p | \hat{p} | p' \rangle = p\delta(p - p'), \quad \hat{P} = \sum_{p, p'} p\delta(p - p') a_p^\dagger a_p = \sum_p p \hat{n}_p \quad (17.31)$$

In particular if $|\alpha\rangle$ are eigenstates of operator R_1 with eigenvalues r_α then we have

$$R = \sum_\alpha r_\alpha a_\alpha^\dagger a_\alpha \quad (17.32)$$

If all the operators are just as above, then there will be no interaction between two particles. Now let's consider interaction between just two particles with interaction potential $u(x_1, x_2)$. The total interaction should be

$$U = \frac{1}{2} \sum_{i \neq j} u(x_i, x_j) \quad (17.33)$$

Now we can introduce the matrix element

$$u_{\alpha\beta, \gamma\delta} = \int dx_1 dx_2 \varphi_\alpha^*(x_1) \varphi_\beta^*(x_2) \varphi_\gamma(x_1) \varphi_\delta(x_2) u(x_1, x_2) \quad (17.34)$$

Note this matrix element is unsymmetrized. The interaction operator becomes

$$U = \frac{1}{2} \sum_{\alpha\beta\delta\gamma} u_{\alpha\beta, \delta\gamma} a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma \quad (17.35)$$

For interactions between 3 or more particles, just more powers of a^\dagger and a are included in the interaction potential, but the form remains the same. Note the introduction of creation and annihilation operators

automatically takes care of symmetrization or antisymmetrization. For example if we have fermions we have

$$U = \frac{1}{2} \sum_{\alpha\beta\delta\gamma} u_{\alpha\beta, [\delta\gamma]} a_{\alpha}^{\dagger} a_{\beta}^{\dagger} a_{\delta} a_{\gamma} \quad (17.36)$$

Let's consider an example where $u(x_1, x_2) = u(x_2 - x_1)$. We can define the center-of-mass coordinates

$$x = x_2 - x_1, \quad X = (x_1 + x_2)/2 \quad (17.37)$$

If the one-particle states are just harmonic waves, then we have the matrix element

$$\begin{aligned} u_{\alpha\beta, \gamma\delta} &= \frac{1}{(2\pi)^6} \int dx dX \exp[-ik_{\alpha}(X - x/2) - ik_{\beta}(X + x/2) + ik_{\gamma}(X - x/2) + ik_{\delta}(X + x/2)] u(x) \\ &= \delta(k_{\alpha} + k_{\beta} - k_{\gamma} - k_{\delta}) \tilde{u}(k_{\gamma} - k_{\alpha}) \end{aligned} \quad (17.38)$$

where \tilde{u} is the Fourier transform of u . The delta function is just an enforcement of momentum conservation.

18 Lecture 18

18.1 Second Quantization Reviewed

Last time we discussed the formalism of second quantization. We introduced the one particle states $|\alpha\rangle$ and corresponding ladder operators a_α^\dagger and a_α . Then we introduced the many particle states which are built up from the vacuum state $|0\rangle$ by repeatedly applying the one-particle creation operator. In order to satisfy the correct statistics for identical particles, we imposed the algebraic relation between the creation and annihilation operators. For fermions we have

$$\{a_\alpha^\dagger, a_\beta^\dagger\} = 0, \quad \{a_\alpha, a_\beta\} = 0, \quad \{a_\alpha, a_\beta^\dagger\} = \delta_{\alpha\beta} \mathbf{1} \quad (18.1)$$

And for bosons we just change the anticommutators into commutators. We also discussed the change of basis for the creation and annihilation operators

$$c_j^\dagger = \sum_\alpha \langle \alpha | j \rangle a_\alpha^\dagger, \quad c_j = \sum_\alpha \langle j | \alpha \rangle a_\alpha \quad (18.2)$$

So we can define the field operators as the creation and annihilation operators at position x :

$$\Psi_\sigma^\dagger(x) = \sum_\alpha \varphi_{\alpha,\sigma}^*(x) a_\alpha^\dagger, \quad \Psi_\sigma(x) = \sum_\alpha \varphi_{\alpha,\sigma}(x) a_\alpha \quad (18.3)$$

We considered different operators on the Hilbert space of many particle states. We defined the additive operators like \mathbf{P} as sum of one-particle operators. In general we have

$$R = \sum_i R_i = \sum_{\alpha\beta} \langle \alpha | R_1 | \beta \rangle a_\alpha^\dagger a_\beta \quad (18.4)$$

where R_1 is the single-particle operator which only acts on single particle states. The more challenging is to define many-body operators. For example we have the two-body interaction operator

$$U = \sum_{i,j} u(x_i, x_j) = \frac{1}{2} \sum_{\alpha\beta,\gamma\delta} \langle \alpha\beta | u | \gamma\delta \rangle a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma \quad (18.5)$$

where the matrix element $\langle \alpha\beta | u | \gamma\delta \rangle$ are symmetrized or antisymmetrized matrix elements depending on whether we are dealing with bosons or fermions:

$$\langle \alpha\beta | u | \gamma\delta \rangle = \begin{cases} u_{\alpha\beta,(\gamma\delta)}, & \text{for bosons} \\ u_{\alpha\beta,[\gamma\delta]}, & \text{for fermions} \end{cases} \quad (18.6)$$

and the unsymmetrized matrix element is just

$$u_{\alpha\beta,\gamma\delta} = \int dx_1 dx_2 \varphi_\alpha^*(x_1) \varphi_\beta^*(x_2) \varphi_\gamma(x_1) \varphi_\delta(x_2) u(x_1, x_2) \quad (18.7)$$

We can choose our state $|\alpha\rangle$ such that the matrix element is just proportional to $\delta_{\alpha\beta,\gamma\delta}$. For fermions we can simplify it as

$$\langle \alpha\beta | u | \gamma\delta \rangle = \frac{1}{2} u^{(\alpha\beta)} (\delta_{\alpha\gamma} \delta_{\beta\delta} - \delta_{\alpha\delta} \delta_{\beta\gamma}) \quad (18.8)$$

So the interaction will be simplified to

$$U = \frac{1}{2} \sum_{\alpha\beta} u^{(\alpha\beta)} a_{\alpha}^{\dagger} a_{\beta}^{\dagger} a_{\beta} a_{\alpha} = \frac{1}{2} \sum_{\alpha \neq \beta} u^{(\alpha\beta)} n_{\alpha} n_{\beta} \quad (18.9)$$

For field operators and interactions of the form $u(x_1 - x_2)$ we can transform the above integral into momentum space, provided the one-particle wave functions are plane waves, so that

$$U = \sum_{k_{\alpha}, k_{\beta}, k_{\gamma}, k_{\delta}} \delta(k_{\alpha} + k_{\beta} - k_{\gamma} - k_{\delta}) a_{k_{\alpha}}^{\dagger} a_{k_{\beta}}^{\dagger} a_{k_{\delta}} a_{k_{\gamma}} \tilde{u}(k_{\gamma} - k_{\alpha}) \quad (18.10)$$

where $\tilde{u}(q)$ is the Fourier transform of $u(x)$, and $q = k_{\alpha} - k_{\gamma}$ is the momentum transfer of the process. The delta function is a result of translational invariance due to the form of the interaction, and from the invariance we have conservation of total momentum.

As a next step, our goal is to present observables as $\langle 0 | a a^{\dagger} \dots | 0 \rangle$. For our purpose it is useful to introduce the normal order where all the a are to the right of all the a^{\dagger} . So the vacuum expectation value of a normal ordered product of creation and annihilation operators vanish identically. This is useful in our later discussion.

Another useful idea to introduce is the Fermi sea, which is the ground state of an ideal Fermi gas. Let N be the number of spin 1/2 fermions and for simplicity consider N to be even. The ground state is constructed by populating the one-particle states with our fermions up to the Fermi level, which is the last occupied state. The energy of that state is called the Fermi energy ϵ_F . In the ground state all the states with $\epsilon < \epsilon_F$ are occupied whereas all states with $\epsilon > \epsilon_F$ are empty. So the occupation number is just

$$n_{\alpha} = \begin{cases} 1, & \text{if } \epsilon_{\alpha} < \epsilon_F \\ 0, & \text{if } \epsilon_{\alpha} > \epsilon_F \end{cases} \quad (18.11)$$

So the total number of particles and the total energy are just

$$N = \sum_{\alpha, \sigma} n_{\alpha, \sigma}, \quad E = \sum_{\alpha, \sigma} \epsilon_{\alpha} n_{\alpha, \sigma} \quad (18.12)$$

This is all very simple. But we will want to introduce interactions between the particles under this formalism.

Let's write the Fermi sea as a state $|F\rangle$. This state can be constructed using the particle creation operators as

$$|F\rangle = \prod_{\epsilon_{\alpha} < \epsilon_F} a_{\alpha, \sigma}^{\dagger} |0\rangle \quad (18.13)$$

Then the expectation value of the total number of particles under this state

$$N = \left\langle F \left| \sum_{\alpha, \sigma} a_{\alpha, \sigma}^{\dagger} a_{\alpha, \sigma} \right| F \right\rangle \quad (18.14)$$

This can be seen from the definition of $|F\rangle$ and we can anticommute the annihilation and creation operators all the way to the vacuum state

$$a_{\alpha, \sigma} |F\rangle = a_{\alpha, \sigma} \prod_{\beta, \sigma'} a_{\beta, \sigma'}^{\dagger} |0\rangle = \begin{cases} 0, & \text{if } \epsilon_{\alpha} > \epsilon_F \\ (-1)^{N-1} |F'\rangle, & \text{if } \epsilon_{\alpha} < \epsilon_F \end{cases} \quad (18.15)$$

where $|F'\rangle$ is the $N - 1$ particle state. We can do this for both a^\dagger and a so after taking sum we have N .

Now we want to consider $|F\rangle$ as a vacuum, and particles as excitations over this vacuum. The creation and annihilation operators for the excitations are defined to be b_α^\dagger and b_α which are

$$b_\alpha^\dagger = \begin{cases} a_\alpha^\dagger, & \text{if } |\alpha\rangle \text{ is empty in } |F\rangle \\ a_\alpha, & \text{if } |\alpha\rangle \text{ is occupied in } |F\rangle \end{cases}, \quad b_\alpha = \begin{cases} a_\alpha, & \text{if } |\alpha\rangle \text{ is empty in } |F\rangle \\ a_\alpha^\dagger, & \text{if } |\alpha\rangle \text{ is occupied in } |F\rangle \end{cases} \quad (18.16)$$

Obviously if $|\alpha\rangle$ is empty in $|F\rangle$ then we have $b_\alpha |F\rangle = 0$. Excitations are “particles” if $E_\alpha > \epsilon_F$ and they are “holes” if $E_\alpha < \epsilon_F$. The energies of the excitations are $E_\alpha - \epsilon_F$ or $\epsilon_F - E_\alpha$ correspondingly.

Let’s look at the density distribution of particles. The density can be easily represented by the field operators

$$n_\sigma(x) = \Psi_\sigma^\dagger(x) \Psi_\sigma(x) \quad (18.17)$$

So the density of the particles, or one-particle distribution function in the Fermi sea is just

$$\langle F | n_\sigma(x) | F \rangle = \sum_{\alpha\beta} \varphi_\alpha^*(x) \varphi_\beta(x) \langle F | a_\alpha^\dagger a_\beta | F \rangle = \sum_{\epsilon_\alpha < \epsilon_F} |\varphi_{\alpha,\sigma}(x)|^2 \quad (18.18)$$

We would have guessed this result from the beginning, but it is reassuring to derive this using our new formalism.

Now it is more interesting to look at the two particle distribution functions

$$\begin{aligned} \langle F | n_{\sigma_1}(x_1) n_{\sigma_2}(x_2) | F \rangle &= \langle F | \Psi_1^\dagger \Psi_1 \Psi_2^\dagger \Psi_2 | F \rangle \\ &= \sum_{\alpha\beta} |\varphi_{\alpha\sigma_1}|^2 |\varphi_{\beta\sigma_2}|^2 + K_{\sigma_1\sigma_2} \\ &= n_1 \cdot n_2 + K_{\sigma_1\sigma_2} \end{aligned} \quad (18.19)$$

Now when $\sigma_1 = \sigma_2$ we know $K_{\sigma_1\sigma_2} = 0$ because the two particles are forced to have different states. And in general $K_{\sigma_1\sigma_2}$ is just the exchange integral we have encountered before.

Let’s come back to the Hartree-Fock method that we discussed a while ago. We applied variational method on the expectation value $\langle H \rangle$ and wanted to find $|\psi\rangle$ which minimizes the expectation value. Now we have multi-particle states $|\psi\rangle = a_{\alpha_1}^\dagger \dots a_{\alpha_N}^\dagger |0\rangle$ and consider variation on this state $|\psi\rangle \rightarrow |\psi\rangle + |\delta\psi\rangle$ such that $\delta \langle H \rangle = 0$. For our purpose it is enough to consider infinitesimal unitary transformation where

$$|\alpha\rangle \rightarrow |\alpha\rangle + i \sum_{\beta} \eta_{\alpha\beta} |\beta\rangle \quad (18.20)$$

where η is a hermitian matrix. We can write our Hamiltonian as

$$\mathcal{H} = \sum_{\alpha\beta} T_{\alpha\beta} a_\alpha^\dagger a_\beta + \frac{1}{2} \sum_{\alpha\beta\gamma\delta} u_{\alpha\beta,\gamma\delta} a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma \quad (18.21)$$

where $T_{\alpha\beta} = \langle \alpha | -\hbar^2 \nabla^2 / 2m + V | \beta \rangle$ is the matrix element of the one particle energy. We want to find the state where the variation of the Hamiltonian is zero, under the constraint that the one-particle states are normalized, but this is ensured by our assumption that the variation is unitary, so

$$\delta \langle \mathcal{H} \rangle = i \sum_{\mu\nu} \eta_{\mu\nu} \left[T_{\mu\nu} + \sum_{\lambda} [u_{\lambda\mu,\lambda\nu} - u_{\lambda\mu,\nu\lambda}] \right] = 0 \quad (18.22)$$

for any $\eta_{\mu\nu}$. We call the sum over λ as $V_{\mu\nu}^{HF}$ and define the Hartree-Fock Hamiltonian

$$\mathcal{H}^{HF} = T + V^{HF} \quad (18.23)$$

and we demand the Hartree-Fock equation to be true as a result of minimizing the expectation value of the Hartree-Fock Hamiltonian

$$\mathcal{H}^{HF} |\alpha\rangle = \epsilon_\alpha |\alpha\rangle \quad (18.24)$$

which says that the stationary points of the functional are just the eigenstates of the Hartree-Fock Hamiltonian, just like what we had in usual variational calculations. The eigenvalues ϵ_α are like energy eigenvalues, but they are not the single-particle energies in this case. The sum of these eigenvalues will be the total energy plus the exchange energy because here we have counted the exchange energy twice. Therefore only in an approximate sense can we think of these eigenvalues as single-particle energies.

The Hartree-Fock equation in position space looks like

$$\begin{aligned} & \left[-\frac{\hbar^2}{2m} \nabla^2 + V(x) \right] \varphi_{\alpha,\sigma}(x) \\ & + \sum_{\alpha',\sigma'} \int dx u(x-x') \left[|\varphi_{\alpha'\sigma'}(x')|^2 \varphi_{\alpha\sigma}(x) - \delta_{\sigma\sigma'} \varphi_{\alpha'\sigma'}^*(x) \varphi_{\alpha'\sigma'}(x') \varphi_{\alpha\sigma}(x') \right] = \epsilon_\alpha \varphi_{\alpha\sigma}(x) \end{aligned} \quad (18.25)$$

And these are usually solved by a computer program.

19 Lecture 19

19.1 Scattering Theory

We are starting a new topic which is scattering theory. In scattering we are mostly interested in incoming and outgoing states. The incoming states are the particle flux we create in the scattering event, and the outgoing states are what we observe in the laboratory. We know the probability flux of a quantum state is

$$j = \frac{\hbar}{m} (\psi^* \nabla \psi - \psi \nabla \psi^*) \quad (19.1)$$

Usually incoming states are approximated as free-particle states which are just momentum eigenstates $|p\rangle$. However these states are infinitely degenerate and we can form a linear combination $|\varphi\rangle$. So we take the Hamiltonian as the free particle Hamiltonian and that

$$\mathcal{H}_0 |\varphi\rangle = E |\varphi\rangle, \quad E = \frac{p^2}{2m} \quad (19.2)$$

Now we want to add a scattering potential V . We want to impose that the potential is localized which means that $V \rightarrow 0$ at $x \rightarrow \infty$. Now the Hamiltonian is $\mathcal{H} = \mathcal{H}_0 + V$ and we want to again solve for the static Schrödinger equation

$$\mathcal{H} |\psi\rangle = E |\psi\rangle \quad (19.3)$$

We can rewrite this equation similar to the way we did in perturbation theory, and introduce a formal solution to this equation

$$|\psi\rangle = |\varphi\rangle + \frac{1}{E - \mathcal{H}_0} V |\psi\rangle \quad (19.4)$$

This is just a way to rewrite the Schrödinger equation and we are not dealing with perturbation. It is a little unpleasant to have the inverse operator, because when the eigenvalue of \mathcal{H}_0 coincide with E the operator will be singular. In order to avoid this we need to regularize it. One way is to add a positive or negative infinitesimal imaginary number in the denominator

$$|\psi^{(\pm)}\rangle = |\varphi\rangle + \frac{1}{E - \mathcal{H}_0 \pm i\epsilon} V |\psi^{(\pm)}\rangle \quad (19.5)$$

This is the celebrated Lippmann-Schwinger equation.

Let's figure out the coordinate form of this equation, by multiplying the coordinate bra $\langle x|$ to the left of the equation

$$\begin{aligned} \psi^{(\pm)}(x) &= \varphi(x) + \left\langle x \left| \frac{1}{E - \mathcal{H}_0 \pm i\epsilon} V \right| \psi^{(\pm)} \right\rangle \\ &= \varphi(x) + \int dx_1 dx_2 \left\langle x \left| \frac{1}{E - \mathcal{H}_0 \pm i\epsilon} \right| x_1 \right\rangle \langle x_1 | V | x_2 \rangle \psi^{(\pm)}(x_2) \end{aligned} \quad (19.6)$$

Now we have converted our equation into an integral equation. The integrand consists of just two matrix elements and the wave function itself.

Before we go any further, let's consider the plane wave state $|p\rangle$. Because it is not a bound state, we need to introduce a kind of normalization. The conventional one is just

$$\langle \mathbf{x} | \mathbf{p} \rangle = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}, \quad \langle \mathbf{p} | \mathbf{p}' \rangle = \delta(\mathbf{p} - \mathbf{p}') \quad (19.7)$$

Fixing this normalization is equivalent to fixing the flux at infinity. Now let's evaluate the matrix element involved in the above integral equation using the momentum eigenstates

$$\begin{aligned}
\left\langle \mathbf{x} \left| \frac{1}{E - \mathcal{H}_0 \pm i\epsilon} \right| \mathbf{x}_1 \right\rangle &= \int d\mathbf{p}_1 d\mathbf{p}_2 \langle \mathbf{x} | \mathbf{p}_1 \rangle \left\langle \mathbf{p}_1 \left| \frac{1}{E - \mathcal{H}_0 \pm i\epsilon} \right| \mathbf{p}_2 \right\rangle \langle \mathbf{p}_2 | \mathbf{x}_1 \rangle \\
&= \int d\mathbf{p} \varphi_{\mathbf{p}}(\mathbf{x}) \frac{1}{E - p^2/2m \pm i\epsilon} \varphi_{\mathbf{p}}^*(\mathbf{x}_1) \\
&= \frac{1}{(2\pi\hbar)^3} \int d\mathbf{p} \frac{e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}_1)/\hbar}}{E - p^2/2m \pm i\epsilon}
\end{aligned} \tag{19.8}$$

If we make the substitution $\mathbf{p} = \mathbf{q}\hbar$, $E = k^2/2m$, then the integral can be written as

$$\begin{aligned}
\left\langle \mathbf{x} \left| \frac{1}{E - \mathcal{H}_0 \pm i\epsilon} \right| \mathbf{x}_1 \right\rangle &= \frac{1}{(2\pi)^3} 2\pi \frac{2m}{\hbar^2} \int_0^\infty q^2 dq \int_{-1}^1 d\cos\theta \frac{e^{iq|\mathbf{x}-\mathbf{x}_1|\cos\theta}}{k^2 - q^2 \pm i\epsilon} \\
&= -\frac{2m}{4\pi^2\hbar^2} \frac{1}{i|\mathbf{x}-\mathbf{x}_1|} \int_0^\infty \frac{dq q (e^{iq|\mathbf{x}-\mathbf{x}_1|} - e^{-iq|\mathbf{x}-\mathbf{x}_1|})}{q^2 - k^2 \pm i\epsilon} \\
&= -\frac{2m}{4\pi\hbar^2} \frac{e^{\pm ik|\mathbf{x}-\mathbf{x}_1|}}{|\mathbf{x}-\mathbf{x}_1|} \\
&= \frac{2m}{\hbar^2} G_{\pm}(\mathbf{x}, \mathbf{x}_1)
\end{aligned} \tag{19.9}$$

where $G_{\pm}(\mathbf{x}, \mathbf{x}_1)$ is accordingly defined. This is just the Green function for the Helmholtz equation

$$(\nabla^2 + k^2)G_{\pm}(\mathbf{x}, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_1) \tag{19.10}$$

With this explicit evaluation, we can write the coordinate Lippmann-Schwinger equation in a new form

$$\psi^{\pm}(\mathbf{x}) = \varphi_{\mathbf{p}}(\mathbf{x}) - \frac{2m}{\hbar^2} \int d\mathbf{x}_1 d\mathbf{x}_2 \frac{e^{\pm ik|\mathbf{x}-\mathbf{x}_1|}}{4\pi|\mathbf{x}-\mathbf{x}_1|} \langle \mathbf{x}_1 | V | \mathbf{x}_2 \rangle \psi^{\pm}(\mathbf{x}_2) \tag{19.11}$$

Finally if V is just some local potential, then we have $\langle \mathbf{x}_1 | V | \mathbf{x}_2 \rangle = V(\mathbf{x}_1)\delta(\mathbf{x}_1 - \mathbf{x}_2)$, so the equation can be further simplified

$$\psi^{\pm}(\mathbf{x}) = \varphi_{\mathbf{p}}(\mathbf{x}) - \frac{2m}{\hbar^2} \int d\mathbf{x}' \frac{e^{\pm ik|\mathbf{x}-\mathbf{x}'|}}{4\pi|\mathbf{x}-\mathbf{x}'|} V(\mathbf{x}') \psi^{\pm}(\mathbf{x}') \tag{19.12}$$

The \pm sign in the exponential labels whether the state is incoming or outgoing. For $-$ sign it means incoming wave, and for $+$ sign it means outgoing wave. Suppose we choose the plus sign, then this equation is just about the dependence of scattered wave on the incoming wave $\varphi_{\mathbf{p}}(\mathbf{x})$. The above integral can be think of running over a localized scatterer, and $\psi^+(\mathbf{x})$ can be thought of the wave function at the observer. We usually take the far-field approximation $|\mathbf{x}| = r \gg r'$. Under this assumption, we can make the following expansion

$$|\mathbf{x} - \mathbf{x}'| \approx r - \frac{\mathbf{x} \cdot \mathbf{x}'}{r} + \dots \tag{19.13}$$

We also want to introduce the momentum in the direction of the observer $\mathbf{k}' = k\mathbf{x}/r$. Therefore the exponential in our integrand of the Lippmann-Schwinger equation will become

$$e^{\pm ik|\mathbf{x}-\mathbf{x}'|} \approx e^{\pm ikr} e^{\mp i\mathbf{k}' \cdot \mathbf{x}_1} \tag{19.14}$$

Now the equation becomes approximately

$$\begin{aligned}\psi^+(\mathbf{x}) &= \varphi_{\mathbf{k}}(\mathbf{x}) - \frac{1}{4\pi} \frac{2m}{\hbar^2} \frac{e^{ikr}}{r} \int d\mathbf{x}' e^{-i\mathbf{k}'\cdot\mathbf{x}'} V(\mathbf{x}') \psi^+(\mathbf{x}') \\ &= \frac{1}{(2\pi)^{3/2}} \left[e^{i\mathbf{k}\cdot\mathbf{x}} + \frac{e^{ikr}}{r} f(\mathbf{k}, \mathbf{k}') \right]\end{aligned}\quad (19.15)$$

where $f(\mathbf{k}, \mathbf{k}')$ is defined as

$$f(\mathbf{k}, \mathbf{k}') = -\sqrt{\frac{\pi}{2}} \frac{2m}{\hbar^2} \int d\mathbf{x}' e^{-i\mathbf{k}'\cdot\mathbf{x}'} V(\mathbf{x}') \psi^+(\mathbf{x}') \quad (19.16)$$

This quantity is called the scattering amplitude, and is the most central object that we study in scattering theory. Note that this quantity is proportional to the matrix element $f \sim \langle \mathbf{k}' | V | \psi^+ \rangle$. Note that this quantity is not limited to quantum mechanics, and everything we consider here can be applied to the scattering of classical waves.

Let's come back to the current, which is the flux or number of particles per unit area per unit time. If we have some unit surface of normal $\hat{\mathbf{n}}$ then the number of particles across this surface per unit time is just $\mathbf{j} \cdot \hat{\mathbf{n}}$. We can consider the quantity $|\mathbf{j}_{\text{scattered}}| / |\mathbf{j}_{\text{incoming}}|$ and this quantity is called the differential cross section

$$\frac{|\mathbf{j}_{\text{sc}}|}{|\mathbf{j}_{\text{in}}|} r^2 = \frac{d\sigma}{d\Omega} \quad (19.17)$$

where $d\Omega$ is the element of solid angle, and $d\sigma$ has the unit of area, which we call the cross section. The relation between the differential cross-section and the scattering amplitude is very simple

$$\frac{d\sigma}{d\Omega} = |f(\mathbf{k}, \mathbf{k}')|^2 \quad (19.18)$$

20 Lecture 20

20.1 Scattering Continued

Let's continue our discussion on scattering in quantum mechanics. Remember our incoming states are assumed to be free particle states $|\mathbf{k}\rangle$. These wave functions are normalized as

$$\langle \mathbf{x} | \mathbf{k} \rangle = \varphi_{\mathbf{k}}(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x}}, \quad \langle \mathbf{k} | \mathbf{k}' \rangle = \delta(\mathbf{k} - \mathbf{k}') \quad (20.1)$$

In addition to this incoming state, we have something called scatterer. It can be some fixed potential, or be another particle. We will put our observation point far away at \mathbf{x} where $|\mathbf{x}| = r \rightarrow \infty$. We then introduced the scattered state $|\psi^\pm\rangle$ and introduced regularization on the inverse operator. We also found that the physical state we are interested in is just $|\psi^+\rangle$. Then the Schrödinger equation was rewritten in a more convenient form

$$\psi^+(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \left[e^{i\mathbf{k}\cdot\mathbf{x}} + \frac{1}{r} e^{ikr} f(\mathbf{k}, \mathbf{k}') \right] \quad (20.2)$$

where $\mathbf{k}' = |\mathbf{k}| \mathbf{x}/r$ and f is the scattering amplitude. Note under this assumption we have $|\mathbf{k}| = |\mathbf{k}'|$ and energy is conserved during scattering. This is completely elastic scattering. We also defined the differential cross section

$$\frac{d\sigma}{d\Omega} = |f(\mathbf{k}, \mathbf{k}')|^2 \quad (20.3)$$

Now let's study the scattering amplitude $f(\mathbf{k}, \mathbf{k}')$. If the scattering potential has rotational invariance, then there is only one vector quantity in kinematic theory. As f is a scalar, $f(\mathbf{k}, \mathbf{k}')$ should only depend on \mathbf{k}' through its dot product: $f(\mathbf{k}, \mathbf{k} \cdot \mathbf{k}')$. We defined the momentum transfer $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ which can be written in terms of the angle between \mathbf{k} and \mathbf{k}'

$$q^2 = 2k^2(1 - \cos\theta) = 4k^2 \sin^2 \frac{\theta}{2} \quad (20.4)$$

20.2 Born Approximation

Now let's introduce approximations that allow us to solve the problem of scattering. First is the Born approximation. Remember we have

$$f = -2\pi^2 \frac{2m}{\hbar^2} \langle \mathbf{k}' | V | \psi^+ \rangle \quad (20.5)$$

We want to consider the case where V is weak, and perturbation is applicable. To first order in perturbation we have

$$f^{(1)}(\mathbf{k}, \mathbf{k}') = -\frac{1}{4\pi} \frac{2m}{\hbar^2} \int d\mathbf{x}' e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}'} V(\mathbf{x}') = -\frac{1}{4\pi} \frac{2m}{\hbar^2} \tilde{V}(\mathbf{q}) \quad (20.6)$$

So the first order approximation to the scattering amplitude is just its Fourier transform. Now as we have argued, if V is a central potential then the scattering amplitude only depends on the magnitude of \mathbf{q} . In this case we can evaluate the angular integral first, which is just an exponential integral. The result will be

$$f^{(1)}(q) = -\frac{1}{2} \frac{2m}{\hbar^2} \frac{1}{iq} \int \frac{r'^2 dr'}{r'} (e^{iqr} - e^{-iqr}) V(r') = -\frac{2m}{\hbar^2} \frac{1}{q} \int dr' r' V(r') \sin(qr') \quad (20.7)$$

Let's consider some examples of evaluating the amplitude. The first one is just a very short ranged potential $V(\mathbf{x}) = V_0\delta(\mathbf{x})$, then we can easily evaluate the integral and get

$$f^{(1)}(q) = \frac{V_0}{4\pi} \frac{2m}{\hbar^2} \quad (20.8)$$

So if our potential has a size r_0 and $r_0q \ll 1$, then the scattering amplitude is determined by the integral of the potential over the small volume, and does not depend on q . This is the simplest example. Let's consider another $V(r) = V_0e^{-\kappa r}$. We define $r_0 = 1/\kappa$. The above integral can be evaluated as

$$f^{(1)}(q) = \frac{2m}{\hbar^2} V_0 \frac{\kappa}{(q^2 + \kappa^2)^2} \quad (20.9)$$

Finally we have the Yukawa potential which is $V(r) = V_0e^{-\kappa r}/r$. This is one of the ways of regularizing the Coulomb potential as usually the slow-decaying potential will bring some problems. Repeating the above integral we will get

$$f^{(1)}(q) = \frac{2m}{\hbar^2} \frac{V_0}{\kappa} \frac{1}{q^2 + \kappa^2} \xrightarrow{q \gg \kappa} \frac{1}{q^2} \quad (20.10)$$

This naturally leads to the Coulomb scattering where $V(r) = q_1q_2/r$ which is just the limit $\kappa \rightarrow 0$ in the Yukawa case. Then we have

$$\frac{d\sigma}{d\Omega} = f^2 \propto \frac{1}{q^4} = \frac{1}{16k^4 \sin^4(\theta/2)} \quad (20.11)$$

In fact if we put back the \hbar from the pre-factor, then the above result has no \hbar dependence, so it is natural that the same result can be calculated classically.

Let's summarize. In Born approximation and spherical symmetric case, we have the following

1. $f(\mathbf{k}, \mathbf{k}') = f(q)$ only depends on the magnitude of the momentum transfer where $q = 2k \sin \theta/2$.
2. $f(q)$ is always real.
3. If $V \rightarrow -V$ then $f(q) \rightarrow -f(q)$ because f always has only one power of V in its integral. So $d\sigma/d\Omega$ remains invariant.
4. If V is monotonic, then $f(q)$ decreases when q increases.
5. Usually when $q \rightarrow 0$ we have $f(q)$ will approach a constant. This is because when the incoming state is very soft then the potential will always look like a δ function.

The reason we discuss Born approximation is that it is the simplest approximation. So it is natural to question its validity. Remember the approximation is equivalent to replacing $|\psi^+\rangle$ by $|\mathbf{k}\rangle$ in the right hand side of the Lippmann-Schwinger equation. In order make the approximation valid, the matrix element should be smaller than 1

$$|\langle \mathbf{k}' | V | \mathbf{k} \rangle| \ll 1 \quad (20.12)$$

In the case where V is the Yukawa potential then there is only one dimensionless quantity we can construct which is interesting, so the criterion becomes

$$\frac{2m}{\hbar^2} \frac{|V_0|}{\kappa^2} \ll 1 \quad (20.13)$$

When $V_0 < 0$ then bound state will appear when the above quantity is of order 1, or more specifically larger than 2.7. So the above range is safe from bound states. For the same reason we have problem with Coulomb potential, because Coulomb potential we always have bound states. And in 2 dimensions we always have bound states for any potential, so it is very hard to apply Born approximation in 2 dimensions.

20.3 Beyond Born Approximation

The above discussion provides enough motivation for us to move beyond the Born approximation. Remember our Lippmann-Schwinger equation connects the incoming state $|\mathbf{k}\rangle$ with the true wave function $|\psi^+\rangle$. This connection is linear as

$$|\psi^+\rangle = |\mathbf{k}\rangle + G_+ V |\psi^+\rangle, \quad G_+ = \frac{1}{E - \mathcal{H}_0 + i\epsilon} \quad (20.14)$$

Let's introduce the transition operator T such that for any incoming state $|\mathbf{k}\rangle$

$$T |\mathbf{k}\rangle = V |\psi^+\rangle \quad (20.15)$$

So if we operate V on both sides of the Lippmann-Schwinger equation we will get

$$T |\mathbf{k}\rangle = V |\mathbf{k}\rangle + V G_+ T |\mathbf{k}\rangle \quad (20.16)$$

Because the equation is valid for any \mathbf{k} then we can drop the state and write it as an operator equation

$$T = V + V G_+ T = V + V G_+ V + V G_+ V G_+ V + \dots \quad (20.17)$$

This provides a very convenient iterative expansion for the operator T .

Remember the scattering amplitude is defined as

$$f = -2\pi^2 \frac{2m}{\hbar^2} \langle \mathbf{k}' | V | \psi^+ \rangle = -A \langle \mathbf{k}' | T | \mathbf{k} \rangle \quad (20.18)$$

So from the above expansion we can write similar one for the scattering amplitude

$$f = \sum_{n=1}^{\infty} f^{(n)}, \quad f^{(n)}(\mathbf{k}, \mathbf{k}') = -A \langle \mathbf{k}' | V (G_+ V)^{n-1} | \mathbf{k} \rangle \quad (20.19)$$

Note in this definition $f^{(1)}$ is the same as what we have in Born approximation. We can also write down the integration

$$f^{(n)}(\mathbf{k}, \mathbf{k}') = -A \int d\mathbf{k}_1 \dots d\mathbf{k}_n \langle \mathbf{k}' | V | \mathbf{k}_1 \rangle \langle \mathbf{k}_1 | G_+ | \mathbf{k}_2 \rangle \dots \langle \mathbf{k}_n | V | \mathbf{k} \rangle \quad (20.20)$$

Now note we have

$$\langle \mathbf{k}_1 | G_+ | \mathbf{k}_2 \rangle = G_+(\mathbf{k}_1, \mathbf{k}_2) = \frac{1}{E - k_1^2/2m + i\epsilon} \delta(\mathbf{k}_1 - \mathbf{k}_2) \quad (20.21)$$

Therefore we can write down an iterative formula

$$f^{(n)}(\mathbf{k}', \mathbf{k}) = -\frac{\hbar^2}{4m\pi^2} \int \frac{d\mathbf{k}_1}{(2\pi)^3} \frac{f^{(n-1)}(\mathbf{k}', \mathbf{k}_1) f^{(1)}(\mathbf{k}_1, \mathbf{k})}{E - k_1^2/2m + i\epsilon} \quad (20.22)$$

In particular we have

$$f^{(2)}(\mathbf{k}', \mathbf{k}) = -\frac{\hbar^2}{4m\pi^2} \int \frac{d\mathbf{k}_1}{(2\pi)^3} \frac{f^{(1)}(\mathbf{k}', \mathbf{k}_1) f^{(1)}(\mathbf{k}_1, \mathbf{k})}{E - k_1^2/2m + i\epsilon} \quad (20.23)$$

Note that k_1 is not required to be on-shell, whereas k and k' are, because k_1 is just an intermediate state and we have the uncertainty relation inherent to quantum mechanics.

We are now ready to introduce the Optical theorem. We define the total cross section

$$\sigma_{\text{total}} = \int \frac{d\sigma}{d\Omega} d\Omega = \int |f|^2 d\Omega \quad (20.24)$$

We are interested in the forward scattering amplitude $f(\theta = 0)$. The optical theorem states that these two quantities are related by

$$\text{Im}f(\theta = 0) = \frac{k}{4\pi} \sigma_{\text{total}} \quad (20.25)$$

This is a very general statement, and does not depend too much on the detail of the interaction.

21 Lecture 21

21.1 Scattering Theory Reviewed

Remember what we have covered so far for scattering theory. We separated the Hamiltonian similar to the way in perturbation theory. We derived the Lippmann-Schwinger equation for the solution of the Schrödinger equation

$$|\psi^\pm\rangle = |\varphi\rangle + G_\pm V |\psi^\pm\rangle \quad (21.1)$$

where $|\varphi\rangle$ is the free particle eigenstate $|\varphi\rangle = |\mathbf{k}\rangle$. We defined the G operator as

$$G_\pm = \frac{1}{E - \mathcal{H}_0 \pm i\epsilon} \quad (21.2)$$

The plus sign is for outgoing states and the minus sign is for incoming states. In terms of wave functions, the Lippmann-Schwinger equation is an integral equation

$$\psi^\pm(\mathbf{x}) = \varphi(\mathbf{x}) - \frac{2m}{\hbar^2} \int d\mathbf{x}' G_\pm(\mathbf{x}, \mathbf{x}') V(\mathbf{x}') \psi^\pm(\mathbf{x}'), \quad G_\pm = -\frac{1}{4\pi} \frac{e^{\pm ik|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|} \quad (21.3)$$

Last time we found that in the limit that the observer is very far away, $|\mathbf{x}| = r \rightarrow \infty$, then we have

$$\psi^\pm(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \left[e^{i\mathbf{k}\cdot\mathbf{x}} + \frac{e^{ikr}}{r} f(\mathbf{k}, \mathbf{k}') \right] \quad (21.4)$$

Note this expression is the definition of the scattering amplitude f , which only makes sense in the limit $r \rightarrow \infty$. In our formalism we can write

$$f(\mathbf{k}, \mathbf{k}') = -2\pi^2 \frac{2m}{\hbar^2} \langle \mathbf{k} | V | \psi^+ \rangle \quad (21.5)$$

The differential cross section is just the absolute value square of the scattering amplitude.

$$\frac{d\sigma}{d\Omega} = |f(\mathbf{k}, \mathbf{k}')|^2 \quad (21.6)$$

Although in this quantity only the magnitude of f matters, as f is a complex number in general its phase also matters in the discussion that follows.

We also introduced the Born approximation, which largely simplified our calculation for the scattering amplitude

$$f^{(1)}(\mathbf{k}, \mathbf{k}') = -\frac{2m}{\hbar^2} \frac{1}{4\pi} \int d\mathbf{x} e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} V(\mathbf{x}) \quad (21.7)$$

so the first order scattering amplitude is just the Fourier transform of the scattering potential. If the potential is spherical symmetric, then the scattering amplitude only depends on θ which is the angle between \mathbf{k} and \mathbf{k}'

$$f(\theta) = -\frac{2m}{\hbar^2} \int_0^\infty r V(r) \frac{\sin qr}{q} dr \quad (21.8)$$

We introduced the transition operator T in order to write Lippmann-Schwinger equation in operator form. We defined $T|\varphi\rangle = V|\psi^+\rangle$, and the equation translates into

$$T = V + VG_+T \quad (21.9)$$

and it has a nice iterative solution. The scattering amplitude can be interpreted nicely as the matrix element of the transition operator

$$f(\mathbf{k}, \mathbf{k}') = -2\pi^2 \frac{2m}{\hbar^2} \langle \mathbf{k}' | T | \mathbf{k} \rangle, \quad f^{(n)} \propto \langle \mathbf{k}' | V(G_+V)^n | \mathbf{k} \rangle \quad (21.10)$$

21.2 Optical Theorem

Last time we introduced the optical theorem but did not prove it. The theorem states that

$$\text{Im } f(\mathbf{k}, \mathbf{k}) = \frac{k}{4\pi} \sigma_{\text{tot}} \quad (21.11)$$

where σ_{tot} is the total cross section, the integral of differential cross section over all solid angle. Now let's try to prove it. For notational simplicity we label $A = -4m\pi^2/\hbar^2$. We use the L-S equation to rewrite

$$|\mathbf{k}\rangle = |\psi^+\rangle - G_+ V |\psi^+\rangle = |\psi^+\rangle - G_+ T |\mathbf{k}\rangle \quad (21.12)$$

Note that $G_+^\dagger = G_-$, so we can take conjugate on the above equation

$$\langle \mathbf{k} | = \langle \psi^+ | - \langle \mathbf{k} | T^\dagger G_- \quad (21.13)$$

Now we can write the imaginary part of f as

$$\begin{aligned} \text{Im } f(\mathbf{k}, \mathbf{k}) &= \text{Im } A \langle \mathbf{k} | T | \mathbf{k} \rangle \\ &= A \text{Im} [\langle \mathbf{k} | V | \mathbf{k} \rangle + \langle \mathbf{k} | V G_+ T | \mathbf{k} \rangle] \\ &= A \text{Im} [\langle \psi^+ | V G_+ T | \mathbf{k} \rangle - \langle \mathbf{k} | T^\dagger G_- V G_+ T | \mathbf{k} \rangle] \\ &= A \text{Im} \sum_{\mathbf{k}_1, \mathbf{k}_2} \langle \mathbf{k} | T^\dagger | \mathbf{k}_1 \rangle \langle \mathbf{k}_1 | G_+ | \mathbf{k}_2 \rangle \langle \mathbf{k}_2 | T | \mathbf{k} \rangle \end{aligned} \quad (21.14)$$

Note that the matrix element of the operator G_+ is

$$\langle \mathbf{k}_1 | G_+ | \mathbf{k}_2 \rangle = \frac{\delta(\mathbf{k}_1 - \mathbf{k}_2)}{E - \hbar^2 k_1^2 / 2m + i\epsilon} \quad (21.15)$$

substituting this into the above equation we can get

$$\text{Im } f(\mathbf{k}, \mathbf{k}) = \frac{1}{A} \sum_{\mathbf{k}_1} |f(\mathbf{k}, \mathbf{k}_1)|^2 \text{Im } G_+(\mathbf{k}_1) \quad (21.16)$$

Now we can use the principal value

$$\text{Im} \frac{1}{x + i\epsilon} = -\pi \delta(x), \quad \delta(F(x)) = \sum_{x_0} \frac{1}{|F'(x_0)|} \delta(x - x_0) \quad (21.17)$$

to unfold the expression and get

$$\text{Im } f(\mathbf{k}, \mathbf{k}) = \frac{1}{4\pi k} \int k^2 |f(\mathbf{k}, \mathbf{k}_1)|^2 d\Omega = \frac{k}{4\pi} \sigma_{\text{tot}} \quad (21.18)$$

This result is useful in checking whether our calculation result of scattering amplitude make sense or not.

21.3 Spherically Symmetric Scattering

Let's consider again a central potential, or spherically symmetric potential $V(\mathbf{x}) = V(r)$. The free Hamiltonian is $\mathcal{H}_0 = \mathbf{p}^2/2m$ and $|\mathbf{k}\rangle$ are eigenstates of the free Hamiltonian. However we know that the angular momentum operators \mathbf{L}^2 and L_z also commute with \mathcal{H}_0 and we can use states with definite angular momentum as the free states $|E, \ell, m\rangle$. These are the spherical wave states.

Now if there is a scatterer which is spherically symmetric, then the operator \mathbf{p} no longer commutes with the Hamiltonian, but \mathbf{L}^2 and L_z still commute with \mathcal{H} . Therefore it is more convenient to use spherical waves as the free states in this situation. So now we need to learn how to operate the mechanism of scattering theory with these states.

To start with, we want to impose a normalization for these states

$$\langle E', \ell', m | E, \ell, m \rangle = \delta_{\ell\ell'} \delta_{mm'} \delta(E - E') \quad (21.19)$$

We would like to expand every state in our scattering theory in terms of these states. At least schematically this is easy and we have

$$|\mathbf{k}\rangle = \sum_{E, \ell, m} |E, \ell, m\rangle \langle \mathbf{k} | E, \ell, m \rangle \quad (21.20)$$

However we need to know the coefficients, which are nothing but some inner products.

Before going any further let's clarify our objective for doing this. Remember last semester we discussed WKB approximation and barrier penetration. Let's consider similar things here. If we have a solid wall and some incoming plane wave with wave vector \mathbf{k} , then upon incidence all the wave will be reflected and by energy conservation the reflected wave will also have wave vector \mathbf{k} . Now suppose we have a potential gradually going to infinity instead of a steep wall, then the reflected wave will still have wave vector \mathbf{k} , but what is different is the phase of the reflected wave. Similar things happen in 3D. If all particles come in with definite angular momenta, scatter off the spherical symmetric potential, then the scattering process will not change the angular momentum quantum numbers. The only thing can change is the phase. So our goal is to derive a formula for the phase shift.

There are a couple of things that we need. The inner product quoted above is

$$\langle \mathbf{k} | E, \ell, m \rangle = \frac{\hbar}{\sqrt{mk}} \delta \left(E - \frac{\hbar^2 k^2}{2m} \right) Y_\ell^m(\theta_k, \phi_k) \quad (21.21)$$

similarly we have the inner product with the position operator

$$\langle \mathbf{x} | E, \ell, m \rangle = \frac{1}{\hbar} i^\ell \sqrt{\frac{2mk}{\pi}} Y_\ell^m(\theta_x, \phi_x) j_\ell(kr) \quad (21.22)$$

where $j_\ell(x)$ are spherical Bessel functions. For our purpose we only need to consider the case where ℓ is an integer, so we have the formula

$$j_\ell(x) = (-x)^\ell \left(\frac{1}{x} \frac{d}{dx} \right)^\ell \frac{\sin x}{x} \quad (21.23)$$

Therefore the plane wave can be expanded using the spherical waves as follows

$$e^{i\mathbf{k}\cdot\mathbf{x}} = 4\pi \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} i^\ell j_\ell(kr) [Y_\ell^m(\theta_k, \phi_k)]^* Y_\ell^m(\theta_x, \phi_x) \quad (21.24)$$

Now let's write our beloved scattering amplitude using this spherical wave representation

$$f(\mathbf{k}, \mathbf{k}') = A \langle \mathbf{k}' | T | \mathbf{k} \rangle = A \sum_{\ell\ell', mm'} \int dE dE' \langle \mathbf{k}' | E', \ell', m' \rangle \langle E', \ell', m' | T | E, \ell, m \rangle \langle E, \ell, m | \mathbf{k} \rangle \quad (21.25)$$

The matrix element for the T operator can't depend on angular momentum quantum numbers, so we can write

$$\langle E', \ell', m' | T | E, \ell, m \rangle = \delta(E - E') \delta_{\ell\ell'} \delta_{mm'} T_\ell(E) \quad (21.26)$$

Therefore the formula for scattering amplitude is written as

$$f(\mathbf{k}', \mathbf{k}) = A \frac{\hbar^2}{mk} \sum_{\ell, m} T_\ell \left(E = \frac{\hbar^2 k^2}{2m} \right) Y_\ell^m(\hat{\mathbf{k}}) \left[Y_\ell^m(\hat{\mathbf{k}}') \right]^* \quad (21.27)$$

We can invoke the spherical harmonics addition theorem to get rid of the m summation

$$P_\ell(\cos \theta) = \frac{4\pi}{2\ell + 1} \sum_{m=-\ell}^{\ell} Y_\ell^m(\hat{\mathbf{k}}')^* Y_\ell^m(\hat{\mathbf{k}}) \quad (21.28)$$

Then we can write the scattering amplitude as

$$f(\mathbf{k}, \mathbf{k}') = - \sum_{\ell=0}^{\infty} \frac{\pi}{k} (2\ell + 1) T_\ell(k) P_\ell(\cos \theta) = \sum_{\ell=0}^{\infty} (2\ell + 1) f_\ell(k) P_\ell(\cos \theta) = f(\theta) \quad (21.29)$$

In the limit $r \rightarrow \infty$ we can write the spherical Bessel function as

$$j_\ell(kr) = (-1)^\ell \frac{1}{kr} \frac{d^\ell}{d(kr)^\ell} \sin(kr) = \frac{1}{2ikr} \left[e^{i(kr - \pi\ell/2)} - e^{-i(kr - \pi\ell/2)} \right] \quad (21.30)$$

where the first term is identified as outgoing partial wave, and the second is identified as incoming partial wave. Therefore the wave function in this limit is written as

$$\psi_{k,\ell,m}(\mathbf{x}) \rightarrow \sum_{\ell=0}^{\infty} (2\ell + 1) P_\ell(\cos \theta) \left[S_\ell \frac{e^{ikr}}{r} - \frac{1}{r} e^{-ikr} e^{-i\pi\ell} \right] \quad (21.31)$$

where $S_\ell = 1 + 2ikf_\ell(k)$. These are the partial scattering amplitudes.

There is one thing we haven't used, which is the conservation of number of particles. This means that the number of incoming particles is the same as number of outgoing particles, and this means that

$$|S_\ell|^2 = 1 \quad (21.32)$$

This condition is called *unitarity*. This condition means that we can write $S_\ell = e^{2i\delta_\ell}$ where δ_ℓ are called scattering phases.

22 Lecture 22

22.1 Spherical Symmetric Scattering Reviewed

Remember we were discussing the special case of spherical symmetric scattering where the scattering potential only depends on the radial distance $V(\mathbf{x}) = V(r)$. Our main object of interest was the scattering amplitude $f(\mathbf{k}', \mathbf{k})$. In plane wave limit we had the result

$$\psi^+(\mathbf{x}) \xrightarrow{r \rightarrow \infty} \frac{1}{(2\pi)^{3/2}} \left[e^{i\mathbf{k} \cdot \mathbf{x}} + \frac{e^{ikr}}{r} f(\mathbf{k}', \mathbf{k}) \right] \quad (22.1)$$

where for elastic scattering we have $k = |\mathbf{k}| = |\mathbf{k}'|$. Then we can write the scattering amplitude as

$$f(\mathbf{k}', \mathbf{k}) = f(|\mathbf{k} - \mathbf{k}'|) = f(k, \theta), \quad \cos \theta = \hat{\mathbf{k}} \cdot \hat{\mathbf{k}}' \quad (22.2)$$

Last time we expanded f in terms of Legendre polynomials

$$f(\mathbf{k}', \mathbf{k}) = \sum_{\ell=0}^{\infty} (2\ell + 1) f_{\ell}(k) P_{\ell}(\cos \theta) \quad (22.3)$$

where the coefficients are nicely related to the matrix elements of the transition operator

$$f_{\ell}(k) = \langle E, \ell, m | T | E, \ell, m \rangle \left(-\frac{\pi}{k} \right) \quad (22.4)$$

Note that T is automatically diagonal in this basis. This is due to conservation of angular momentum and energy, which is the consequence of elastic scattering. Our basis in momentum and coordinate representations are respectively

$$\langle \mathbf{k} | E, \ell, m \rangle = \frac{\hbar}{\sqrt{mk}} \delta \left(E - \frac{\hbar^2 k^2}{2m} \right) Y_{\ell}^m(\theta_k, \phi_k), \quad \langle \mathbf{x} | E, \ell, m \rangle = \frac{1}{\hbar} i^{\ell} \sqrt{\frac{2mk}{\pi}} Y_{\ell}^m(\theta_x, \phi_x) j_{\ell}(kr) \quad (22.5)$$

Then we can evaluate the expansion of plane wave in terms of these spherical waves

$$\langle \mathbf{x} | \mathbf{k} \rangle = \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{k} \cdot \mathbf{x}} = \frac{1}{(2\pi)^{3/2}} \sum (2\ell + 1) P_{\ell}(\cos \theta) j_{\ell}(kr) i^{\ell} \quad (22.6)$$

Remember the asymptotic form of the spherical Bessel functions are

$$j_{\ell}(kr) \xrightarrow{r \rightarrow \infty} (-1)^{\ell} \frac{1}{kr} \frac{d^{\ell}}{d(kr)^{\ell}} \sin(kr) = \frac{1}{2ikr} \left[e^{i(kr - \pi\ell/2)} - e^{-i(kr - \pi\ell/2)} \right] \quad (22.7)$$

so in this limit the plane wave becomes

$$\langle \mathbf{x} | \mathbf{k} \rangle \xrightarrow{r \rightarrow \infty} \frac{1}{(2\pi)^{3/2}} \sum_{\ell=0}^{\infty} \frac{2\ell + 1}{2kr} i^{\ell-1} P_{\ell}(\cos \theta) \left[e^{ikr - i\ell\pi/2} - e^{-ikr + i\ell\pi/2} \right] \quad (22.8)$$

We identify the first term in the square bracket as the outgoing wave, and the second incoming wave, both of which spherical. It is important that this equation is only an asymptotic form and is only valid at $r \rightarrow \infty$. Now if there is scattering potential, then the wave function should become

$$\langle \mathbf{x} | \psi^+ \rangle = \frac{1}{(2\pi)^{3/2}} \sum_{\ell=0}^{\infty} \frac{2\ell + 1}{2ik} P_{\ell}(\cos \theta) \left[(1 + 2ikf_{\ell}(k)) \frac{e^{ikr}}{r} - (-1)^{\ell} \frac{e^{-ikr}}{r} \right] \quad (22.9)$$

We will denote the term $(1 + 2ikf_\ell(k))$ by $S_\ell(k)$. By conservation of number of particles and rotation invariance, we know that the number of incoming particles with angular momentum ℓ will be equal to the number of outgoing particles with angular momentum ℓ . So we require

$$|S_\ell|^2 = 1, \quad S_\ell = e^{2i\delta_\ell} \quad (22.10)$$

where δ_ℓ are called scattering phases. The condition above is called unitarity and is believed to hold for all quantum scattering process. If we label the incoming channels using index i and outgoing channels by j , then the matrix S_{ij} is called the S -matrix, which encodes all the information of scattering process, and our unitarity condition is equivalent to $S^\dagger S = 1$. In our case the channels are labelled by their angular momentum quantum numbers, and the matrix is diagonal.

Now with an expression for S_ℓ we can get f_ℓ under some approximation

$$f_\ell(k) = \frac{S_\ell - 1}{2ik} = e^{i\delta_\ell} \frac{\sin \delta_\ell}{k} \xrightarrow{\delta_\ell \ll 1} \frac{\delta_\ell}{k} \quad (22.11)$$

Then the scattering amplitude is reduced to

$$f(k, \theta) = \frac{1}{k} \sum_{\ell=0}^{\infty} (2\ell + 1) e^{i\delta_\ell} \sin \delta_\ell P_\ell(\cos \theta) \quad (22.12)$$

Note that the summation term is dimensionless, and $f \sim 1/k$ has the dimension of length, and f^2 has the right dimension of cross section. And accordingly we can write down the scattering cross section

$$\sigma_{\text{tot}} = \int \frac{d\sigma}{d\Omega} d\Omega = \int_0^{2\pi} d\phi \int_{-1}^1 d(\cos \theta) \frac{1}{k^2} \sum_{\ell, \ell'} (2\ell + 1)(2\ell' + 1) \sin \delta_\ell \sin \delta_{\ell'} P_\ell(\cos \theta) P_{\ell'}(\cos \theta) \quad (22.13)$$

Remember the orthogonality condition for Legendre polynomials

$$\int d(\cos \theta) P_\ell(\cos \theta) P_{\ell'}(\cos \theta) = \frac{2}{2\ell + 1} \delta_{\ell, \ell'} \quad (22.14)$$

So the total cross section is much simplified

$$\sigma_{\text{tot}} = \frac{4\pi}{k^2} \sum_{\ell} (2\ell + 1) \sin^2 \delta_\ell = \sum_{\ell} \sigma_\ell \quad (22.15)$$

where σ_ℓ can be interpreted as total scattering cross sections for partial waves of angular momentum ℓ . We can also use this expression to again verify the optical theorem.

If we are dealing with weak scattering, for example when Born approximation is applicable, then we have $\delta_\ell \ll 1$. In this case we can expand the exponential and sine function in the expression for f and find that f_ℓ is small. However for strong scattering $\sin^2 \delta_\ell \leq 1$, so the total cross section has an upper bound

$$\sigma_\ell \leq \frac{4\pi}{k^2} (2\ell + 1) \quad (22.16)$$

This limit is called the unitary limit and however strong the potential is, this is always the upper bound for the total cross section as long as unitarity is preserved. The equality happens usually when there is a resonance, where the scattering potential has a bound state.

22.2 Examples

For our convenience we will assume that our potential is of finite range, so that $V(r) = 0$ when $r > R$ for some radius R . Our spherical Bessel functions come from the solution of the radial equation

$$\frac{d^2}{dx^2} j_\ell(x) + \frac{2}{x} \frac{d}{dx} j_\ell(x) + \left[1 - \frac{\ell(\ell+1)}{x^2} \right] j_\ell(x) = 0 \quad (22.17)$$

Note that $j_\ell(x)$ is regular at $x \rightarrow 0$. However every second order ordinary differential equation has two linearly independent sets of solutions and we have another set of spherical Bessel functions

$$n_\ell(x) = (-x)^\ell \left(\frac{1}{x} \frac{d}{dx} \right)^\ell \frac{\cos x}{x} \quad (22.18)$$

These are also solutions but they are singular at $x \rightarrow 0$. We discarded this set of functions in our discussion without potential because we want our wave function to be regular everywhere. However now if we don't worry about what happens in a region $r < R$ due to the existence of a potential, then we need to use n_ℓ in addition to j_ℓ . In this case we need to write the free wave function as

$$\langle \mathbf{x} | \mathbf{k} \rangle = \frac{1}{(2\pi)^{3/2}} \sum_\ell i^\ell (2\ell+1) P_\ell(\cos \theta) A_\ell(r) \quad (22.19)$$

Before A_ℓ was just j_ℓ , but now we need

$$A_\ell = a j_\ell(kr) + b n_\ell(kr) = c_\ell^{(1)} h^{(1)}(kr) + c_\ell^{(2)} h^{(2)}(kr) \quad (22.20)$$

where for convenience we define the functions

$$h_\ell^{(1,2)}(x) = j_\ell(x) \pm i n_\ell(x) \sim e^{\pm ix} \quad (22.21)$$

If there is no scattering potential then we should have our original result, so we expect $c^{(1)} = c^{(2)} = 1/2$. In the case of ψ^+ we need to replace

$$A_\ell = e^{i\delta_\ell} [\cos \delta_\ell j_\ell(kr) - \sin \delta_\ell n_\ell(kr)] \quad (22.22)$$

And to solve for δ_ℓ we need to assume appropriate boundary conditions on $r = R$.

Before we do any serious calculations, let's consider some examples where we can get the result without too much effort. Let's consider the potential

$$V(\mathbf{x}) = \begin{cases} 0 & \text{if } r > R \\ \infty & \text{if } r < R \end{cases} \quad (22.23)$$

This is the so-called hard-sphere scattering. The boundary conditions on $r = R$ is simply that the wave function vanishes, so $A_\ell(R) = 0$ for any ℓ

$$j_\ell(kR) \cos \delta_\ell - n_\ell(kR) \sin \delta_\ell = 0 \quad \implies \quad \tan \delta_\ell = \frac{j_\ell(kR)}{n_\ell(kR)} \quad (22.24)$$

So this is the explicit form of the scattering phase. And know the scattering phase we will know f and therefore will know $d\sigma/d\Omega$.

Let's look at the meaning of this result. Let's consider the case $\ell = 0$, then we can explicitly write down the expression

$$\tan \delta_0 = \frac{\sin(kR)/kR}{\cos(kR)/kR} \implies \delta_0 = -kR \quad (22.25)$$

Therefore the s -wave will always scatter with a constant phase shift which is proportional to the radius of the sphere. Now let's consider the higher spherical waves in the limit $kR \ll 1$ which is the low energy limit. A straightforward calculation of asymptotics will give

$$j_\ell(kr) \xrightarrow{kr \ll 1} \frac{(kr)^\ell}{(2\ell + 1)!!}, \quad n_\ell(kr) \xrightarrow{kr \ll 1} -\frac{(2\ell - 1)!!}{(kr)^{\ell+1}} \quad (22.26)$$

So the scattering phase shift is

$$\tan \delta_\ell \sim \frac{(-kR)^{2\ell+1}}{(2\ell + 1) [(2\ell - 1)!!]^2} \quad (22.27)$$

So δ_ℓ quickly decays for large ℓ . The sum for total cross section will be dominated by the $\ell = 0$ term

$$\sigma_{\text{tot}} \sim \frac{4\pi}{k^2} \sin^2 \delta_0 = \frac{4\pi}{k^2} (kR)^2 = 4\pi R^2 \quad (22.28)$$

This is a well-known result for hard sphere scattering.

23 Lecture 23

23.1 Scattering off a Central Potential, Review

Remember we have been doing scattering off a central potential for a while. The scattering potential was written as $V(\mathbf{x}) = V(r)$. We wrote the scattering amplitude $f(\mathbf{k}, \mathbf{k}')$ as a function of the magnitude of the wave vectors (assuming elastic scattering) and their angle $f(k, \theta)$. Remember we expanded the function in Legendre polynomials and wrote

$$f(k, \theta) = \sum_{\ell=0}^{\infty} (2\ell + 1) f_{\ell}(k) P_{\ell}(\cos \theta) \quad (23.1)$$

We defined the S -matrix element $S_{\ell} = 1 + 2ik f_{\ell}$ and from the unitarity requirement $|S_{\ell}|^2 = 1$ we wrote it as $S_{\ell} = e^{2i\delta_{\ell}}$. Then we were able to write the partial wave scattering amplitude as

$$f_{\ell}(k) = e^{i\delta_{\ell}} \frac{\sin \delta_{\ell}}{k} \quad (23.2)$$

For simplicity we considered the case where the potential is of finite range, i.e. $V(r) = 0$ for $r > R$ where R is finite. Therefore when outside the potential region we have the free wave function

$$\psi^+(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \sum_{\ell=0}^{\infty} i^{\ell} (2\ell + 1) P_{\ell}(\cos \theta) A_{\ell}(r) \quad (23.3)$$

where $A_{\ell}(r)$ is the solution of the radial equation. If we just plug in the Laplacian operator in spherical coordinates and write the radial equation for definite ℓ then we can find

$$A_{\ell}'' + \frac{2}{x} A_{\ell}' + \left[1 - \frac{\ell(\ell + 1)}{x^2} \right] A = 0 \quad (23.4)$$

where we have changed the variable $x = kr$. The solutions are two kinds of spherical Bessel functions

$$j_{\ell}(x) = (-x)^{\ell} \left(\frac{1}{x} \frac{d}{dx} \right)^{\ell} \frac{\sin x}{x}, \quad n_{\ell}(x) = (-x)^{\ell} \left(\frac{1}{x} \frac{d}{dx} \right)^{\ell} \frac{\cos x}{x} \quad (23.5)$$

Note $n_{\ell}(x)$ is singular at the origin, but if we only consider solutions outside the potential region then we have to take both into our solution and take a linear combination

$$A_{\ell}(r) = e^{i\delta_{\ell}} [\cos \delta_{\ell} j_{\ell}(kr) - \sin \delta_{\ell} n_{\ell}(kr)] \quad (23.6)$$

We considered the hard-sphere scattering where the potential is infinity at $r < R$. The boundary condition we need to use is $A_{\ell}(R) = 0$ and this provides us the complete solution for δ_{ℓ}

$$\tan \delta_{\ell} = \frac{j_{\ell}(kR)}{n_{\ell}(kR)} \quad (23.7)$$

We considered the limit when $kR \ll 1$ which is the limit of low energy scattering. In this limit we can just limit our consideration to s -wave scattering by the asymptotics of the Bessel functions, because we found $\delta_{\ell} \sim (kR)^{2\ell+1}$. So in the hard-sphere scattering we found that

$$\delta_{\ell} \approx kR \implies \sigma_{\text{tot}} = 4\pi R^2 \quad (23.8)$$

In the opposite limit when $kR \gg 1$, if the impact parameter of the particle is ρ , then the angular momentum $L \sim \hbar \rho k \lesssim kR$, so we have to sum the angular momentum ℓ all the way up to the order kR .

23.2 Finite Potential at Finite Distance

Now let's go beyond the simple case of hard sphere scattering and assume that $V(r) \neq \infty$ when $r < R$. In this case we need to actually solve the radial equation, just like we solve the Schrödinger equation for the wave function. Traditionally it is more convenient to write $u_\ell = rA_\ell(r)$ and write the radial equation as follows

$$\frac{d^2 u_\ell}{dr^2} + \left[k^2 - \frac{2m}{\hbar^2} V(r) - \frac{\ell(\ell+1)}{r^2} \right] u_\ell = 0 \quad (23.9)$$

Now we need to worry about boundary conditions. One obvious condition is that we want $A_\ell(0)$ to be finite, which means that $u_\ell(0) = 0$. For $r > R$ the above equation reduces to the free equation, and we should use the free wave as the boundary condition at $r = R$. However if we specify $u_\ell(R)$ then we need the normalization to make sense of the number, and we can't fix the normalization by the information at $r > R$ alone. So we need a more clever way of specifying the boundary condition. We can take the derivative

$$\frac{d \ln u_\ell}{d \ln r} = 1 + \frac{d \ln A_\ell}{d \ln r} = 1 + \beta_\ell \quad (23.10)$$

At $r = R_+$ we can write the function β_ℓ using the free wave function

$$\beta_\ell^{(+)} = kR \left[\frac{j'_\ell \cos \delta_\ell - n'_\ell \sin \delta_\ell}{j_\ell \cos \delta_\ell - n_\ell \sin \delta_\ell} \right] \quad (23.11)$$

where prime means derivative taken with respect to kR . We can also write the β function at $r = R_-$ in terms of u_ℓ inside the potential region

$$\beta_\ell^{(-)} = \left(\frac{d \ln u_\ell}{d \ln r} - 1 \right)_{r=R} \quad (23.12)$$

Requiring $\beta_\ell^{(+)} = \beta_\ell^{(-)}$ will give us the required boundary condition at $r = R$ independent of normalization. If we determine β using this condition, then we can get the scattering phase shift as

$$\tan \delta_\ell = \frac{kR j'_\ell - \beta_\ell j_\ell}{kR n'_\ell - \beta_\ell n_\ell} \quad (23.13)$$

This is the general scheme that we use to determine the solution of the scattering wave function in the scattering region.

Let's work out some realistic things. Again we assume low energy scattering $kR \ll 1$. Actually we don't need to require $V(r)$ to vanish when $r > R$, only that it falls rapid enough to zero. However let's still assume that $V(r) = 0$ for $r > R$. Again we would have similar behavior as in hard sphere scattering, that $\delta_\ell \sim (kR)^{2\ell+1}$ and we can neglect all partial waves except for $\ell = 0$ which is the s -wave. We use the following finite potential

$$V(r) = \begin{cases} 0 & \text{if } r > R \\ V_0 & \text{if } r < R \end{cases} \quad (23.14)$$

where V_0 is a finite number. When $V_0 > 0$ we have a repulsive potential and when $V_0 < 0$ we have an attractive potential. This is a simple case because we can solve for the wave function inside and outside $r = R$ explicitly. The wave function outside is just

$$\varphi = \frac{e^{i\delta_0}}{kr} \sin(kr + \delta_0) \quad (23.15)$$

The radial equation inside is

$$u_0'' + k'^2 u_0 = 0, \quad k'^2 = k^2 - \frac{2mV_0}{\hbar^2} \quad (23.16)$$

So the solution is just

$$u_0 \sim \sin k' r \quad (23.17)$$

Note if $V_0 > \hbar^2 k^2 / 2m$ then k' will be imaginary and $k' = i\kappa$, then the sine will become hyperbolic sine, and we have

$$u_0 \sim \sinh \kappa r \quad (23.18)$$

What we need to do next is to match at the boundary $r = R$ and solve for δ_0 . Before we do the calculation we expect that $\delta_0 < 0$ for repulsive potential and $\delta_0 > 0$ for attractive potential. Let's specialize to the attractive case and consider some extremes. When $k'R = \pi/2$, then because of the assumption $kR \ll 1$ we have $\delta_0 \approx \pi/2$, so $\sin^2 \delta_0 \sim 1$ and the total cross section will be

$$\sigma_{\text{tot}} = \sigma_{\ell=0} = \frac{4\pi}{k^2} \gg 4\pi R^2 \quad (23.19)$$

This is the unitary limit we discussed last time. When we increase even more $k'R = \pi$, then $\delta_0 \approx \pi$ and $\sin \delta_0 \sim 0$. In this case $\sigma_0 \rightarrow 0$ we have almost perfect transmission. This effect is known as Ramsauer-Townsend effect, which says that in an attractive potential we can find some point where the wave is transmitted almost completely without scatter. Let's find the condition that this happens:

$$k' = \frac{\pi}{R} = \sqrt{k^2 + \frac{2m|V_0|}{\hbar^2}} \implies kR = \sqrt{\pi^2 - \frac{2m|V_0|}{\hbar^2}} R^2 \quad (23.20)$$

So this happens when

$$|V_0| \approx \frac{\pi^2 \hbar^2}{2mR^2} \quad (23.21)$$

which means that there appears a bound state in the attractive potential.

Now let's get to the limit where $k \rightarrow 0$, which means very very low energy. In this limit the outside equation for u_0 will be $u_0'' = 0$, which means that the solution is some linear function $u_0 \propto r - a$. If we calculate the logarithmic derivative of this u_0 then we will find

$$\frac{d \ln u_0}{d \ln r} = \frac{1}{r - a} \quad (23.22)$$

For our inside wave function we have already found that

$$\frac{d \ln u_0}{d \ln r} = k \cot \left[k \left(r + \frac{\delta_0}{k} \right) \right] \xrightarrow{k \rightarrow 0} \frac{k}{kr + \delta_0} \implies -\frac{1}{a} = k \cot \delta_0 \quad (23.23)$$

And the partial scattering amplitude is

$$f_0 = e^{i\delta_0} \sin \delta_0 \frac{1}{k} = \frac{1}{k \cot \delta_0 - ik} \quad (23.24)$$

and the total cross section

$$\sigma_{\text{tot}} = 4\pi |f_0|^2 = 4\pi a^2 \quad (23.25)$$

Now we have a characteristic scattering length a with the above expression. This scattering length can be quite different from the range of potential R . The visualization of this length is that if we extrapolate the linear wave function outside the potential region we will hit $u_0 = 0$ at some finite r and that place is a . When the potential is repulsive $V_0 > 0$ then we will find that the wave function inside the potential region to be exponential decaying and a will be positive. When the potential is attractive $V_0 < 0$ then the wave function inside is sinusoidal and a will be negative.

Note in the latter case when $a < 0$ we can tune up the interaction strength, $|a|$ will be larger and larger and at some point we will get $a \rightarrow -\infty$, and after that we will get $a > 0$. Keep tuning up the strength and this pattern will repeat again and again. The relation between a and $|V_0|$ will be something like a $-\tan$ function. When $a \rightarrow \pm\infty$ the scattering hits the unitary limit, and each time when $a \rightarrow \pm\infty$ there is a bound state for the attractive potential.

24 Lecture 24

Last time we discussed scattering at very low energies, i.e. $kR \ll 1$, where R is the scale of the potential. We took R as the range of the potential, but everything we discussed is applicable to the case where the potential is everywhere finite, as long as it decays fast enough at long range. One good example is exponential decay. Power law is also fine, with the exception of Coulomb potential which we will discuss later.

When we are far away from the potential region, the wave function is oscillating, and the only thing we need to know about them is the phase. However in the potential region we have either exponential or sinusoidal behavior, as shown last time. For the range $R < r \ll 2\pi/k$, we could equivalently assume k to be very small and approximate the radial solution to be linear $u_0 \sim r - a$. Recall the radial equation is

$$\frac{d^2 u_\ell}{dr^2} + \left[k^2 - \frac{2m}{\hbar^2} V(r) - \frac{\ell(\ell+1)}{r^2} \right] u_\ell = 0 \quad (24.1)$$

where $u_\ell(r) = rA_\ell(r)$. So when $r \rightarrow \infty$ we can see that both the potential and centrifugal barrier term are unimportant and the solution is just

$$u_0(r) \sim \sin(kr + \delta_0) \sim \sin[k(r - a)] \quad (24.2)$$

where $a = -\delta_0/k$ is called the scattering length. It is just a representative of the scattering phase, and it could be either positive or negative. We discussed at length about the dependence of a on the sign as well as magnitude of V_0 .

In order to match the boundary condition without solving the whole problem (not caring the normalization), we used the logarithmic derivative $d \ln u_0 / dr$. We found that for $E > 0$ the outside wave boundary condition is

$$\frac{d \ln u_0}{dr} = k \cot [k(r - a)] \xrightarrow{k \rightarrow 0} \frac{1}{r - a} \xrightarrow{r \rightarrow 0} -\frac{1}{a} \quad (24.3)$$

However the same thing when $E < 0$ is

$$\frac{d \ln u_0}{dr} = -\kappa \implies \kappa = \frac{1}{a} \implies E_{\text{BS}} = -\frac{\kappa^2 \hbar^2}{2m} = -\frac{\hbar^2}{2ma^2} \quad (24.4)$$

This is how we get the bound state energy.

Recall the scattering wave function is proportional to

$$S_\ell(k) \frac{e^{i(kr - \pi\ell/2)}}{r} - \frac{e^{-i(kr - \pi\ell/2)}}{r} = S_\ell(k) \chi_{\ell,k}^{(+)}(r) - \chi_{\ell,k}^{(-)}(r) \quad (24.5)$$

Recall that unitarity requires that $S_\ell = e^{2i\delta_\ell}$ for real k . However now we need to deal with $k = i\kappa$ which is an imaginary quantity. But in this limit one of the spherical waves will blow up, and what we want to do is to get rid of the incoming wave which will blow up in the limit $r \rightarrow \infty$. Equivalently we can say that at $r \rightarrow \infty$ the first term is infinitely more important, so we should require that $S_\ell(k)$ diverge when $k \rightarrow i\kappa$, which means that $S_\ell(k)$ has a pole there. So we reach a conclusion that whenever the scattering matrix has a pole it corresponds to a bound state in the scattering potential. In anticipation that S_ℓ has a pole, we write

$$S_\ell(k) \xrightarrow{k \approx i\kappa} \frac{?}{k - i\kappa} \xrightarrow{k \rightarrow 0, \kappa \rightarrow 0} \frac{-k - i\kappa}{k - i\kappa} \quad (24.6)$$

so if we work out the scattering amplitude we can get

$$f_\ell = \frac{S_\ell - 1}{2ik} = \frac{1}{-\kappa - ik} \implies f_0 = \frac{1}{k \cot \delta_0 - ik}, a = -\frac{\delta_0}{k} = \frac{1}{\kappa} \quad (24.7)$$

Let's consider the effective potential

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2}{2mr^2} \ell(\ell + 1) \quad (24.8)$$

The second term is the centrifugal barrier term, and for large ℓ it will create a large barrier in the effective potential. When the barrier is infinite we have a stable bound state in the middle when $V(r)$ is more important. However when the barrier is finite the bound state has finite lifetime $\tau = 2\hbar/\Gamma$, and we call it a quasistationary state. This kind of states will bring resonances, and we will talk about it now.

Suppose the energy of the bound state is E_{res} . The scattering phase will approach $\pi/2$ when $E \rightarrow E_{\text{res}}$ and the scattering cross section will reach the unitary limit $\sigma_\ell = 4\pi/k^2$. If we call k_{res} the wave vector corresponding to E_{res} , then what we expect is that $\sigma_\ell(k)$ will have a peak at k_{res} with the value approaching the unitary limit. This is called resonance. Now recall that

$$f_\ell = \frac{1}{k \cot \delta_\ell - ik} \quad (24.9)$$

Note $\cot \delta_\ell = 0$ when $\delta_\ell = \pi/2$, and $\cot \delta_\ell = -\delta_\ell + 2\pi$ in the neighborhood of $\pi/2$. So we write

$$\delta_\ell(E) \approx -\frac{2}{\Gamma}(E - E_{\text{res}}) \implies f_\ell = -\frac{1}{k} \frac{\Gamma/2}{(E - E_{\text{res}}) + i\Gamma/2} \quad (24.10)$$

So the total cross section is

$$\sigma_\ell = |f_\ell|^2 = \frac{4\pi(2\ell + 1)}{k^2} \frac{(\Gamma/2)^2}{(E - E_{\text{res}})^2 + \Gamma^2/4} \quad (24.11)$$

So the resonance is of a Lorentzian type.

Let's return to the S -matrix element and discuss its properties. We write the wave function as

$$a_\ell(k)\chi_{\ell,k}^{(-)}(r) - b_\ell(k)\chi_{\ell,k}^{(+)} \implies S_\ell(k) = \frac{b_\ell(k)}{a_\ell(k)} \quad (24.12)$$

We assume that the scattering is elastic, or E is real, and it only depends on k^2 , so that the solution is invariant under $k \rightarrow -k$ and

$$\chi_{\ell,k}^{(\pm)} \rightarrow (-1)^\ell \chi_{\ell,k}^{(\mp)}, \quad S_\ell(k) = S_\ell^{-1}(-k) \quad (24.13)$$

We also assume T invariance, and if $\psi(x)$ is a solution then $\psi^*(x)$ is also a solution. The whole wave function conjugated will be

$$\left[a_\ell(k)\chi_{\ell,k}^{(-)}(r) - b_\ell(k)\chi_{\ell,k}^{(+)} \right]^* = a_\ell^*(k)(-1)^\ell \chi_{\ell,k}^{(+)} - b_\ell^*(-1)^\ell \chi_{\ell,k}^{(-)} \quad (24.14)$$

However we know that there is no degeneracy here, so the two states ψ and ψ^* will be proportional, and

$$\frac{a_\ell(k)}{b_\ell(k)} = \frac{b_\ell^*(k)}{a_\ell^*(k)} \implies S_\ell(k) = [S_\ell^{-1}(k)]^* \quad (24.15)$$

And in general we also have the relation for complex k

$$S_\ell(k) = [S_\ell^{-1}(k^*)]^* \quad (24.16)$$

These relations provide a strict constraint on the S matrix, and we only need to work out the function in one quadrant, and its value will be determined over the whole complex plane of k .

Recall that the poles of the S -matrix at $k = i\kappa$ correspond to bound states $E = -\kappa^2\hbar^2/2m$. From the above relations, we know that the poles in the upper half plane will correspond to zeros in the lower half plane. On the upper half plane we have $\text{Im } k > 0$ and the poles can only be on the imaginary axis, or the energy will be imaginary. On the lower half plane we can have poles anywhere, however they have to come in pairs due to the above constraints, and the pairs should be symmetric over the imaginary axis. This is because we can kill $\chi^{(+)}$ without worrying about the negative energy problem. These poles correspond to resonances and the resonance energy is

$$E_{\text{res}} = \frac{\hbar^2(\text{Re } k)^2}{2m}, \quad \frac{\Gamma}{2} = \frac{\hbar^2}{2m}\text{Im } k^2 \quad (24.17)$$

25 Lecture 25

We have been discussing the analytic properties of the S -matrix element. Remember the radial wave function was

$$u_{k\ell}(r) = rR_{k\ell}(r) \propto \left[e^{-ikr - i\pi\ell/2} - S_\ell(k)e^{ikr + i\pi\ell/2} \right] \quad (25.1)$$

where $S_\ell(k) = e^{i2\delta_\ell(k)}$ which is followed from unitarity. The general property of S_ℓ were found last time, assuming k is a complex number

$$S_\ell(k) = S_\ell^{-1}(-k), \quad S_\ell(k) = [S_\ell^{-1}(k^*)]^* \quad (25.2)$$

Also the phase is an odd function in k : $\delta_\ell(k) = -\delta_\ell(-k)$. We also found last time that there are two kinds of singularities. The first lies in the upper halfplane, on the imaginary axis. They are the true bound states. The second kind lie on the lower halfplane, which correspond to resonances in the scattering amplitude.

Now in order to derive some more important properties of S -matrix element, we write the function S_ℓ in terms of the energy E instead of k . Recall that the energy of the free particle is just

$$E = \frac{\hbar^2 k^2}{2m} \quad (25.3)$$

The time dependence of the plane wave state is just like in ordinary quantum mechanics

$$\psi(t) \sim \left[e^{-ikr} - S(E)e^{ikr} \right] e^{-iEt/\hbar} \quad (25.4)$$

Now instead of considering the eigenstates of the Hamiltonian, which by definition are states infinitely spread over all the position space, we consider wave packets, which are localized in space. We can consider linear combinations of energy eigenstates

$$\psi_{\text{in}}(t) \propto \int_0^\infty dE' F(E') e^{-ik'r - iE't/\hbar} \quad (25.5)$$

In response to the incoming particles scattering on target, we have the outgoing wave functions

$$\psi_{\text{out}}(t) \propto \int_0^\infty dE' G(E') e^{ik'r - iE't/\hbar} \quad (25.6)$$

Because our Schrödinger equation is linear, therefore our response must be linear, and we can define a Green function, or propagator connecting the outgoing wave function with the incoming wave function

$$\psi_{\text{out}}(t) = \int_{-\infty}^\infty H(t - t') \psi_{\text{in}}(t') dt' \quad (25.7)$$

where H is a function which is independent of either the incoming state or the outgoing state. There is obviously constraints on the kernel H . One is causality, which requires that $H(\tau) = 0$ for $\tau < 0$. This is natural because there can be no outgoing wave when there is no incoming wave. We can Fourier transform the integration kernel and write the above equation into

$$\psi_{\text{out}}(t) = \int dt' \int d\omega \exp[i\omega(t - t')] h(\omega) \psi_{\text{in}}(t') \quad (25.8)$$

Comparing this with our scattering wave function we can identify that

$$h(\omega) = h(E/\hbar) = \frac{1}{2\pi} e^{2ikr} S(E) \quad (25.9)$$

which means that

$$e^{2ikr} S(E) = \int e^{iE\tau/\hbar} H(\tau) d\tau = \int_0^\infty e^{iE\tau/\hbar} H(\tau) d\tau \quad (25.10)$$

The second equality is due to causality. Because we are discussing the analytic properties, here the argument E will be complex $E = \text{Re } E + i \text{Im } E$. If $\text{Im } E < 0$ we have exponential decay in the above integrand, and the integral will converge in general, and the left hand side will not have any singularities. Therefore we have the important result that causality forces $S(E)$ to be analytic at $\text{Im } E > 0$. Notice that this region corresponds to the first quadrant of the complex k plane. It also corresponds to the third quadrant of the k plane but it is on the second part of the Riemann sheet and it is not physical from a QFT point of view. So $S(k)$ will be analytic in the first quadrant in k plane, and we immediately know that it has no poles in both the first and second quadrant, and it has no zeroes in the third and fourth quadrant. This is in accordance with our discussion above, and it is connected to causality in a very basic level.

25.1 Coulomb Scattering

Now we consider Coulomb scattering. The scattering potential is

$$V = \frac{Q_1 Q_2}{r} \quad (25.11)$$

In Born approximation we have

$$f(\theta) = - \left(\frac{2mQ_1Q_2}{\hbar^2} \right) \frac{1}{q^2} = \frac{Q_1Q_2}{4E} \frac{1}{\sin^2 \theta/2} \quad (25.12)$$

therefore we have

$$\frac{d\sigma}{d\Omega} = \frac{Q_1^2 Q_2^2}{16E^2} \frac{1}{\sin^4 \theta/2} \quad (25.13)$$

Notice that if we integrate over all solid angle we will get a divergent total cross section. This is also the classical result of Rutherford scattering. We define the quantum mechanical quantity

$$\gamma = \frac{Q_1 Q_2}{\hbar v} = \frac{m Q_1 Q_2}{\hbar^2 k} \quad (25.14)$$

This parameter is the only meaningful parameter in Coulomb scattering and it provides a convenient scale in the problem. The Born approximation corresponds to $\gamma \rightarrow 0$ while the classical limit corresponds to $\gamma \rightarrow \infty$, so the fact that the Born result corresponds to the classical limit means that the scattering amplitude should be constant for any γ . This peculiarity will be considered in a moment.

Let's consider first an interacting fermion gas. The ground state of the system will be a Fermi sea where all the energy levels below the Fermi level will be occupied whereas none of the levels above is occupied. We denote the Fermi energy by E_F and the corresponding momentum p_F and we have the expression for Fermi momentum

$$p_F = \hbar(3\pi^2 n)^{1/3} \quad (25.15)$$

where n is the number density of the fermions. If we have short range interaction $V = \lambda\delta(\mathbf{x}_1 - \mathbf{x}_2)$, then we will have potential energy per particle equal to λn because each particle sees a potential proportional to the number density at that point. The kinetic energy per particle will be of the order E_F which goes like $n^{2/3}$. Therefore when we increase the density the potential energy contribution gets more important, which is in accordance with our intuition because if we increase the density it will not remain a gas, but will become liquid and solid.

Now if we consider Coulomb interaction $V = Q_1Q_2/r$, then γ will be

$$\gamma = \frac{e^2}{\hbar v} = \frac{e^2 m}{\hbar^2 k_F} \sim n^{-1/3} \quad (25.16)$$

So contrary to the above intuition, when the number density becomes large enough we have an ideal gas, and only when the number density is low will the interaction be important.

Let's come back to Coulomb scattering. Remember we have the transition operator T and we have the operator equation

$$T = V + VG_+T, \quad G_+ = \frac{1}{E - \mathcal{H}_0 + i\epsilon} \quad (25.17)$$

The scattering amplitude is just the matrix element of the transition operator

$$f(\mathbf{k}', \mathbf{k}) = -\frac{4\pi^2 m}{\hbar^2} \langle \mathbf{k}' | T | \mathbf{k} \rangle \quad (25.18)$$

The Born approximation is just $T = V$. This gives us the nice result which is equal to the classical result, however we know from the above equation that $T = V$ is not an exact solution. But this seems to contradict with our previous observation that the first order is already the classical result. The resolution here is that $d\sigma/d\Omega$ is the absolute value squared of the scattering amplitude. In fact it is true that the amplitude does not get corrected in perturbation theory, however the scattering phase is not perturbative, and it has very nontrivial dependence on γ .

Remember we had the following asymptotics

$$\psi^+(\mathbf{x}) \xrightarrow{r \rightarrow \infty} \frac{1}{(2\pi)^{3/2}} \left[e^{-i\mathbf{k}\cdot\mathbf{x}} + \frac{1}{r} e^{ik'r} f \right] \quad (25.19)$$

which we justify by requiring that $V \rightarrow 0$ when $r \rightarrow 0$. We found an equation for $u_{k\ell}$ which is just $rR_{k\ell}(r)$

$$\frac{d^2 u}{dr^2} + \left[k^2 - \frac{2m}{\hbar^2} V(r) - \frac{\ell(\ell+1)}{r^2} \right] u = 0 \quad (25.20)$$

We argued that we can discard the potential term at infinity because of the above assumption, and we also can discard the centrifugal barrier because $r \rightarrow \infty$. Only then do we have the simple equation that $u'' + k^2 u = 0$. However now we need to consider this in more detail. Let us write $u_{k\ell} = \exp[iw(kr)]$ and we can get

$$-(w')^2 + iw'' + k^2 - \frac{2m}{\hbar^2} V - \frac{\ell(\ell+1)}{r^2} = 0 \quad (25.21)$$

To kill the k^2 term we can just shift w' by k and define $w = kr + w_1$, so that we have

$$-2kw'_1 + (w'_1)^2 + iw'' = \frac{2m}{\hbar^2} V + \frac{\ell(\ell+1)}{r^2} \quad (25.22)$$

Now if $V \sim 1/r^{1+\epsilon}$ then w'_1 will also be proportional to $1/r^{1+\epsilon}$ and we will get

$$w_1 \sim \frac{1}{r^\epsilon} \xrightarrow{r \rightarrow \infty} 0 \quad (25.23)$$

However for Coulomb case we have $V \sim 1/r$. Then we have

$$w'_1 \sim \frac{1}{r} \implies w_1 \sim \ln(kr) \xrightarrow{r \rightarrow \infty} \infty \quad (25.24)$$

Therefore we have a problem. Due to this logarithmic divergence we can't apply our previous techniques and we are forced to write

$$u_{k\ell} \propto e^{i(kr + \gamma \ln kr + \dots)} \quad (25.25)$$

and the logarithm in the exponential is called the Coulomb logarithm. Since the current is

$$j = -\frac{i\hbar}{m} [\psi^* \nabla \psi - \psi \nabla \psi^*] \quad (25.26)$$

so the logarithm will cancel in the exponent and because its derivative is $1/r$ so its contribution is not important in $r \rightarrow \infty$ limit. This is why we don't have this effect in the classical limit.

We can write the Schrödinger equation as

$$\left(\nabla^2 + k^2 + \frac{2\gamma k}{r} \right) \psi = 0 \quad (25.27)$$

It is convenient to separate variables in parabolic coordinates (ζ, η, φ) where

$$\zeta = r + z, \quad \eta = r - z, \quad \tan \varphi = \frac{y}{x} \quad (25.28)$$

If we plug this into the above equation and solve it then we get the following result

$$\psi_k(r, \theta) = \left[1 + \frac{1}{i\gamma^2 kr(1 - \cos \theta)} \right] \left\{ e^{ikz - i\gamma \ln[kr(1 - \cos \theta)]} + \frac{f(\theta)}{r} e^{ikr + i\gamma \ln 2kr} \right\} \quad (25.29)$$

where the scattering amplitude is

$$f(\theta) = \frac{\gamma}{2k \sin^2 \theta / 2} \frac{\Gamma(1 - i\gamma)}{\Gamma(1 + i\gamma)} e^{2i\gamma \ln(\sin \theta / 2)} \quad (25.30)$$

Recall some important identities of the Γ function

$$\Gamma(s) = \int_0^\infty e^{-t} t^{s-1} dt, \quad \Gamma(1) = 1, \quad \Gamma(s+1) = s\Gamma(s), \quad \Gamma(s^*) = \Gamma(s)^* \quad (25.31)$$

So the scattering amplitude is only corrected by a complicated phase of Γ functions. If we write out the scattering phase shift then we get

$$e^{2i\delta_\ell} = \frac{\Gamma(\ell + 1 - i\gamma)}{\Gamma(\ell + 1 + i\gamma)} \quad (25.32)$$

Note that Γ function has poles at $s = 0, -1, -2, \dots$. This means that the above combination has poles at $i\gamma = n \geq \ell + 1$. If we write this condition out using the definition of γ then we will have

$$\gamma = \frac{e^2 m}{\hbar^2} \frac{1}{k} = \frac{1}{ka_B} \implies k_n = \frac{1}{-ina_B} \implies E_n = -\frac{\hbar^2}{2ma_B^2 n^2} = -\frac{E_0}{n^2} \quad (25.33)$$

So we recover all of the hydrogen spectrum. By only analyzing the scattering problem and the analytic structure of the scattering amplitude, we recovered the whole spectrum of the bound states of the Coulomb potential.

26 Lecture 26

We have been talking about S -matrices and long formulae for some time. Now let's change topic and discuss the theory of quantum computation. It is usually a mathematical topic, however we know that there are only two kinds of math books, one you can only understand the first page, whereas the other one you can only understand the first line. So we will take a physicist's view, and give an introduction on the theory of quantum information.

26.1 Quantum Computation

We will start with something from information theory. It is a beautiful mathematical theory, but it has correspondence in physics, because information is always encoded in some physical system. This connection between information and physics is two-fold. One is that we realize that some physical systems need information theory to describe properly. One example is Brownian motion. It is just the random motion of atoms and molecules inside matter. The discovery of Brownian motion raised a debate about whether matter is continuous or discrete. For classical mechanics we have nice continuum and time-reversability, however in thermodynamics we have the second law which favors one direction of time. In fact people built the foundations of information theory to explain the various concepts in thermodynamics, and paradoxes like Maxwell's demon. It was during the study of the Maxwell's demon that people realize that there is entropy and energy associated with energy.

This was followed by some development of information theory. One important concept is that it is not the creation of information that requires energy, but rather the destruction of information, in particular shrinking the phase space volume and decreasing entropy. Let's suppose we have a and b which can take 0 or 1 as value. If we do an operation on them which yields 0 if both of the bits are 1, and 0 otherwise, then we obviously destroyed information because the 1 case corresponds to 3 originally different configurations. However if we use the following operation

$$(a, b, c) \longrightarrow \begin{cases} (a, b, \bar{c}) & \text{if } a = b = 1 \\ (a, b, c) & \text{otherwise} \end{cases} \quad (26.1)$$

Note that this operation is reversible and we don't require energy in the process. This is made possible by keeping track of the original data, thereby retaining some junk that we don't actually want.

The above is about how classical computers work. What is different in quantum information? The difference lies in many aspects. The foremost is that quantum mechanics is truly random. In classical mechanics we may approximate random process by some chaotic process, but in quantum mechanics even the simplest process has inherent randomness. The second is that in quantum mechanics some observables don't commute. It is also because of this reason that we have the uncertainty principle, so we have the third difference which is that a quantum system is not exactly reproducible. If we can exactly copy a system, then we can measure different observables on different copies, and this will violate the uncertainty principle. The fourth difference is what we discussed which is the Bell inequalities and absence of hidden variables.

Now let's discuss some basic concepts in information theory. One basic concept is algorithmic complexity. Suppose we have some computational problem and we produce some output from some input. There is always a number N associated with the problem which is the size of the problem. For example if we want to sum two N -digit numbers A and B , then the size of the input will be equal to N . The complexity of the algorithm is just the time it takes to complete this operation, or equivalently the number of elementary

steps. We denote it by $T(N)$. We are mainly interested in the behavior when $N \rightarrow \infty$. In the previous example the elementary step is to sum the individual digits, so we should expect $T(N) \propto N$.

There is an important class of algorithms which is called the P -class, or the “polynomial class”, where $T(N) \leq N^\alpha$. Therefore in these problems the complexity are just polynomials, and they are considered easy to solve. There is another important class of algorithms which is called NP -class, or “non-polynomial class”. The solution of problems in this type can be verified in polynomial time. An example of problems in this class is the decomposition of large numbers into primes. However there is no proof yet that this statement is true, and is still a hot topic today, because it lies in the heart of cryptography. Inside this NP -class we have a subclass which is called the NP -complete. This is defined by the fact that any NP problem in polynomial time can be reduced to an NP -complete problem. An open question now is whether $P = NP$. This is one of the problems from Clay institute which is worth a million dollars.

Let’s come to some concrete problems. Let’s consider the Shor algorithm, which concerns the decomposition of primes. There was a group of scientists in 2009 decomposed a 232-digit number which took 2 years of continuous computation. The best algorithm we have now has the time complexity

$$T(N) = O\left(\exp\left(\frac{64}{9}N\right)^{1/3} (\ln N)^{2/3}\right) \quad (26.2)$$

What Shor proposed was a quantum algorithm, and its complexity is greatly reduced

$$T(N) = O(N^3) \quad (26.3)$$

This is one of the important incentives for people to consider quantum computers.

In classical computers a indivisible unit is a bit, which is just a number a which can take value either 1 or 0. The analog in quantum computers is a *qubit*, or a quantum bit. In our language it is just a 2-level system or a spin 1/2 system. It can take values $|0\rangle$ or $|1\rangle$ and these are the basis states. Any state of qubit is a linear combination of these basis states

$$|\alpha\rangle = a|0\rangle + b|1\rangle, \quad |a|^2 + |b|^2 = 1 \quad (26.4)$$

Now the story becomes exciting when we have n qubits. We can again consider basis states which are direct products of the individual basis states

$$|\mathbf{x}\rangle = |x_1\rangle \otimes |x_2\rangle \otimes \cdots \otimes |x_n\rangle, \quad x_1, \dots, x_n = 0, 1 \quad (26.5)$$

So we will have 2^n different basis states. Therefore we have a huge basis states and this is one of the reasons that this is very efficient. We can draw the diagram of these states

The basis states of n qubits correspond to the 2^n vertices of the n -dimensional hypercube. Now any state can be written as

$$|\alpha\rangle = \sum_{\{\mathbf{x}\}} a_{\mathbf{x}} |\mathbf{x}\rangle \quad (26.6)$$

where we have 2^n coefficients $a_{\mathbf{x}}$.

Now what is quantum computation? The general principle is as follows. Suppose we start with an initial state $|0\rangle = |0, 0, \dots, 0\rangle$. This is called initialization. We construct a unitary operator

$$U = \prod u_i^{\text{st}} \quad (26.7)$$

where U means standard unitary operator. Now we can construct a state $|\alpha\rangle = U|0\rangle$, which will constitute the output of the operation. And in order to read out the output we just need to project the components of $|\alpha\rangle$ by taking its inner product with $|\mathbf{x}\rangle$.

One simple case where quantum computers can beat classical computers is to simulate a quantum system. Obviously if we use a classical computer to simulate the time evolution of n 2-level systems we will encounter problems when n becomes very large, but obviously quantum computers don't have any problem in simulating itself. The more important question is that whether we can make the quantum computer do something other than simulating itself.

Let's assume that $n = 3m$. Instead of measuring all the n qubits at the same time, we can measure the three groups of m qubits separately, but because of correlation this kind of measurement will destroy all the data and we only have an exponentially small portion left. Therefore we should think of the quantum information to be stored in nonlocal correlations, which is also called in the more fashionable word *entanglement*. So we know:

$$|m_1, m_2\rangle \neq |m_1\rangle \otimes |m_2\rangle \quad (26.8)$$

Let's consider an operation which takes $x = 0, 1$ to a result $f(x) = 0, 1$, and this operation takes time T . Now classically to figure out the result of this operation we need to evaluate $f(0)$ and $f(1)$ separately and we need time $2T$. In fact we only want to know whether $f(0) = f(1)$ or not. If it is then f is constant, otherwise we say it is balanced. The classical computer obviously takes $2T$ to answer this question. Now we want to see what can we do in quantum computer case. We introduce the state $|x, y\rangle$ and want to implement an operation: if $f(x) = 1$ then we flip y , and if $f(x) = 0$ then we don't flip y . The matrix of this operation is

$$U_f = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad (26.9)$$

which takes $|x, y\rangle$ to $|f(x), y\rangle$. If we feed the state $|f(x), y\rangle$ to the operator, and take $|y\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$, then we get

$$U_f |f(x), y\rangle = |x\rangle \otimes \frac{1}{\sqrt{2}} (|f(x)\rangle - |1 \oplus f(x)\rangle) = |x\rangle \otimes \frac{1}{\sqrt{2}} (-1)^{f(x)} (|0\rangle - |1\rangle) \quad (26.10)$$

Now if we also take $|x\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, then the result will become

$$\frac{1}{2} \left((-1)^{f(0)} |0\rangle + (-1)^{f(1)} |1\rangle \right) (|0\rangle - |1\rangle) \quad (26.11)$$

So if we write $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ then we will know

$$|+\rangle \rightarrow \begin{cases} |+\rangle & \text{if } f \text{ is constant} \\ |-\rangle & \text{if } f \text{ is balanced} \end{cases} \quad (26.12)$$

This is the main advantage of the quantum computer, as the whole operation is just one step.

27 Lecture 27

27.1 Quantum Information Continued

As discussed last time, although information is not usually in the domain of physics, any information must be carried by physical media. It is especially so for quantum information theory as it is interesting both from a computer scientist's view and from a physicist's view, and it stands at the boundary between the two disciplines.

Recall that classical information is encoded in binary code, or in the unit "bit". A bit is just a number a which can take value 1 or 0. In the quantum case our unit of information becomes qubits, or quantum bits. A qubit is a quantum state in a two-level system with basis $|0\rangle$ and $|1\rangle$. The state of the qubit can be written as $|\alpha\rangle = a|0\rangle + b|1\rangle$. An N -qubit system would then have basis $|\mathbf{x}\rangle = |x_1\rangle \otimes \cdots \otimes |x_N\rangle$, and any state should be expanded in this basis as

$$|\alpha\rangle = \sum_{\{\mathbf{x}\}} a_{\mathbf{x}} |\mathbf{x}\rangle \quad (27.1)$$

Any expansion has 2^N complex coefficients, so a huge amount of information is encoded in this system. This is why quantum information attracts interest in the first place because relatively small number of qubits will give us a huge capacity for storage. The only restriction that we have is just the normalization condition $\sum |a_{\mathbf{x}}|^2 = 1$.

One of the quantities that we want to calculate from an N -qubit system is the correlation function $\langle x_i x_j \rangle$. Note that this is a kind of nonlocal correlation, because as we discussed last time that if we only measure part of the system at a time, then the information will be destroyed. Also as we considered last time that obviously quantum computers can simulate themselves, and as QM is hard to compute classically, we expect that quantum computers will also excel in other difficult computations. And we already saw that the quantum version of the prime decomposition algorithm has much less complexity than the classical algorithm.

So much for review. Let's now move on to the idea of decoherence. We said that quantum mechanics is hard to compute classically, but that is not the only part of physics that is hard to compute. It is also hard to compute, say turbulence in fluid dynamics. In the early stage of computers there were two kinds of computers, namely digital and analog. However nowadays there is no analog computers anymore. One of the reasons is that computers usually require very high precision in the calculation, but chaotic behavior due to the interaction with the environment will magnify any small perturbation and in the end will render the result not reliable.

This is for classical system only. For quantum mechanical system this is even more important as we have the concept of entanglement. Any physically interesting state is not a direct product

$$|\alpha\rangle \neq |\alpha_1\rangle \otimes |\beta_1\rangle \quad (27.2)$$

So most of the quantum information is stored in this entanglement. However we know that the quantum computer can't be decoupled from the environment because we need to read out data and manipulate it. If we think of the environment as a huge quantum system which is out of our control, then we need to study the entanglement between the quantum computer and the environment. Due to this interaction between the computer and the environment, information is gradually lost to the environment from the computer, because the environment constantly "measures" the quantum computer, projecting it to certain basis states. The flow of information into the environment is called *decoherence*.

This dates back to the discussion of the Schrödinger's cat, who exists in a linear combination state

$$|\alpha_{\text{cat}}\rangle = \frac{1}{\sqrt{2}} (|\text{alive}\rangle + |\text{dead}\rangle) \quad (27.3)$$

If there were no environment then a cat will be able to maintain this state. But the interaction with the environment will project its state into either $|\text{alive}\rangle$ or $|\text{dead}\rangle$ in no time. This does not only apply to cats but also to all macroscopic objects, even large atoms.

In order to incorporate this effect between the environment and the quantum computer, we inevitably will start considering methods of error correction. Classically error correction is done by adding redundancy in the data. For example we can replace 0 by a triplet (0, 0, 0) and 1 by (1, 1, 1). Then if our data is somehow changed by one bit, then we can still recover the original data by taking majority. The fact that this method works is based on the following calculation. If p is the probability of flipping one bit due to error, then in the example we considered just now, the probability of having two flips will be $3p^2(1-p)$ and the probability of having three flips will be p^3 . The sum of these two will be $3p^2 - 2p^3$ which is smaller than p if $p < 1/2$. So even if p is not particularly small, this is a relatively reliable method of correcting the error. Now we can obviously generalize this into using n bits as error correction, where we calculate the sum of all the n bits and if the sum is larger than $n/2$ then we take the value to be 1 and if it is smaller than $n/2$ we will take it to be 0. In this method if we write p as the probability of flip and $p = 1/2 - \epsilon$, then the probability of finding $s > 1/2$ will be

$$P\left(s > \frac{1}{2}\right) \propto e^{-2n\epsilon^2} \quad (27.4)$$

Therefore even for very small ϵ we can find large enough redundancy n that can make the probability of changing the information to be exponentially small.

Now in quantum information we have more possible kinds of error. Obviously we can flip one qubit $|0\rangle \rightarrow |1\rangle$. We can also flip the phase of one state $|1\rangle \rightarrow -|1\rangle$ and because of entanglement we will have disastrous results. And instead of in classical case where we either flip or don't flip, we can have small errors in quantum systems

$$|\alpha\rangle = a|0\rangle + b|1\rangle \rightarrow (a + \epsilon)|0\rangle + (b - \epsilon)|1\rangle \quad (27.5)$$

In addition to these, the redundancy method in classical case will not work here because of the non-clonability of the quantum system.

What do we do to correct errors in quantum computers? Instead of using one qubit which has $|0\rangle$ and $|1\rangle$, we use the new states $|\tilde{0}\rangle = |0, 0, 0\rangle$ and $|\tilde{1}\rangle = |1, 1, 1\rangle$. So a usual state will become

$$a|0\rangle + b|1\rangle \rightarrow a|0, 0, 0\rangle + b|1, 1, 1\rangle \quad (27.6)$$

We introduce the notation \oplus which means addition modulo 2, so $0 \oplus 1 = 1$ but $1 \oplus 1 = 0$. Now if we call the basis states $|x, y, z\rangle$ then we can measure $x \oplus z$ and $y \oplus z$. If there is no flipping then both will be 0. If one qubit is flipped then one of these quantities will be 1, and we will know there is something wrong, and find out the error by comparing the result. This can also correct the small ϵ error because if we have some error like that we will have a state $|0, 0, 0\rangle(1 - \epsilon) + |1, 0, 0\rangle\epsilon$ which is detectable by the above method. Now if we want to correct phase error we can enlarge the basis states into 9 qubits and write

$$|\tilde{0}\rangle = \frac{1}{2^{3/2}} (|000\rangle + |111\rangle) (|000\rangle + |111\rangle) (|000\rangle + |111\rangle) \quad (27.7)$$

$$|\tilde{1}\rangle = \frac{1}{2^{3/2}} (|000\rangle - |111\rangle) (|000\rangle - |111\rangle) (|000\rangle - |111\rangle) \quad (27.8)$$

However above discussion relies on the fact that the error events are causally independent. In reality much error are correlated, so we need to take that into account. One thing people use to constrain the change of system due to small perturbations and increase robustness is topology. Because topological properties are global quantities it is stable under small perturbations, so there are many ideas recently to exploit the topological properties of materials to limit the error in the system.

One last thing we want to discuss is called adiabatic computation. The idea is as following. Let's consider a system governed by a Hamiltonian which we can design. For example suppose we have a Hamiltonian $\mathcal{H}_p \{x_i\}$ where x_i are the bits in the sense of classical computing. In classical computing we often want to find the minimum of a function $F \{x_i\}$ with a lot of bits as the argument. We can solve this problem using our Hamiltonian and replacing bits by qubits and tune it such that

$$\mathcal{H}_p \{x_i\} = F \{x_i\} \quad (27.9)$$

If we replace the bits by qubits we will write $x_i \rightarrow \sigma_i^z + 1/2$ where σ^z is the third Pauli matrix. In order to solve the above problem we can write it as a sum with another Hamiltonian

$$\mathcal{H} = s\mathcal{H}_p \{\sigma_i^z\} + (1 - s)\mathcal{H}_0 \{\sigma_i^x\} \quad (27.10)$$

where $\mathcal{H}_0 = \sum \sigma_i^x$. Now obviously we can prepare a state as the ground state of \mathcal{H}_0 . If we adiabatically change the parameter s from $s = 0$ to $s = 1$ then adiabatic evolution will turn the ground state of \mathcal{H}_0 to the ground state of \mathcal{H}_p and then if we measure the resulting ground state we can immediately get the result of the above classical minimizing problem.