The key new concept in quantum mechanics (in comparison with classical mechanics) is the idea of a quantum state. In classical mechanics, we characterize a system using measurable quantities (e.g., mass, velocity, angular velocity...).

In the study of quantum physics, it took people a long time to realize that we can no longer use measurable quantities to describe a system. Instead, we need something that cannot be measured experimentally to describe a system, i.e., a quantum state. Once we realized that, quantum mechanics arises naturally.

1.1. States

**Ket vector:** For a quantum system, a state is described by a ket vector $|\psi\rangle$, which is similar to a column vector in linear algebra.

**Bra vector:** Each state has a conjugate state $\langle\psi|$ which is known as a bra vector. A bra vector is similar to a row vector in linear algebra.

**Inner product:** We can compute the inner product for a bra vector $|\psi\rangle$ and a ket vector $\langle\chi|$: $\langle\chi|\psi\rangle$. The inner product is a scalar (i.e., a complex number), which is similar to the inner product for a column and a row vector.

**Normalization:** For a quantum state, we usually choose a normalization scheme, e.g., $\langle\psi|\psi\rangle = 1$.

**Bottom line:** in quantum mechanics, a quantum state is a normalized vector.

1.2. Operators

Since vectors describe quantum states, what do the matrices represent? They are quantum operators. Each quantum operator is a matrix.

In linear algebra, we know that if we multiply a matrix with a column vector, we get a column vector, so a quantum operator times a ket vector is another ket vector, i.e., a quantum operator “acting” on a quantum state results in another quantum state (up to a complex prefactor)

$$\hat{X} |\psi\rangle = C |\chi\rangle$$

(1.1)

where $C$ is a complex number. To determine the value of $C$, we use the following formula

$$\langle\phi|\hat{X}^\dagger\hat{X}|\psi\rangle = \langle\chi|C^* |\chi\rangle$$

(1.2)

So

$$\langle\phi|\hat{X}^\dagger\hat{X}|\psi\rangle = \langle\chi|C^* |\chi\rangle = |C|^2 \langle\chi|\chi\rangle = |C|^2$$

(1.3)

Here, we used the fact that $\langle\chi|\chi\rangle = 1$. And therefore

$$|C| = \sqrt{\langle\phi|\hat{X}^\dagger\hat{X}|\psi\rangle}$$

(1.4)

or say
\[ C = e^{i\phi} \sqrt{\langle \psi | \hat{X}^\dagger \hat{X} | \psi \rangle} \]  

(1.5)

where \( \phi \) is an arbitrary phase.

Observables: Physics observables corresponds to Hermitian operators \( \hat{X}^\dagger = \hat{X} \). From now on, we only consider Hermitian operators, unless noted otherwise.

Eigenstates: In general, a quantum operator “acting” on a quantum state leads to another quantum state. But for some states, they remains the same, up to a prefactor

\[ \hat{X} | \psi_n \rangle = X_n | \psi_n \rangle \]  

(1.6)

This type of states are known as eigenstates of the operator \( \hat{X} \). And the number \( X_n \) are called eigenvalues. For Hermitian operators, the eigenvalues are always real.

Completeness: All eigenstates of a quantum operator forms a complete orthonormal basis. So for any quantum state, we can write it as

\[ | \psi \rangle = \sum C_n | \psi_n \rangle \]  

(1.7)

### 1.3. Measurement

Consider a quantum state \( | \psi \rangle \), and we try to measure some observable \( \hat{X} \). As shown above, we can write \( | \psi \rangle \) as linear superposition of \( | \psi_n \rangle \)

\[ | \psi \rangle = \sum C_n | \psi_n \rangle \]  

(1.8)

where \( | \psi_n \rangle \)s are eigenstates of \( \hat{X} \). When we measure \( \hat{X} \), for each measurement, the output takes randomly one eigenvalues \( X_n \) of \( \hat{X} \). The probability for getting \( X_n \) is \( C_n^2 \).

Consequence: Only an eigenstate can produce a definite measured value.

### 1.4. Uncertainty relation

If two operators \( \hat{A} \) and \( \hat{B} \) commute with each other, \([\hat{A}, \hat{B}] = 0\), we can find common eigenstates for both operators

\[ \hat{A} | \psi_n \rangle = A_n | \psi_n \rangle \]  

(1.9)

\[ \hat{B} | \psi_n \rangle = B_n | \psi_n \rangle \]  

(1.10)

And thus, for \( | \psi_n \rangle \), we have definite measured values for both physical observables. In other words, the two quantities CAN be measured accurately simultaneously.

If \([\hat{A}, \hat{B}] \neq 0\), in general, it is impossible to get common eigenstates, and thus in general, it is impossible to measure both quantities simultaneously.

### 1.5. A special quantum operator: the Hamiltonian

The Hamiltonian is the quantum operator for energy. It is of special significance in quantum mechanics. This is because the Hamiltonian dictates the time-evolution of quantum states. In other words, if we know a quantum state at time \( t_0 \), the Hamiltonian allows us to predict the state at a later time \( (t > t_0) \), using time-dependent Schrodinger equations. Here, we start from the time-independent Schrodinger equations:

\[ \hat{H} | \psi_n \rangle = E_n | \psi_n \rangle \]  

(1.11)

For a eigenstate of \( \hat{H} \), the time evolution is

\[ | \psi_n(t) \rangle = e^{-i \frac{E_n}{\hbar} (t-t_0)} | \psi_n(t_0) \rangle \]  

(1.12)

For these states, the only thing changes is a phase factor (i.e. the complex phase increases at the rate of \( -E_n / \hbar \)).
Bottom line: up to a phase factor, these eigenstates doesn’t change at all. And thus, they are known as stationary states, and the equation for finding these states, $\hat{H} | \psi_n \rangle = E_n | \psi_n \rangle$, is known as the time-independent Schrödinger equation.

For a general quantum state, $| \psi \rangle$, we write it in terms of $| \psi_n \rangle$s at time $t_0$

$$| \psi(t_0) \rangle = \sum C_n | \psi_n \rangle$$

then at time $t$

$$| \psi(t) \rangle = \sum C_n e^{-\frac{i}{\hbar} \frac{E_n}{T} (t-t_0)} | \psi_n \rangle$$

Note: measuring the time-evolution of a states is in fact measuring the energy. This is the origin for the uncertainly relation between $t$ and $E$. If one wants to measure $E$ accurately, we need to measure the time-evolution accurately (which takes infinite amount to time).

We can verify this time-evolution does follow the Schrödinger equation:

$$i \hbar \frac{\partial}{\partial t} | \psi(t) \rangle = \sum C_n i \hbar \frac{\partial}{\partial t} e^{-\frac{i}{\hbar} \frac{E_n}{T} (t-t_0)} | \psi_n \rangle = \sum C_n e^{-\frac{i}{\hbar} \frac{E_n}{T} (t-t_0)} \hat{H} | \psi_n \rangle = \hat{H} \sum C_n e^{-\frac{i}{\hbar} \frac{E_n}{T} (t-t_0)} | \psi_n \rangle$$

This is the Schrödinger equation!

### 1.6. Wave functions: one set of basis for quantum states

Consider the quantum operator of position $\hat{x}$, which measures the location of a particle. The eigenstates of this operator are states where the particle has a definite position.

$$\hat{x} | x_0 \rangle = x_0 | x_0 \rangle$$

where the eigenvalue $x_0$ is the coordinate of the particle (here we consider a 1D system as an example).

With the eigenstates of $\hat{x}$, we can write down the quantum wavefunctions for any states very easily.

$$\langle x | \psi \rangle = \psi(x)$$

This is the wavefunction. And we can easily prove:

$$| \psi \rangle = \sum | x \rangle \langle x | \psi \rangle = \sum | x \rangle \psi(x) \langle x |$$

i.e., the wavefunction carries all the information of the quantum state. Once we know the wavefunction, we can easily write down the quantum state as shown in the formula above. This is the reason why in QMI, we can mostly use wavefunctions to describe a quantum state.

Note: although quantum wavefunctions probably looks more familiar for you, please keep in mind that quantum state is more fundamental. A quantum wavefunction is a specific way to write down a quantum state, using a specific basis. We could choose any other basis to present the same information. And any basis would describe the same physics equivalently well. So although we are used to quantum wavefunctions in QMI, it is not special and is not better than other ways to present quantum states.
Time-Independent Perturbation Theory

2.1. Overview

2.1.1. General question
Assuming that we have a Hamiltonian,
\[ H = H^0 + \lambda H^1 \]  \hspace{1cm} (2.1)
where \( \lambda \) is a very small real number. The eigenstates of the Hamiltonian should not be very different from the eigenstates of \( H^0 \). If we already know all eigenstates of \( H^0 \), can we get eigenstates of \( H^1 \) approximately?

Bottom line: we are studying an approximate method.

2.1.2. Why perturbation theory?
Why we need to study this approximation methods? (considering the fact that numerical methods can compute the eigenstates very efficiently and accurately for any Hamiltonian that we consider in this course)

Reason number I: It is part of the history (QM was born before electronic computer becomes a powerful tool in scientific research).

Reason number II: It reveals to us universal principles, which are very important and cannot be obtained from just numerical simulations

Reason number III: The idea of perturbation theory has very deep and broad impact in many branches of physics. Perturbation theories is in many cases the only theoretical technique that we have to handle various complex systems (quantum and classical). Examples: in quantum field theory (which is in fact a nonlinear generalization of QM), most of the efforts is to develop new ways to do perturbation theory (Loop expansions, 1/N expansions, 4-\( \epsilon \) expansions).

2.1.3. Assumptions
Assumption #1: we know all eigenstates of \( H^0 \), as well as their corresponding eigenenergies
\[ H^0 \left| \psi_n^0 \right> = E_n^0 \left| \psi_n^0 \right> \]  \hspace{1cm} (2.2)

Assumption #2: we know the perturbation \( H^1 \). What do we mean by knowing \( H^1 \)? Here, we mean that we can write down \( H^1 \) using the complete basis of \( \left| \psi_n^0 \right> \), i.e., we know the value of \( \left< \psi_n^0 \left| H^1 \right| \psi_m^0 \right> \) for any \( m \) and \( n \).

Assumption #3: we only consider quantum states with discrete eigenenergies

In general, the energy spectrum of a quantum system (i.e. all eigenvalues of the Hamiltonian) falls into one of the following three general possibilities

- A discrete spectrum: eigenenergies can only take certain discredited values (example: infinite deep potential wells, e.g. harmonic potential \( E_n = (n + 1/2) \hbar \omega \))
A continuous spectrum: eigenenergies can take any (real) values in certain allowed range (example: a constant potential. Here, any \( E \geq V \) is an eigenenergy)

A mixed spectrum: some parts of the spectrum are continuous, while other parts have discrete eigenenergies. (example: a finite potential well. Here, we may have some discrete states inside the well. But for \( E \) above the top of the potential well, we have a continuous spectrum).

Q: Consider the energy spectrum of an attractive Coulomb (\( 1/r \)) potential. Is it discrete, continuous or mixed?

A: It is mixed. When we consider the an attractive Coulomb potential, we mostly focus on the negative energy states (\( E < 0 \)). This part of the spectrum is discrete, as we all know very well from the study of a hydrogen atom. But if we look at states with positive energies, there is a continuous spectrum for \( E > 0 \). For \( E > 0 \), the system is NOT a bound state, i.e. the proton and the electron doesn’t form an atom. In other words, we have a high probability found the proton and the electron to be separated far from each other. There, the attractive potential is very small and negligible, so we have two free particles and only need to consider their kinetic energies. For free particles, we know that any positive energy is an allowed eigenenergy (i.e. we have a continuum spectrum for \( E > 0 \)).

Bottom line: in this chapter, our perturbation theory only consider discrete spectrum or the discrete part of a mixed spectrum.

Another version of assumption #3: we only consider confined states. (In QM, in most cases, confined states=discrete energy and unconfined states=continuous energy).

Comment: In QM, we only study discrete states in a perturbation theory. But this is NOT true for other branches of physics. For example, in quantum field theory, perturbation theory is applied to continuous spectral.

### 2.2. Non-degenerate Perturbation Theory

#### 2.2.1. Assumptions

**Key assumption:** we consider a specific state \( |\psi_n^0\rangle \). Here, we assume that \(|E_n^0 - E_m^0|\) is much larger than \( \lambda H \) for any other eigenstate \( |\psi_m^0\rangle \)

#### 2.2.2. Preparation #1 wavefunctions

Since the eigenstates of \( H_0 \) form a compete basis, we can write down any quantum state as a linear superposition of \( |\psi_n^0\rangle \)

\[
|\psi\rangle = \sum_n a_n \left| \psi_n^0 \rightangle
\]

(2.3)

Now, if we consider an eigenstate of \( H \), \( |\tilde{\psi}_n\rangle \), it can also be written in a similar fashion

\[
|\tilde{\psi}_n\rangle = \sum_m a_m \left| \psi_m^0 \rightangle
\]

(2.4)

As discussion above, if \( \lambda \) is small, an eigenstate of \( H \) would be similar to an eigenstate of \( H_0 \). Here, we assume that \( |\tilde{\psi}_n\rangle \) is very close to \( |\psi_n^0\rangle \). This means that \( a_n \approx 1 \) and for other values of \( m \neq n \), \( a_m \sim 0 \). To highlight this, we separate the term for \( |\psi_n^0\rangle \) out from the sum

\[
|\tilde{\psi}_n\rangle = a_n \left| \psi_n^0 \rightangle + \sum_{m \neq n} a_m \left| \psi_m^0 \rightangle
\]

(2.5)

It turns out that it is usually more convenient to use unnormalized eigenstates. Now, let us define unnormalized eigenstates of \( H \)

\[
|\psi_n\rangle = \frac{1}{a_n} \left| \tilde{\psi}_n \rightangle = \left| \psi_n^0 \rightangle + \sum_{m \neq n} \frac{a_m}{a_n} \left| \psi_m^0 \rightangle
\]

(2.6)

For simplicity, we will now call \( a_m/ a_n = c_m \)

\[
|\psi_n\rangle = \left| \psi_n^0 \rightangle + \sum_{m \neq n} c_n \left| \psi_m^0 \rightangle
\]

(2.7)

Because \( a_m \sim 0 \) and \( a_n \sim 0 \), we know that \( c_m \sim 0 \) for small \( \lambda \).

Comment #1: This state is NOT normalized

\[
\langle \psi_n | \psi_n \rangle = 1 + \sum_{m \neq n} |c_m|^2 \geq 1
\]

(2.8)
As a result, the leading order terms in the equation is

2.2.4. Schrodinger Equation in the perturbation theory

and for small Eigenenergies of $\lambda$, we can use the Taylor series:

$$ H = H_0 + \lambda H' \psi_0 + \lambda^2 H' \psi_0^2 + \ldots $$

2.2.3. Preparation #2 eigenenergies

Eigenenergies of $H$ are also functions of $\lambda$, and for small $\lambda$, we can use the Taylor series:

$$ E_n(\lambda) = E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \ldots $$

2.2.4. Schrodinger Equation in the perturbation theory

$$ H | \psi_n \rangle = E_n | \psi_n \rangle $$

In the perturbation theory, we need to compute two sets of quantities (1) energy corrections at each order $E_n^1, E_n^2,...$ and (2) wavefunction corrections at each order, $| \psi_n^1 \rangle, | \psi_n^2 \rangle, | \psi_n^3 \rangle$. It turns out that these two set of quantities are entangled together and we need to compute both of them. At each order, we will first compute energy corrections, and then the wavefunction corrections.

2.2.5. Zeroth order

The leading order terms in the equation is $\lambda^0 = \text{constant}$

$$ H_0 | \psi_0 \rangle = E_n^0 | \psi_0 \rangle $$

Comment #3: (almost) any quantum states can be written in the form of Eq. (2.3). This is because $| \psi_n^0 \rangle$ forms a complete basis.

Q: what does the word “almost” mean here?

A: If a state is orthogonal to $| \psi_n^0 \rangle$, we cannot write the state the form of Eq. (2.3). But we don’t need to worry about it here, because we are doing perturbation theory and we know that the eigenstates of $H$ is close to eigenstates of $H_0$. So it can not be orthogonal to $| \psi_n^0 \rangle$.

Bottom line: we are not making any assumptions or approximations here. It is just a new way to write down eigenstates of $H$.

Comment #3: $c_{m,n}$ are functions of $\lambda$, i.e. $c_m(\lambda)$. For small $\lambda$, we can use the Taylor series:

$$ c_m = c_m^{(1)} \lambda + c_m^{(2)} \lambda^2 + c_m^{(3)} \lambda^3 + \ldots $$

Here, the Taylor series doesn’t contain the 0th order term of $\lambda$ (i.e. the constant term). This is because when $\lambda = 0$, $| \psi_n \rangle = | \psi_n^0 \rangle$, and thus $c_m(\lambda) = 0$ at $\lambda = 0$.

As a result,

$$ | \psi_n \rangle = | \psi_n^0 \rangle + \sum_{m \neq n} \sum_{k=1}^{\infty} c_m^{(k)} \lambda^k | \psi_m^k \rangle = | \psi_n^0 \rangle + \sum_{k=1}^{\infty} \lambda^k \sum_{m \neq n} c_m^{(k)} | \psi_m^k \rangle $$

If we define

$$ | \psi_n^k \rangle = \sum_{m \neq n} c_m^{(k)} | \psi_m^k \rangle $$

we get

$$ | \psi_n \rangle = | \psi_n^0 \rangle + \lambda \lambda^1 \lambda^2 \lambda^3 + \ldots $$

This is Eq. [6.5] in the textbook.

Important: $| \psi_n^1 \rangle, | \psi_n^2 \rangle \ldots$ doesn’t contain $| \psi_n^0 \rangle$. In other words, all corrections are orthogonal to $| \psi_n^0 \rangle$.

$$ E_n(\lambda) = E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \ldots $$

This is Eq. [6.6] in the textbook.

$$ H | \psi_n \rangle = E_n | \psi_n \rangle $$

(2.15)

(2.16)

(2.17)

In the perturbation theory, we need to compute two sets of quantities (1) energy corrections at each order $E_n^1, E_n^2,...$ and (2) wavefunction corrections at each order, $| \psi_n^1 \rangle, | \psi_n^2 \rangle, | \psi_n^3 \rangle$. It turns out that these two set of quantities are entangled together and we need to compute both of them. At each order, we will first compute energy corrections, and then the wavefunction corrections.
This is identical to the case of $\lambda = 0$, i.e. the unperturbed system.

### 2.2.6. First order

To the order of $\lambda$, we have

$$H_0 \left[ \psi_n^1 \right] + H^\prime \left[ \psi_n^0 \right] = E_n^0 \left[ \psi_n^0 \right] + E_n^1 \left[ \psi_n^0 \right]$$  \hspace{1cm} (2.19)

Here, we first compute the energy correction $E_n^1$. This is done by multiplying on both sides $\left[ \psi_n^0 \right]

$$\left\langle \psi_n^0 \right| H_0 \left| \psi_n^0 \right\rangle + \left\langle \psi_n^0 \right| H^\prime \left| \psi_n^0 \right\rangle = \left\langle \psi_n^0 \right| E_n^0 \left| \psi_n^0 \right\rangle + \left\langle \psi_n^0 \right| E_n^1 \left| \psi_n^0 \right\rangle$$  \hspace{1cm} (2.20)

For the first term on the l.h.s., we use the fact that

$$\left\langle \psi_n^0 \right| H_0 \left| \psi_n^0 \right\rangle = \left\langle \psi_n^0 \right| E_n^0 \left| \psi_n^0 \right\rangle$$  \hspace{1cm} (2.21)

For the last term on the r.h.s., we use the fact that $E_n^1$ is a number (not a quantum operator), and thus

$$\left\langle \psi_n^0 \right| E_n^1 \left| \psi_n^0 \right\rangle = E_n^1 \left\langle \psi_n^0 \right| \psi_n^0 \right\rangle = E_n^1$$  \hspace{1cm} (2.22)

The first order correction in energy is the expectation value of $H^\prime$.

$$E_n = E_n^0 + \lambda \left\langle \psi_n^0 \right| H^\prime \left| \psi_n^0 \right\rangle + O(\lambda^2) = \left\langle \psi_n^0 \right| H_0 \left| \psi_n^0 \right\rangle + \left\langle \psi_n^0 \right| \lambda H^\prime \left| \psi_n^0 \right\rangle + O(\lambda^2) = \left\langle \psi_n^0 \right| H \left| \psi_n^0 \right\rangle + O(\lambda^2)$$  \hspace{1cm} (2.24)

**Bottom line: to the first order (or say up to corrections at the order of $\lambda^2$), we can use the old wavefunction (the zeroth order wavefunction).**

Then we compute the first order correction for the wavefunction $\left| \psi_n^1 \right\rangle$. To do that, we multiply both sides of the equation with $\left| \psi_m^0 \right\rangle$ where $m \neq n$

$$\left\langle \psi_m^0 \right| H_0 \left| \psi_n^1 \right\rangle + \left\langle \psi_m^0 \right| H^\prime \left| \psi_n^0 \right\rangle = \left\langle \psi_m^0 \right| E_n^0 \left| \psi_n^0 \right\rangle + \left\langle \psi_m^0 \right| E_n^1 \left| \psi_n^0 \right\rangle$$  \hspace{1cm} (2.25)

For the first term on the l.h.s., we use the fact that

$$\left\langle \psi_m^0 \right| H_0 \left| \psi_m^0 \right\rangle = \left\langle \psi_m^0 \right| E_n^0 \left| \psi_m^0 \right\rangle$$  \hspace{1cm} (2.26)

For the two terms on the r.h.s., we use the fact that $E_n^0$ and $E_n^1$ are both numbers (not quantum operators), so $\left\langle \psi_n^0 \right| E_n^0 \left| \psi_n^1 \right\rangle = E_n^0 \left\langle \psi_n^0 \right| \psi_n^1 \right\rangle$ and $\left\langle \psi_n^0 \right| E_n^1 \left| \psi_n^0 \right\rangle = E_n^1 \left\langle \psi_n^0 \right| \psi_n^0 \right\rangle = 0$. Here, we used the fact that when $m \neq n$, the two quantum states are orthogonal and thus

$$\left\langle \psi_m^0 \right| \psi_n^0 \right\rangle = 0.$$  \hspace{1cm} (2.27)

So

$$\left\langle \psi_m^0 \right| \psi_n^1 \right\rangle = \frac{\left\langle \psi_m^0 \right| H^\prime \left| \psi_n^0 \right\rangle}{E_n^0 - E_n^0}$$  \hspace{1cm} (2.28)

According to the definition of $\left| \psi_n^1 \right\rangle$

$$\left| \psi_n^1 \right\rangle = \sum_n c_{m} \left| \psi_m^0 \right\rangle$$  \hspace{1cm} (2.29)

we have

$$c_{m} \left| \psi_m^0 \right\rangle \left| \psi_n^1 \right\rangle = \frac{\left\langle \psi_m^0 \right| H^\prime \left| \psi_n^0 \right\rangle}{E_n^0 - E_m^0}$$  \hspace{1cm} (2.30)

And therefore,
\[ |\psi_n^2| = \sum_{m \neq n} |\psi_m^0| \frac{\langle \psi_m^0 | H' | \psi_n^0 \rangle}{E_m^0 - E_n^0} \]  
(2.31)

So

\[ |\psi_n^2| = |\psi_n^0| + \lambda \sum_{m \neq n} |\psi_m^0| \frac{\langle \psi_m^0 | H' | \psi_n^0 \rangle}{E_m^0 - E_n^0} + \ldots \]  
(2.32)

### 2.2.7. Second order

\[ H_0 |\psi_n^2| + H' |\psi_n^1| = E_n^0 |\psi_n^2| + E_n^1 |\psi_n^1| + E_n^2 |\psi_n^0| \]  
(2.33)

Here, we first compute the energy correction \( E_n^2 \). This is done by multiplying on both sides \( \langle \psi_n^0 | \) \n
\[ \langle \psi_n^0 | H_0 |\psi_n^2| + \langle \psi_n^0 | H' |\psi_n^1| = \langle \psi_n^0 | E_n^0 |\psi_n^2| + \langle \psi_n^0 | E_n^1 |\psi_n^1| + \langle \psi_n^0 | E_n^2 |\psi_n^0| \]  
(2.34)

\[ \langle \psi_n^0 | E_n^0 |\psi_n^2| + \langle \psi_n^0 | H' |\psi_n^1| = \langle \psi_n^0 | E_n^0 |\psi_n^2| + E_n^1 \langle \psi_n^0 |\psi_n^1| + E_n^2 \langle \psi_n^0 |\psi_n^0| \]  
(2.35)

The second term on the r.h.s. is zero, because we required \( \langle \psi_n^0 |\psi_n^1| = 0 \) at the beginning.

\[ E_n^2 = \langle \psi_n^0 | H' |\psi_n^1| \]  
(2.36)

**Bottom line: to compute the second order perturbation, we need to know wavefunction at the first order.**

This conclusion is in fact generically true. We need wavefunction at lower order to compute energy correction at one order higher.

\[ E_n^2 = \langle \psi_n^0 | H' |\psi_n^1| = \sum_{m \neq n} |\psi_m^0| \frac{\langle \psi_m^0 | H' | \psi_n^1|}{E_m^0 - E_n^0} = \sum_{m \neq n} \frac{\langle \psi_n^0 | H' | \psi_m^0|}{E_m^0 - E_n^0} \frac{\langle \psi_m^0 | H' | \psi_n^0|}{E_m^0 - E_n^0} \]  
(2.37)

\[ E_n = E_n^0 + \lambda \sum_{m \neq n} |\psi_m^0| \frac{\langle \psi_m^0 | H' | \psi_n^0|}{E_m^0 - E_n^0} = \sum_{m \neq n} \frac{\langle \psi_n^0 | H' | \psi_m^0|}{E_m^0 - E_n^0} \frac{\langle \psi_m^0 | H' | \psi_n^0|}{E_m^0 - E_n^0} + O(\lambda^2) \]  
(2.38)

The we compute the second order correction for the wavefunction \( |\psi_n^2| \). To do that, we multiply both sides of the equation with \( \langle \psi_m^0 | \) where \( m \neq n \)

\[ \langle \psi_m^0 | H_0 |\psi_n^2| + \langle \psi_m^0 | H' |\psi_n^1| = \langle \psi_m^0 | E_n^0 |\psi_n^2| + \langle \psi_m^0 | E_n^1 |\psi_n^1| + \langle \psi_m^0 | E_n^2 |\psi_n^0| \]  
(2.39)

\[ E_n^0 \langle \psi_m^0 |\psi_n^2| + \langle \psi_m^0 | H' |\psi_n^1| = E_n^0 \langle \psi_m^0 |\psi_n^2| + E_n^1 \langle \psi_m^0 |\psi_n^1| + E_n^2 \langle \psi_m^0 |\psi_n^0| \]  
(2.40)

\[ (E_n^0 - E_m^0) \langle \psi_m^0 |\psi_n^2| = (\psi_m^0 | H' |\psi_n^1| - E_n^1 \langle \psi_m^0 |\psi_n^1| \]  
(2.41)

\[ \langle \psi_m^0 |\psi_n^2| = \frac{\langle \psi_m^0 | H' |\psi_n^1|}{E_n^0 - E_m^0} \]  
(2.42)

According to the definition of \( |\psi_n^2| \)

\[ |\psi_n^2| = \sum_{m \neq n} c_m^{(2)} |\psi_m^0| \]  
(2.43)

We have

\[ c_m^{(2)} = \langle \psi_m^0 |\psi_n^2| \]  
(2.44)
and thus

\[ |\psi_n|^2 = \sum_{m \neq n} \sum_{m' \neq n} \left( \frac{\langle \psi_n^0 | H' | \psi_m^0 \rangle \langle \psi_m^0 | H' | \psi_n^0 \rangle}{E_n^0 - E_m^0} - \frac{\langle \psi_n^0 | H' | \psi_n^0 \rangle}{(E_n^0 - E_m^0)^2} \right) |\psi_m^0|^2 \]  

(2.45)

2.2.8. Third order

Same as the second order, we can use the same method to show that

\[ E_n^3 = \langle \psi_n^0 | H^1 | \psi_n^0 \rangle \]  

(2.46)

So,

\[ E_n^3 = \sum_{m \neq n} \sum_{m' \neq n} \left( \frac{\langle \psi_n^0 | H' | \psi_m^0 \rangle \langle \psi_m^0 | H' | \psi_n^0 \rangle}{(E_n^0 - E_m^0)^2} \right) \left( \frac{\langle \psi_n^0 | H' | \psi_n^0 \rangle}{E_n^0 - E_m^0} \right) - \frac{\langle \psi_n^0 | H' | \psi_n^0 \rangle}{(E_n^0 - E_m^0)^2} \]  

(2.47)

And one can keep doing this for higher and higher order

2.2.9. Summary

For a Hamiltonian

\[ H = H^0 + \lambda H^1 \]  

(2.48)

assuming that we know all the eigenstates of \( H^0 \) (\( |\psi_n^0\rangle \)), and we know the expectation values \( \langle \psi_m^0 | H' | \psi_n^0 \rangle \) for any two eigenstates of \( H_0 \), \( |\psi_m^0\rangle \) and \( |\psi_n^0\rangle \), then we can write down eigenstates of \( H \) as a power series expansions of \( \lambda \)

\[ E_n = E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \ldots = E_n^0 + \lambda \langle \psi_n^0 | H' | \psi_n^0 \rangle + \lambda^2 \sum_{m \neq n} \frac{\langle \psi_n^0 | H' | \psi_m^0 \rangle \langle \psi_m^0 | H' | \psi_n^0 \rangle}{E_n^0 - E_m^0} + \ldots \]  

(2.49)

and

\[ |\psi_n\rangle = |\psi_n^0\rangle + \lambda |\psi_n^1\rangle + \lambda^2 |\psi_n^2\rangle + \ldots = |\psi_n^0\rangle + \lambda \sum_{m \neq n} \frac{\langle \psi_n^0 | H' | \psi_m^0 \rangle}{E_n^0 - E_m^0} + \ldots \]  

(2.50)

2.2.10. Second order perturbation always reduces the energy of the ground state

One key conclusion from the perturbation theory is that the second order correction always makes the energy of the ground state lower (in comparison to the unperturbed one). This can be seen by looking at \( E_0^2 \)

\[ E_0^2 = \sum_{m \neq n} \frac{\langle \psi_n^0 | H' | \psi_m^0 \rangle \langle \psi_m^0 | H' | \psi_n^0 \rangle}{E_n^0 - E_m^0} = \sum_{m \neq n} \left( \frac{\langle \psi_n^0 | H' | \psi_n^0 \rangle}{E_n^0 - E_m^0} \right)^2 \]  

(2.51)

In the numerator, \( \langle \psi_n^0 | H' | \psi_m^0 \rangle \) is the complex conjugate of \( \langle \psi_n^0 | H' | \psi_n^0 \rangle \), so it is \( |\langle \psi_n^0 | H' | \psi_n^0 \rangle|^2 \), which is non-negative

\[ |\langle \psi_n^0 | H' | \psi_n^0 \rangle|^2 \geq 0 \]  

(2.52)

The denominator \( E_n^0 - E_m^0 < 0 \), if \( n \) is the ground state for the unperturbed Hamiltonian (if it is the ground state, then its eigenenergy must be smaller than eigenenergy of any other states). And therefore

\[ \left( \frac{\langle \psi_n^0 | H' | \psi_n^0 \rangle}{E_n^0 - E_m^0} \right)^2 \leq 0 \]  

(2.53)

So

\[ E_0^2 = \sum_{m \neq n} \left( \frac{\langle \psi_n^0 | H' | \psi_n^0 \rangle}{E_n^0 - E_m^0} \right)^2 \leq 0 \]  

(2.54)

The equal sign only arise when \( \langle \psi_n^0 | H' | \psi_n^0 \rangle = 0 \) for ALL \( m \neq n \). (if this is the case, we don’t need to do perturbation theory. The first order perturbation become exact). As long as we ignore this very special case, we find that \( E_n^2 < 0 \) for the ground state, regardless of details.
This conclusion is very important in quantum mechanics, because in many systems, the first order perturbation of the ground state happens to be zero. $E_n^1 = 0$. There,

$$E_n = E_n^0 + \lambda^2 E_n^2 + \ldots$$  \hspace{1cm} (2.55)

The energy correction is dominated by the second order term, which must be negative for the ground state. Without any calculation, we know immediately that

$$E_n < E_n^0$$  \hspace{1cm} (2.56)

In the first homework, we will see that this relation implies that the speed of light in a (linear) medium can only be slower than the vacuum. (i.e., if $E_n > E_n^0$, we will violate the special relativity).

### 2.3. Brillouin-Wigner Perturbation Theory

#### 2.3.1. Negative sides of Rayleigh–Schrödinger perturbation theory

The perturbation theory discussed above is known as Rayleigh–Schrödinger perturbation theory. It is presented for most of the textbooks. However, this approach has some limitations and is not sufficient enough for some cases.

1. Too complicated to go to higher order (e.g. third order or fourth order correction)
2. The physical meaning is less clear (Why do we need to sum over all other quantum state? How should we think about the sum.)
3. One needs to compute energy and wavefunctions at the same time (if we only want to know the eigenenergy, can we compute only energy without bothering to do wavefunction?)

One way to resolve these problems: Brillouin-Wigner Perturbation Theory

#### 2.3.2. Brillouin-Wigner Perturbation Theory

Brillouin-Wigner Perturbation Theory considers the same setup and the final conclusions are exactly the same. However it has a couple of advantages

1. It offers a nice and simple physical interpretation (a baby version of Feynman diagrams used in quantum field theory)
2. It is easier to compute higher order corrections (If we want to compute the eigenenergy using a computer, this perturbation theory just needs one very simple iteration)
3. One can compute energy along, without worry about wavefunctions.

Let’s start from the same setup

$$(H_0 + \lambda H')\ket{\psi_n} = E_n \ket{\psi_n}$$  \hspace{1cm} (2.57)

where $\ket{\psi_n}$ represents the same unnormalized eigenstate of $H$

$$\ket{\psi_n} = \ket{\psi_n^0} + \sum_{m \neq n} c_n \ket{\psi_m^0}$$  \hspace{1cm} (2.58)

We can rewrite the equation above as

$$(E - H_0)\ket{\psi_n} = \lambda H'\ket{\psi_n}$$  \hspace{1cm} (2.59)

and

$$\ket{\psi_n} = \lambda (E - H_0)^{-1} H'\ket{\psi_n}$$  \hspace{1cm} (2.60)

Note: here $(E - H_0)^{-1}$ is the matrix inverse, instead of a number inverse, because $H_0$ is an operator, instead of a number.

**Q:** What is a function of operator? e.g., $f(\hat{Q})$?

**A:** First, we write down the same function as a number function and do a power-law expansion

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots$$  \hspace{1cm} (2.61)
then, \( f(\hat{Q}) \) represents the same power series, but with number \( x \) substitute by operator \( \hat{Q} \)

\[
f(\hat{Q}) = a_0 + a_1 \hat{Q} + a_2 \hat{Q}^2 + \ldots
\]

(2.62)

where \( \hat{Q}^2 = \hat{Q} \hat{Q}, \) etc.

Here, the inverse function of operator \((E - H_0)^{-1}\) shall be understood the same way

Because

\[
\frac{1}{E - \hat{H}_0} = \frac{1}{E} + \frac{x}{E^2} + \frac{x^2}{E^3} + \ldots
\]

(2.63)

we know that

\[
(E - \hat{H}_0)^{-1} = \frac{1}{E} + \frac{\hat{H}_0}{E^2} + \frac{\hat{H}_0^2}{E^3} + \ldots
\]

(2.64)

Now, back to the derivation above:

\[
\langle \phi_m^0 \mid \psi_n \rangle = \langle \phi_m^0 \mid \lambda (E - H_0)^{-1} H' \mid \psi_n \rangle = \frac{\lambda}{E - E_m^0} \langle \phi_m^0 \mid H' \mid \psi_n \rangle
\]

(2.65)

Remember that from the definition, of \( \mid \psi_n \rangle \)

\[
\mid \psi_n \rangle = \mid \psi_0 \rangle + \sum_{m \neq n} c_m \mid \psi_m^0 \rangle
\]

(2.66)

we have

\[
\langle \phi_m^0 \mid \psi_n \rangle = c_m
\]

(2.67)

for any \( m \neq n \). And thus, we get

\[
\mid \psi_n \rangle = \mid \psi_0 \rangle + \lambda \sum_{m \neq n} \langle \phi_m^0 \mid \psi_n \rangle = \frac{\lambda}{E - E_m^0} \langle \phi_m^0 \mid H' \mid \psi_n \rangle
\]

(2.68)

If we define a quantum operator \( \hat{R} \) as

\[
\hat{R} = \sum_{m \neq n} \mid \phi_m^0 \rangle \frac{1}{E - E_m^0} \langle \phi_m^0 \mid
\]

(2.69)

we get

\[
\mid \psi_n \rangle = \mid \psi_0 \rangle + \lambda \hat{R} \hat{H}' \mid \psi_n \rangle
\]

(2.70)

Thus,

\[
(\hat{I} - \lambda \hat{R} \hat{H}') \mid \psi_n \rangle = \mid \psi_n^0 \rangle
\]

(2.71)

where \( \hat{I} \) is the identity operator

So,

\[
\mid \psi_n \rangle = (\hat{I} - \lambda \hat{R} \hat{H}')^{-1} \mid \psi_n^0 \rangle
\]

(2.72)

Again, we emphasize that here, \((\hat{I} - \lambda \hat{R} \hat{H}')^{-1}\) represent matrix inverse. For matrix inverse, we can use Taylor expansions to write it out. We know that

\[
(1 - a)^{-1} = 1 + a + a^2 + a^3 + \ldots
\]

(2.73)

So similarly, we have

\[
(\hat{I} - \lambda \hat{R} \hat{H}')^{-1} = \hat{I} + \lambda \hat{R} \hat{H}' + \lambda^2 \hat{R} \hat{H}' \hat{R} \hat{H}' + \lambda^3 \hat{R} \hat{H}' \hat{R} \hat{H}' \hat{R} \hat{H}' + \ldots
\]

(2.74)

So we have

\[
\mid \psi_n \rangle = (\hat{I} - \lambda \hat{R} \hat{H}')^{-1} \mid \psi_n^0 \rangle = \mid \psi_0 \rangle + \lambda \hat{R} \hat{H}' \mid \psi_n \rangle + \lambda^2 \hat{R} \hat{H}' \hat{R} \hat{H}' \mid \psi_n \rangle + \lambda^3 \hat{R} \hat{H}' \hat{R} \hat{H}' \hat{R} \hat{H}' \mid \psi_n \rangle + \ldots
\]

(2.75)
So we find that

\[ |\psi_n^k\rangle = (\hat{R} \hat{H}^k) |\psi_n^0\rangle \]  

(2.76)

Previous, we found that

\[ E_n^1 = \left< \psi_n^0 | H^+ | \psi_n^0 \right> \]  

(2.77)

\[ E_n^2 = \left< \psi_n^0 | H^+ | \psi_n^1 \right> \]  

(2.78)

\[ E_n^3 = \left< \psi_n^0 | H^+ | \psi_n^2 \right> \]  

(2.79)

In fact, we can use the same procedure to show that for \( k \)th order,

\[ E_n^k = \left< \psi_n^0 | H^+ | \psi_n^{k-1} \right> \]  

(2.80)

Because we have found that

\[ |\psi_n^{k-1}\rangle = \left( \hat{R} \hat{H}^k \right)^{k-1} |\psi_n^0\rangle \]

\[ E_n^k = \left< \psi_n^0 | H^+ | \psi_n^{k-1} \right> = \left< \psi_n^0 | H^+ (\hat{R} \hat{H})^{k-1} | \psi_n^0 \right> \]  

(2.81)

So, we have

\[ E_n^1 = \left< \psi_n^0 | H^+ | \psi_n^0 \right> \]  

(2.82)

\[ E_n^2 = \left< \psi_n^0 | H^+ R H^+ | \psi_n^0 \right> = \sum_{m \neq n} \left< \psi_m^0 | H^+ | \psi_n^0 \right> \frac{1}{E_n^0 - E_m^0} \left< \psi_n^0 | H^+ | \psi_n^0 \right> \]  

(2.83)

\[ E_n^3 = \left< \psi_n^0 | H^+ R H^+ R H^+ | \psi_n^0 \right> \]  

(2.84)

\[ \ldots \]  

(2.85)

From these formula we see a pattern.

1. For any \( E_n^k \), if we look at the formula from right to left, one always start from unperturbed state \( |\psi_n^0\rangle \) and eventually goes back to the same state \( \langle \psi_n^0 | \).

2. In the path from \( |\psi_n^0\rangle \) to \( \langle \psi_n^0 | \), we go through several intermediate states \( |\psi_m^0\rangle, |\psi_m'^0\rangle \ldots \). For \( k \)th order perturbation, we have \( k - 1 \) intermediate states.

3. To turn from a state to another along the path (e.g. from \( n \) to \( m' \) or from \( m' \) to \( m \) in \( E_n^3 \)), we use the perturbation \( H^+ \).

4. For each intermediate state, we have a denominator \( \frac{1}{E_n^0 - E_m^0} \).

### 2.3.3. Diagrammatic representation

We can represent the \( E_n^k \) using diagrams.

1. For each intermediate state, we represent \( \frac{1}{E_n^0 - E_m^0} \) as a solid line with integer \( m \) labeling the state.

2. For each \( \langle \psi_m^0 | H^+ | \psi_m'^0 \rangle \), we represent it as a dot. And we use \( V_{nn} \) to represent \( \langle \psi_n^0 | H^+ | \psi_n^0 \rangle \)

3. Connect everything together in the same order as in \( E_n^k \)

4. At the two ends of the line, we use two short line to present that we start from and end at the same state \( |\psi_n^0\rangle \)

First order:

\[ V_{nn} |n \rangle \langle n| \]
Second order:

![Diagram of second order perturbation term](image)

Third order

![Diagram of third order perturbation term](image)

By making the line longer, we can write down easily perturbation terms to any order.

**Relations to QFT:**

In QFT, we use very similar diagrams, known as the Feynman diagrams. There, solid lines are propagator of a particle \( \frac{1}{\omega - \epsilon_0} \) where \( \omega \) is frequency, pretty much the same as energy \( E_n \) and \( \epsilon_0 \) is the unperturbed energy of the particle (energy ignore interactions between particles). In fact, the diagrams we show here are baby versions of the diagrams of Feynman.

**Physics meaning discussed in class:** (example: two electrons exchange photons to get E&M interactions).

### 2.3.4. How to compute the energy using Brillouin-Wigner Perturbation Theory?

First, let’s define some abbreviation to make the formula shorter,

\[
V_{ij} = \langle \psi^0_i | H^* | \psi^0_j \rangle
\]

and thus

\[
E_n = E_n^0 + \lambda V_{nn} + \lambda^2 \frac{V_{nm} V_{mn}}{E_n - E_n^0} + \lambda^3 \frac{V_{nm} V_{mn} V_{mn}}{(E_n - E_n^0)(E_n - E_m^0)} + \ldots
\]

(2.87)

Here all the \( m \)s are summed over but they cannot be the same as \( n \). It may looks like that we can find \( E_n \) using this formula, but it is not quite the case yet. This is because on the r.h.s., the denominator contains also \( E_n \), i.e. IT is a equation for \( E_n \) and \( E_n \) arises on both sides.

This equation can be solved easily using **iterative method (e.g. using a computer code)**. One start from zeroth order, and then go to first, second, third order ..., every time we need \( E_n \) in the \( k \)th order calculation, we just use the \((k - 1)\)th order \( E_n \) on the r.h.s. Here is how it is done

**First run**

\[
E_n^{(1)} = E_n^0 + \lambda V_{nn}
\]

(2.88)

**Second run**

\[
E_n^{(2)} = E_n^0 + \lambda V_{nn} + \lambda^2 \frac{V_{nm} V_{mn}}{E_n^{(1)} - E_n^0}
\]

(2.89)

**Third run**

\[
E_n^{(3)} = E_n^0 + \lambda V_{nn} + \lambda^2 \frac{V_{nm} V_{mn}}{E_n^{(2)} - E_n^0} + \lambda^3 \frac{V_{nm} V_{mn} V_{mn}}{(E_n^{(2)} - E_n^0)(E_n^{(2)} - E_m^0)}
\]

(2.90)

**Fourth run**
2.3.5. Preparation

Consider the following function

\[ f(x) = x^2, \quad g(x) = x^2(1 + ax + bx^2 + cx^3 + \ldots) \]  

(2.92)

If we want to keep \( f(x) \) to \( O(x^n) \), we only need to keep \( g(x) \) to \( O(x^{n+2}) \). Similarly, for the following function

\[ f(x) = \frac{x^2}{g(x)} = \frac{x^2}{1 + ax + bx^2 + cx^3 + \ldots} \]  

(2.93)

If we want to keep \( f(x) \) to \( O(x^n) \), we only need to keep \( g(x) \) to \( O(x^{n+2}) \). This will be something that useful for us latter

2.3.6. iterative method

\[ E_n = E_n^0 + \lambda V_m + \lambda^2 \frac{V_{m,m} V_{m,n}}{E_n^0 - E_m^0} + \lambda^3 \left( \frac{V_{m,m} V_{m,n} V_{n,n}}{(E_n^0 - E_m^0)(E_n^0 - E_m^0)} \right) + \lambda^4 \left( \frac{V_{m,m} V_{m,n} V_{n,n} V_{m,n}}{(E_n^0 - E_m^0)(E_n^0 - E_m^0)(E_n^0 - E_m^0)} \right) + \ldots \]  

(2.91)

Another option is using analytic methods, as will be discussed below.

...
$$E_n = E_n^0 + \lambda V_{mn} + \lambda^2 \frac{V_{mn} V_{mn'}}{E_n^0 - E_m^0} - E_{m}^0$$  

$$\lambda^3 \frac{V_{mn} V_{mn'} V_{mn''}}{(E_n^0 - E_m^0)(E_n^0 - E_{m'}^0)(E_n^0 - E_{m''}^0)} + \lambda^4 \frac{V_{mn} V_{mn'} V_{mn''} V_{mn'''}}{(E_n^0 - E_m^0)^2(E_n^0 - E_{m'}^0)^2(E_n^0 - E_{m''}^0)^2} + O(\lambda^5)$$  

(2.101)

$$E_n = E_n^0 + \lambda V_{mn} + \lambda^2 \frac{V_{mn} V_{mn'} V_{mn''}}{E_n^0 - E_m^0} + \lambda^3 \frac{V_{mn} V_{mn'} V_{mn''} V_{mn'''}}{(E_n^0 - E_m^0)(E_n^0 - E_{m'}^0)(E_n^0 - E_{m''}^0)} + \lambda^4 \frac{V_{mn} V_{mn'} V_{mn''} V_{mn'''} V_{mn''''}}{(E_n^0 - E_m^0)^2(E_n^0 - E_{m'}^0)^2(E_n^0 - E_{m''}^0)^2(E_n^0 - E_{m'''})} + O(\lambda^5)$$  

(2.102)

### 2.4. Degenerate Perturbation Theory

In the previous section, we studied the effect of a small perturbation $\lambda H'$ on an eigenstate of $H_0$, $|\psi_0^0\rangle$. The key assumption there is that before we turn on the perturbation (i.e. at $\lambda = 0$), the eigenenergies of all other eigenstates of $H_0$ are very far away from $E_n^0$

$$|E_n^0 - E_m^0| > |\langle \psi_0^0 | \lambda H' | \psi_0^0 \rangle|$$  

(2.103)

This section, we will consider the opposite situation, where there is at least one other eigenstate of $H_0$ which has the same eigenenergy as $|\psi_0^0\rangle$. Two states having the same eigenenergy is known as “degeneracy”. So this perturbation theory is known as the degenerate perturbation theory.

#### 2.4.1. Why non-degenerate perturbation theory fails in the presence of degeneracy?

In the presence of degeneracy, the perturbation theory that we learned before will fail. To see this, we just need to look at the second order perturbation of the eigenenergy

$$E_n = E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \ldots = E_n^0 + \lambda \langle \psi_0^0 | H' | \psi_0^0 \rangle + \lambda^2 \sum_{m \neq n} \frac{\langle \psi_0^0 | H' | \psi_m^0 \rangle \langle \psi_m^0 | H' | \psi_0^0 \rangle}{E_n^0 - E_m^0} + \ldots$$  

(2.104)

Here, we focus on the second order correction:

$$\lambda^2 \sum_{m \neq n} \frac{\langle \psi_0^0 | H' | \psi_m^0 \rangle \langle \psi_m^0 | H' | \psi_0^0 \rangle}{E_n^0 - E_m^0}$$  

(2.105)

If $H_0$ has another eigenstate $|\psi_m^0\rangle$ with the same eigenenergy, at least one term in this sum will have zero in the denominator and thus will diverge, i.e., when $E_n^0 = E_m^0$, $\frac{1}{E_n^0 - E_m^0} \rightarrow \infty$, and thus the theory becomes ill-defined.

**NOTE:** the same divergence will arise also in higher order corrections. But there is no divergence in the first order correction $E_n^1$.

In power-law expansions, infinite coefficient doesn’t always mean singularity. It means that we missed something in the lower order correction. Here is a simple example: Let’s consider a function $f(x)$, which can be written as the following Taylor expansion at small $x$

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots$$  

(2.106)

Now, assume that I made a mistake in the Taylor expansion for the coefficient $a_1$. Instead of the correction value, $a_1$, I used a wrong coefficient for the linear term, say $b_1$.

$$f(x) = a_0 + b_1 x + (a_1 - b_1) x + a_2 x^2 + \ldots$$  

(2.107)

In other words, here I missed part of the linear term, $(a_1 - b_1) x$. And thus coefficients of the higher order terms will also need to be adjusted to absorb this mistake. Let’s try to use the $x^2$ term to correct this error, i.e.

$$f(x) = a_0 + b_1 x + \left(\frac{a_1 - b_1}{x} + a_2\right) x^2 + \ldots$$  

(2.108)
Let me define $b_2 = a_2 + \frac{a_1 - b_1}{x}$

$$f(x) = a_0 + b_1 x + b_2 x^2 + \ldots$$

Now, once again, I wrote my function as a power-law expansion. Because I used a wrong coefficient for the linear term, $b_1$, my second order term needs to use this new coefficient. This new coefficient $b_2$ is infinite at small $x$. This is transparent if we notice that when $x \to 0$

$$b_2 = a_2 + \frac{a_1 - b_1}{x} \to \infty$$

**Bottom line:** infinite coefficient in the second order term (and higher order term) means that the first order result is incorrect and needs to be revised.

### 2.4.2. What to do?

Here, let’s first take another look at the second order correction

$$E_n^2 = \sum_{m,n} \frac{\langle \phi_n^0 \mid H^\dagger \mid \phi_m^0 \rangle \langle \phi_m^0 \mid H^\dagger \mid \phi_n^0 \rangle}{E_m^0 - E_n^0}$$

As we know, the problem arises because $E_n^0 = E_m^0$ for certain $m$, and thus we get $\frac{1}{0} = \infty$. To avoid this singularity, the only thing that we need to do is to request that the numerator also vanish whenever the denominator is zero. i.e., if $E_n^0 = E_m^0$, we must make sure that $\langle \psi_n^0 \mid H^\dagger \mid \psi_n^0 \rangle = 0$.

**NOTE:** the two factors in the numerator are complex conjugate to each other: $\langle \phi_n^0 \mid H^\dagger \mid \phi_m^0 \rangle = \langle \phi_m^0 \mid H^\dagger \mid \phi_n^0 \rangle^*$, and thus if one of them is zero, the other is also zero.

**Bottom line:** For degenerate states, before we start the procedure described in the non-degenerate perturbation theory, we need to first make sure that for any degenerate states, $\langle \phi_n^0 \mid H^\dagger \mid \phi_n^0 \rangle = 0$

### 2.4.3. Whenever there is an degeneracy, we have an option to choose the basis

A good example, a free particle. Consider a free particle with mass $m$.

$$H_0 = \frac{p^2}{2 m} = -\frac{\hbar^2}{2 m} \frac{d^2}{dx^2}$$

The eigenstates of $H_0$ arises in pairs (i.e. there is a degeneracy for any excited states). The static Schrodinger equation here is

$$-\frac{\hbar^2}{2 m} \frac{d^2}{dx^2} \psi(x) = E \psi(x)$$

It is a second order differential equation and we know the solution are just plane waves

$$\psi = A e^{ikx} + B e^{-ikx}$$

The eigenenergy for this state is $E = \frac{p^2}{2 m} = (h k)^2 / 2 m$, i.e. the kinetic energy. Here, $A$ and $B$ are two arbitrary coefficients.

For each fixed $k$, we have one eigenenergy $E = (h k)^2 / 2 m$, but infinite number of eigenstates $\psi = A e^{ikx} + B e^{-ikx}$, i.e. a **degeneracy**. This example is known as two-fold degeneracy, or we say that two states have the same energy. The reason we say “two states” here is because not all the eigenstates are linear independent. In fact, we just need two states, $e^{ikx}$ and $e^{-ikx}$, all other eigenstates can be written as linear superposition of these two. **Bottom line:** two-fold degeneracy means that any linear combination of these two states is an eigenstate of $H_0$ with the same eigenenergy.

Now, let’s look at the same second order differential equation again.

$$-\frac{\hbar^2}{2 m} \frac{d^2}{dx^2} \psi(x) = E \psi(x)$$

we know that we can also write the solution for this equation as

$$\psi = C \cos kx + D \sin ky$$
i.e., instead of using exponentials, we can use sin or cos functions to represent plane waves. Here, once again we find infinite number of eigenstates with the same eigenenergy, and once again, they are not all independent. We just need two states \( \cos kx \) or \( \sin kx \). And all other eigenstates are just linear superpositions of them two. So, again, we reach the same conclusion, the system have a two-fold degeneracy. But early on, we said that the two states are \( e^{i kx} \) and \( e^{-i kx} \), but now for the two degenerate states, we use \( \cos kx \) or \( \sin kx \).

These different choices are just different basis to represent all the eigenstates. We can choose to use \( e^{i kx} \) and \( e^{-i kx} \) or \( \cos kx \) or \( \sin kx \). There is no difference between them. In fact, we can choose any two linear independent states

\[
\psi_1 = A_1 e^{i kx} + B_2 e^{-i kx} \tag{2.117}
\]

\[
\psi_4 = A_2 e^{i kx} + B_2 e^{-i kx} \tag{2.118}
\]

And then, we can say that we have two degenerate states \( \psi_1(x) \) and \( \psi_2(x) \). and then, we can represent any other eigenstates (with the same eigenenergy) as

\[
\psi(x) = X\psi_1(x) + Y\psi_2(x) \tag{2.119}
\]

Q: Why do we usually use \( e^{i kx} \) and \( e^{-i kx} \) or \( \cos kx \) or \( \sin kx \)? Why not use \( e^{i kx} \) and \( \cos kx \).

A: There is no problem (mathematically) if we choose to use \( e^{i kx} \) instead of \( e^{i kx} \) or \( \sin kx \). Here, once again we find infinite number of eigenstates with the same eigenenergy, and once again, they are not all independent. We just need two states \( \cos kx \) or \( \sin kx \). However, for convenience, it is usually better using orthonormal bases

\[
\int dx (e^{i k_1 x})^* e^{i k_2 x} = 2\pi \delta(k_1 - k_2) \tag{2.120}
\]

\[
\int dx (\cos kx)^* \sin kx = 0 \tag{2.121}
\]

Q: Why do we use \( e^{i kx} \) and \( e^{-i kx} \) more often than \( \cos kx \) or \( \sin kx \) in quantum mechanics?

A: For \( H_0 \), there is little difference between the two choices. However, if we consider other quantum operators, like momentum, \( e^{i kx} \) and \( e^{-i kx} \) is a better choice. This is because \( e^{i kx} \) and \( e^{-i kx} \) are eigenstates of the momentum operator too! So they have not only well-defined energy, but also well defined momenta \((h k \text{ and } -h k \text{ respectively})\). \( \cos kx \) or \( \sin kx \) don’t have well defined momentum. They have 50% chance having momentum \( h k \) and another 50% chance having momentum \( -h k \).

**Bottom line: when you cannot decide which choice of basis is better, look at another quantum operator.**

These conclusions are true generically. Let’s start from two-fold degenerate. Assuming that for \( H_0 \), there are two degenerate eigenstates:

\[
H_0 \left| \psi_0^0 \right> = E^0 \left| \psi_0^0 \right>
\]

and

\[
H_0 \left| \psi_0^1 \right> = E^0 \left| \psi_0^1 \right>
\]

We assume that these two states are orthogonal to each other (otherwise, we make them orthogonal, using Gram-Schmidt procedure). We assume that \( \left< \psi_0^0 \right| H^* \left| \psi_0^0 \right> \neq 0 \). Here we first prove a fact: if \( \left| \psi_0^0 \right> \) and \( \left| \psi_0^1 \right> \) are both eigenstates of \( H_0 \) and they have the same eigenvalue \( E^0 \), then any linear superposition of \( \left| \psi_0^0 \right> \) and \( \left| \psi_0^1 \right> \) is also an eigenstate of \( H_0 \) with the same eigenenergy. Let’s define

\[
\left| \phi^0 \right> = \alpha \left| \psi_0^0 \right> + \beta \left| \psi_0^1 \right>
\]

\[
H_0 \left| \phi^0 \right> = H_0 (\alpha \left| \psi_0^0 \right> + \beta \left| \psi_0^1 \right>) = \alpha H_0 \left| \psi_0^0 \right> + \beta H_0 \left| \psi_0^1 \right> = \alpha E_0 \left| \psi_0^0 \right> + \beta E_0 \left| \psi_0^1 \right> = E_0 (\alpha \left| \psi_0^0 \right> + \beta \left| \psi_0^1 \right>) = E_0 \left| \phi^0 \right>
\]

**Bottom line: if \( H_0 \) has two degenerate eigenstates \( \left| \psi_0^0 \right> \) and \( \left| \psi_0^1 \right> \), we have infinite eigenstates with the same eigenvalue \( \left| \phi^0 \right> = \alpha \left| \psi_0^0 \right> + \beta \left| \psi_0^1 \right> \)**

To represent these infinite eigenstates, we need to choose two states as basis, e.g. \( \left| \psi_0^0 \right> \) and \( \left| \psi_0^1 \right> \). Then, any eigenstates with eigenenergy \( E^0 \) can be written as a superposition of them. When we have one set of basis, we know that we can choose another set of basis (i.e. we can change to a different set of basis): for example, we can define \( \left| \psi_1^0 \right> \) and \( \left| \psi_1^1 \right> \), where

\[
\left| \psi_1^0 \right> = \alpha_1 \left| \psi_0^0 \right> + \beta_1 \left| \psi_0^1 \right>
\]

(2.125)
As mentioned above, in general, $W$ is a Hermitian matrix ($W = W^\dagger$). Here, we request $\psi_a$ and $\psi_b$ to be orthogonal to each other (otherwise, we make them orthogonal, using Gram-Schmidt procedure).

$$\langle \psi_a | H' | \psi_b \rangle = 0$$

If so, the native perturbation theory will have singularity at the second order. Now, we learned that we can choose to use a different set of unperturbed eigenstates $|\psi_1^{(0)}\rangle$ and $|\psi_2^{(0)}\rangle$, so can we make $\langle \psi_1^{(0)} | H' | \psi_2^{(0)} \rangle = 0$? If so, it will save the day and get ride of the singularity. The way to do it is very simple. As we know early on, if we cannot decide which basis to use, we shall look at another quantum operator. Here we do have one more quantum operator, which is $H'$.

We first write down a $2 \times 2$ matrix

$$W = \begin{pmatrix} \langle \psi_a | H' | \psi_a \rangle & \langle \psi_a | H' | \psi_b \rangle \\ \langle \psi_b | H' | \psi_a \rangle & \langle \psi_b | H' | \psi_b \rangle \end{pmatrix} \quad (2.130)$$

To make the formula shorter, we define

$$W_{ij} = \langle \psi_i | H' | \psi_j \rangle \quad (2.131)$$

So

$$W = \begin{pmatrix} W_{aa} & W_{ab} \\ W_{ba} & W_{bb} \end{pmatrix} \quad (2.132)$$

1. $W$ is a Hermitian matrix ($W^\dagger = W$), so its eigenvalues are real

This is pretty straightforward to prove, because $\langle \psi_1 | X^\dagger | \psi_2 \rangle = \langle \psi_2 | X^\dagger | \psi_1 \rangle$. In particular, if $X$ is an Hermitian operator (the quantum operator of any physics observable is Hermitian), we have $X = X^\dagger$ and thus $\langle \psi_1 | X | \psi_2 \rangle^* = \langle \psi_2 | X | \psi_1 \rangle$. Thus it is easy to notice that

$$W_{aa} = W_{aa}^* \quad \text{and} \quad W_{bb} = W_{bb}^* \quad \text{and} \quad W_{ba} = W_{ab}^* \quad \text{and} \quad W_{ab} = W_{ba}^* \quad (2.133)$$

$$W^\dagger = (W^\dagger W) = \begin{pmatrix} W_{aa}^* & W_{ab}^* \\ W_{ba}^* & W_{bb}^* \end{pmatrix} = \begin{pmatrix} W_{aa} & W_{ab} \\ W_{ba} & W_{bb} \end{pmatrix} = W \quad (2.134)$$

2. $W$ has two eigenvalues $E_+$ and $E_-$, and each of them has a vector, $\begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix}$ for $E_+$ and $\begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix}$ for $E_-

$$\begin{pmatrix} W_{aa} & W_{ab} \\ W_{ba} & W_{bb} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} = E_+ \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} \quad (2.135)$$

and

$$\begin{pmatrix} W_{aa} & W_{ab} \\ W_{ba} & W_{bb} \end{pmatrix} \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix} = E_- \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix} \quad (2.136)$$

where $E_+$ and $E_-$ are the two eigenvalues.

3. We can use these two eigenvectors to define our $|\psi_a^{(0)}\rangle$ and $|\psi_b^{(0)}\rangle$ as

$$|\psi_a^{(0)}\rangle = \alpha_1 |\psi_a^{(0)}\rangle + \beta_1 |\psi_b^{(0)}\rangle \quad (2.137)$$

$$|\psi_b^{(0)}\rangle = \alpha_2 |\psi_a^{(0)}\rangle + \beta_2 |\psi_b^{(0)}\rangle \quad (2.138)$$

As will be shown below, these two states are precisely what we should use for the perturbation theory.
4. The two eigenvalues are the first order corrections to the eigenenergy

\[ E_1 = E^0 + \lambda E_+ + O(\lambda^2) \]  
(2.139)

\[ E_2 = E^0 + \lambda E_- + O(\lambda^2) \]  
(2.140)

Q: Why do we have two eigenenergies here?

A: Because we started from two degenerate eigenstates. At \( \lambda = 0 \), the two states \(| \psi_1^0 \rangle \) and \(| \psi_2^0 \rangle \) have the same energy. Now, if we turn on a small perturbation \( \lambda H' \), we find that these two states (in general) have different eigenenergies. One of them is \( E_1 = E^0 + \lambda E_+ + O(\lambda^2) \) and the other \( E_2 = E^0 + \lambda E_- + O(\lambda^2) \).

NOTE #1: We say that the perturbation “lifted the degeneracy”.

NOTE #2: After we lift the degeneracy, \(| \psi_1^0 \rangle \) and \(| \psi_2^0 \rangle \) no longer have the same energy. If the perturbation is small enough, we can now do non-degenerate perturbation theory, i.e. problem solved.

2.4.5. A key conclusion: in quantum mechanics, perturbations will in general lift all degeneracy, unless there is a reason saying that the degeneracy shall not be lifted.

In general, in the study of quantum physics, we can never include all terms in the Hamiltonian in our theoretical calculation. We always need some approximations (i.e. drop some small/unimportant part of the Hamiltonian). For example, in the study of a Hydrogen atom, we ignored relativistic effects. We also ignored the magnetic interactions between the electron and the nucleon (remember that both particles have spin. Whenever a charged particle starts to spin, there is a magnetic dipole. Because both the electron and the proton have magnetic dipoles, there should be an dipole-dipole interaction between them, which was ignored). In addition, we also ignored the earth magnetic field, which is always in presence when we do an experiment (unless we screen it out using some special devices.)

Let’s use \( H_{\text{real}} \) to represent the full Hamiltonian of a real system and \( H_{\text{model}} \) to represent the Hamiltonian that we used to theoretically analyze the system. We know that these two are not the same, because we always need some approximation to simplify a real problem, i.e.

\[ H_{\text{real}} = H_{\text{model}} + \delta H \]  
(2.141)

we can treat \( \delta H \) as a perturbation.

Now here comes the questions, if we found that two (or more) states have the same eigenenergy (degeneracy) using \( H_{\text{model}} \), are these states really degenerate in a real system? The general answer is no (unless there is a reason), because our first order degenerate perturbation theory told us that any small perturbation will in general lift the degeneracy.

The only except is: if there is a reason (usually based on symmetry) to tell us that \( W_{aa} \) is exactly the same as \( W_{bb} \) and \( W_{ab} = W_{ba} = 0 \) precisely. (in a real physics system, in most of the case, we cannot say that the value of a quantity is precisely this number. What we really mean is that there is an argument to show that the difference between \( W_{aa} \) and \( W_{bb} \) is unmeasurably small and \( W_{ab} \) and \( W_{ba} \) is unmeasurably small).

2.4.6. Prove: \( \langle \psi_1^0 \mid H' \mid \psi_2^0 \rangle = 0 \)

Here the proof contains two steps: first

\[ \langle \psi_1^0 \mid H' \mid \psi_2^0 \rangle = (\alpha_1^* \beta_1^*) \left( \begin{array}{cc} W_{aa} & W_{ab} \\ W_{ba} & W_{bb} \end{array} \right) \left( \begin{array}{c} \alpha_2 \\ \beta_2 \end{array} \right) \]  
(2.142)

and then we will show

\[ (\alpha_1^* \beta_1^*) \left( \begin{array}{cc} W_{aa} & W_{ab} \\ W_{ba} & W_{bb} \end{array} \right) \left( \begin{array}{c} \alpha_2 \\ \beta_2 \end{array} \right) = 0 \]  
(2.143)

The first step is very straightforward (it is from the definition of \( | \psi_1^0 \rangle \) and \( | \psi_2^0 \rangle \))

\[ | \psi_1^0 \rangle = \alpha_1 \mid \phi_a^0 \rangle + \beta_1 \mid \phi_b^0 \rangle \]  
(2.144)

\[ | \psi_2^0 \rangle = \alpha_2 \mid \phi_a^0 \rangle + \beta_2 \mid \phi_b^0 \rangle \]  
(2.145)

so
\((\psi_0^1 \mid H' \mid \psi_0^1) = ((\psi_0^0 \mid \alpha_1^* + (\psi_0^0 \mid \beta_1^*)H' (\alpha_2^* + (\psi_0^0 \mid \beta_2^*) + (\psi_0^0 \mid \beta_1^*) (\alpha_2^* H' \mid \psi_0^0) + (\beta_2^*) H' \mid \psi_0^0) = \\
\alpha_1^* \alpha_2^* (\psi_0^0 \mid H' \mid \psi_0^0) + (\alpha_1^* \beta_2^*) (\psi_0^0 \mid H' \mid \psi_0^0) + (\beta_1^* \alpha_2^*) (\psi_0^0 \mid H' \mid \psi_0^0) + (\beta_1^* \beta_2^*) (\psi_0^0 \mid H' \mid \psi_0^0)
\)  

(2.146)

The r.h.s. of the equation is in fact exactly the same, if we remember the definition of the \(W\) matrix \(W_{ij} = (\psi_i^0 \mid H' \mid \psi_j^0)\)

\((\alpha_1^* \beta_1^*) (W_{aa} W_{ab}) (\alpha_2^* \beta_2^*) = (\alpha_1^* \alpha_2^* W_{aa} + (\alpha_1^* \beta_2^*) W_{ab} + (\beta_1^* \alpha_2^*) W_{ba} + (\beta_1^* \beta_2^*) W_{bb})
\)

(2.147)

So, we proved that \((\psi_0^0 \mid H' \mid \psi_0^0) = (\alpha_1^* \beta_1^*) (W_{aa} W_{ab}) (\alpha_2^* \beta_2^*)
\)

(2.149)

Similarly, if we start from the other eigenquation

\((W_{aa} W_{ab}) (\alpha_1^* \beta_1^*) = E (\alpha_1^* \beta_1^*)
\)

(2.150)

we get

\((\alpha_2^* \beta_2^*) (W_{aa} W_{ab}) (\alpha_1^* \beta_1^*) = (\alpha_2^* \beta_2^*) E (\alpha_1^* \beta_1^*)
\)

(2.151)

If we take a conjugate on both sides

\((\alpha_1^* \beta_1^*) (W_{aa} W_{ab}) (\alpha_2^* \beta_2^*) = E (\alpha_1^* \beta_1^*) (\alpha_2^* \beta_2^*)
\)

(2.152)

here we used the fact that \(W\) is Hermitian and thus \(W^* = W\). Notice that we have shown

\((\alpha_1^* \beta_1^*) (W_{aa} W_{ab}) (\alpha_2^* \beta_2^*) = E (\alpha_1^* \beta_1^*) (\alpha_2^* \beta_2^*)
\)

(2.153)

and

\((\alpha_1^* \beta_1^*) (W_{aa} W_{ab}) (\alpha_2^* \beta_2^*) = E (\alpha_1^* \beta_1^*) (\alpha_2^* \beta_2^*)
\)

(2.154)

If \(E_+ \neq E_-\), the only way that these two equations can both be valid is that

\((\alpha_1^* \beta_1^*) (W_{aa} W_{ab}) (\alpha_2^* \beta_2^*) = E (\alpha_1^* \beta_1^*) (\alpha_2^* \beta_2^*)
\)

(2.155)

So we proved that \((\psi_0^0 \mid H' \mid \psi_0^0) = 0.
\)

Q: what will happen if \(E_+ = E_-\)?

A: Turns out that this is the simple case. If \(E_+ = E_-\), as will be shown below, \(W_{ab} = (\psi_a^0 \mid H' \mid \psi_b^0) = 0\). So, there is no divergence from the beginning. We can start the perturbation theory without worrying about these divergences.

### 2.4.7 First order perturbation

The calculation described above provides to us the zeroth order wavefunctions (i.e., we should use \(|\psi_0^0\rangle\) or \(|\psi_2^0\rangle\), instead of \(|\psi_0^{\alpha}\rangle\) or \(|\psi_0^{\beta}\rangle\) as our unperturbed wavefunction). As we learned early on (in non-degenerate perturbation theory), the first order correction of energy is just

\(E_0^1 = (\psi_0^0 \mid H' \mid \psi_0^0)
\)

(2.156)

i.e., we use the zeroth order wavefunction and compute the expectation value for \(H'\). Here, for the zeroth order wavefunctions, we have two of them, \(|\psi_0^\alpha\rangle\) and \(|\psi_0^\beta\rangle\), so we need to compute the first order energy correction for each of them. And we will prove in this section...
In this part, we review basic ideas of eigenvalues and eigenvectors. We starts from the eigenequation defined in the previous section

\[ E_1 = \langle \phi_0^i | H^* | \phi_0^j \rangle = E \]  

(2.157)

and

\[ E_2 = \langle \phi_0^o | H^* | \phi_0^o \rangle = E \]  

(2.158)

i.e., the first order energy corrections for | \psi_1^2 \rangle and | \psi_2^2 \rangle are precisely the two eigenvalues of the \( W \) matrix.

\[ \langle \phi_0^i | H^* | \phi_0^i \rangle = \left( \langle \phi_0^i | | \alpha_1^* + \langle \phi_0^i | \beta_1^* \rangle \right) H^* \left( \langle \phi_0^i | \beta_1^* \right) \right) = \left( \langle \phi_0^0 | | \alpha_1^* + \langle \phi_0^0 | \beta_1^* \rangle \right) \left( \langle \phi_0^0 | H^* | \phi_0^0 \rangle + \beta_1^* \right) = \alpha_1^* \alpha_1 \left( \phi_0^0 | H^* | \phi_0^0 \right) + \beta_1^* \alpha_1 \left( \phi_0^0 | H^* | \phi_0^0 \right) + \beta_1^* \alpha_1 \left( \phi_0^0 | H^* | \phi_0^0 \right)

(2.159)

If we remember the definition of the \( W \) matrix \( W_{ij} = \langle \psi_i^0 | H^* | \psi_j^0 \rangle \), we realized immediately that this formula is exactly the same as

\[ \langle \alpha_1^* \beta_1^* \left( W_{aa} W_{ab} \left( \alpha_1 \beta_1 \right) = \alpha_1^* \beta_1^* \left( W_{aa} W_{ab} \right) \left( \alpha_1 \beta_1 \right) \right)

(2.160)

So we found

\[ E_1 = \langle \phi_0^i | H^* | \phi_0^i \rangle = \left( \alpha_1^* \beta_1^* \right) \left( W_{aa} W_{ab} \right) \left( \alpha_1 \beta_1 \right)

(2.161)

Because \( \left( \frac{\alpha_1}{\beta_1} \right) \) is an eigenvector of \( B \)

\[ \left( W_{aa} W_{ab} \right) \left( \frac{\alpha_1}{\beta_1} \right) = E_1 \left( \frac{\alpha_1}{\beta_1} \right) \]  

(2.162)

\[ E_1 = \langle \phi_0^i | H^* | \phi_0^i \rangle = \left( \alpha_1^* \beta_1^* \right) \left( W_{aa} W_{ab} \right) \left( \frac{\alpha_1}{\beta_1} \right) = E_1 \left( \frac{\alpha_1^* \beta_1^*}{\beta_1^*} \right) \left( \frac{\alpha_1}{\beta_1} \right)

(2.163)

Because we have required \( \left( \frac{\alpha_1}{\beta_1} \right) \) to be normalized, \( \left( \alpha_1^* \beta_1^* \right) \left( \frac{\alpha_1}{\beta_1} \right) = \alpha_1^* \alpha_1 + \beta_1^* \beta_1 = 1 \)

\[ E_1 = \langle \phi_0^i | H^* | \phi_0^i \rangle = \left( \alpha_1^* \beta_1^* \right) \left( W_{aa} W_{ab} \right) \left( \frac{\alpha_1}{\beta_1} \right) = E_1 \left( \frac{\alpha_1^* \beta_1^*}{\beta_1^*} \right) \left( \frac{\alpha_1}{\beta_1} \right) = E

(2.164)

Similarly, we can show that

\[ E_2 = \langle \phi_0^o | H^* | \phi_0^o \rangle = \left( \alpha_2^* \beta_2^* \right) \left( W_{aa} W_{ab} \right) \left( \frac{\alpha_2}{\beta_2} \right) = E_2 \left( \frac{\alpha_2^* \beta_2^*}{\beta_2^*} \right) \left( \frac{\alpha_2}{\beta_2} \right) = E

(2.165)

2.4.8. Eigenvalues of the matrix \( W \)

In this part, we review basic ideas of eigenvalues and eigenvectors. We starts from the eigenequation defined in the previous section

\[ \left( W_{aa} W_{ab} \right) \left( \frac{\alpha}{\beta} \right) = E \left( \frac{\alpha}{\beta} \right)

(2.166)

This means that

\[ W_{aa} \alpha + W_{ab} \beta = E \alpha \]  

(2.167)

\[ W_{ba} \alpha + W_{bb} \beta = E \beta \]  

(2.168)

These two equations have an obvious and trivial solution \( \alpha = \beta = 0 \). This solution is NOT what we want and we will not consider this trivial solution. To get a nontrivial solution, the eigenvalue \( E \) cannot be an arbitrary value. It can only be one of two values, as will be seeing below.

Using the first equation, we get

\[ \alpha = \frac{W_{ab}}{E - W_{aa}} \beta

(2.169)

Using the second equation, we get

\[ \alpha = \frac{E - W_{ab}}{W_{ba}} \beta

(2.170)

The first relation means
\[
\frac{\alpha}{\beta} = \frac{W_{ab}}{E - W_{aa}}
\]

but the second relation requires
\[
\frac{\alpha}{\beta} = \frac{E - W_{bb}}{W_{ba}}
\]

So we have
\[
\frac{\alpha}{\beta} = \frac{W_{ab}}{E - W_{aa}} = \frac{E - W_{bb}}{W_{ba}}
\]

In general, \( \frac{W_{ab}}{E - W_{aa}} \neq \frac{E - W_{bb}}{W_{ba}} \), so we find an contradiction. This contradiction means that for a general value of \( E \), we will only have the trivial solution \( \alpha = \beta = 0 \). To get a nontrivial solution, we have to request \( \frac{W_{ab}}{E - W_{aa}} = \frac{E - W_{bb}}{W_{ba}} \). This equation is often written in a different form

\[
\frac{W_{ab}}{E - W_{aa}} = \frac{E - W_{bb}}{W_{ba}}
\]

or equivalently
\[
\det \left[ \begin{array}{cc}
E & -W_{ab} \\
-W_{ba} & E - W_{bb}
\end{array} \right] = 0
\]

or equivalently
\[
\det \left[ \begin{array}{cc}
E & -E \\
0 & E
\end{array} \right] - \left[ \begin{array}{cc}
W_{aa} & W_{ab} \\
W_{ba} & W_{bb}
\end{array} \right] = 0
\]

or equivalently
\[
\det \left( E - W \right) = 0
\]

Here, \( W \) is the matrix that we define above
\[
W = \left[ \begin{array}{cc}
W_{aa} & W_{ab} \\
W_{ba} & W_{bb}
\end{array} \right]
\]

and number \( E \) here means \( E \) times the identity matrix
\[
E \left( \begin{array}{cc}
1 & 0 \\
0 & 1
\end{array} \right) = \left( \begin{array}{cc}
E & 0 \\
0 & E
\end{array} \right)
\]

\[
\det \left( E - W \right) = 0 \text{ means}
\]

\[
(W_{aa} - E)(W_{bb} - E) - W_{ab}W_{ba} = 0
\]

And thus
\[
E^2 - (W_{aa} + W_{bb})E + (W_{aa}W_{bb} - W_{ab}W_{ba}) = 0
\]

By definition, \( \text{tr} \ W = W_{aa} + W_{bb} \) and \( \det \ W = W_{aa}W_{bb} - W_{ab}W_{ba} \). Therefore, we can write the same equation as
\[
E^2 - \text{tr} \ W E + \det \ W = 0
\]

This equation has two solutions
\[
E_A = \frac{\text{tr} \ W \pm \sqrt{(\text{tr} \ W)^2 - 4 \det \ W}}{2}
\]

As shown above, these two solutions, \( E_A \) are the first order correction to the eigenenergy. In the perturbation theory, the eigenenergies of these two quantum states are
\[
E = E^0 + E_A \lambda + O(\lambda^2)
\]
\[ E = E^0 + E_\lambda + O(\lambda^2) \]  
\[ \text{(2.187)} \]

at small \( \lambda \).

**Comment #1.** \( \text{tr} W \) and \( \det W \) are both real. This is straightforward to prove, if we notice that \( W \) is Hermitian. Because \( W_{aa} = W_{aa}^* \) and \( W_{bb} = W_{bb}^* \), and \( W_{ba} = W_{ab}^\dagger \). \( W_{aa} \) and \( W_{bb} \) are real. And \( W_{ab} W_{ba} = |W_{ab}|^2 \) is also real, so \( \text{tr} W = W_{aa} + W_{bb} \) and \( \det W = W_{aa} W_{bb} - W_{ab} W_{ba} \) are both real.

**Comment #2.** \( (\text{tr} W)^2 - 4 \det W \geq 0 \). Therefore, \( E_\lambda \) are both real.

\[ (\text{tr} W)^2 - 4 \det W = (W_{aa} + W_{bb})^2 - 4 W_{aa} W_{bb} + 4 W_{ab} W_{ba} = (W_{aa} - W_{bb})^2 + 4 |W_{ab}|^2 \geq 0 \]  
\[ \text{(2.188)} \]

here we used the fact that \( \bar{W}_{ab} = W_{ba} \).

**Comment #3.** There are in general two possible situations (a) If \( (\text{tr} W)^2 - 4 \det W > 0 \), \( E_\lambda > E_- \), i.e. the two eigenvalues are NOT the same. (b) If \( W_{aa} = W_{bb} \) and \( W_{ab} = W_{ba} = 0 \), \( (\text{tr} W)^2 - 4 \det W = 0 \), and thus \( E_+ = E_- = \text{tr} W/2 \).

The situation (b) is the easy case, because \( W_{ab} = W_{ba} = 0 \) means \( \langle \psi_0^\dagger | H' | \psi_0^\dagger \rangle = 0 \). Remember that the problem we had from the beginning is that the second order correction will diverge, \( \langle \psi_0^\dagger | H' | \psi_0^\dagger \rangle \), because \( E_0^0 = E_0^0 \). For situation (b), the numerator is zero, so there is no divergence. And thus we can just do non-degenerate perturbation theory. The situation (a) is the more generic case. There, as we have shown early on, the perturbation \( H' \) lift the degeneracy.

### 2.4.9. Eigenvectors of the matrix \( W \)

In this section, we will assume that \( (\text{tr} W)^2 - 4 \det W > 0 \), i.e. situation (a) discussed in the previous section. We have two eigenvalues. For each eigenvalue, we can solve for the corresponding eigenvector. For the eigenvalue \( E_+ = \frac{\text{tr} W + \sqrt{(\text{tr} W)^2 - 4 \det W}}{2} \), we have

\[ \begin{pmatrix} W_{ab} \\ W_{ba} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} = E_+ \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} \]  
\[ \text{(2.189)} \]

and for the other eigenvalue \( E_- = \frac{\text{tr} W - \sqrt{(\text{tr} W)^2 - 4 \det W}}{2} \), we have

\[ \begin{pmatrix} W_{ab} \\ W_{ba} \end{pmatrix} \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix} = E_- \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix} \]  
\[ \text{(2.190)} \]

We will use the first one as example (equation for \( E_+ \)). There, the matrix equation can be written as two separate equations

\[ W_{aa} \alpha_1 + W_{ab} \beta_1 = E_+ \alpha_1 \]  
\[ \text{(2.191)} \]

\[ W_{ba} \alpha_1 + W_{bb} \beta_1 = E_+ \beta_1 \]  
\[ \text{(2.192)} \]

Using the first equation, we get

\[ \alpha_1 = \frac{W_{ab}}{E_+ - W_{aa}} \beta_1 \]  
\[ \text{(2.193)} \]

Using the second equation, we get

\[ \alpha_1 = \frac{E_+ - W_{bb}}{W_{ba}} \beta_1 \]  
\[ \text{(2.194)} \]

These two relations are actually identical, because for any eigenvalue \( E \), we have \( \frac{W_{ab}}{E - W_{aa}} = \frac{E - W_{bb}}{W_{ba}} \) as we proved early on.

In addition, we know that \( \alpha_1^* \alpha_1 + \beta_1^* \beta_1 = 1 \), i.e. the normalization condition. So we have

\[ \alpha_1 = \frac{W_{ab}}{\sqrt{|W_{ab}|^2 + (E_+ - W_{aa})^2}} \]  
\[ \text{(2.195)} \]

\[ \beta_1 = \frac{E_+ - W_{aa}}{\sqrt{|W_{ab}|^2 + (E_+ - W_{ab})^2}} \]  
\[ \text{(2.196)} \]

Similarly, we have
\[
\alpha_2 = \frac{W_{ab}}{\sqrt{|W_{ab}|^2 + (E_+ - W_{aa})^2}} \\
\beta_2 = \frac{E_- - W_{aa}}{\sqrt{|W_{ab}|^2 + (E_- - W_{ab})^2}} \tag{2.197}
\]

In conclusion, we found that

\[
\psi_1 = \alpha \psi_a + \beta \psi_b = \frac{W_{ab}}{\sqrt{|W_{ab}|^2 + (E_+ - W_{aa})^2}} \psi_a + \frac{E_+ - W_{aa}}{\sqrt{|W_{ab}|^2 + (E_+ - W_{ab})^2}} \psi_b \tag{2.198}
\]

\[
|\psi_0\rangle = \alpha_1 |\psi_0\rangle + \beta_1 |\psi_0\rangle = \frac{W_{ab}}{\sqrt{|W_{ab}|^2 + (E_+ - W_{aa})^2}} |\psi_0\rangle + \frac{E_+ - W_{aa}}{\sqrt{|W_{ab}|^2 + (E_+ - W_{ab})^2}} |\psi_0\rangle \tag{2.199}
\]

\[
|\psi_2\rangle = \alpha_2 |\psi_0\rangle + \beta_2 |\psi_0\rangle = \frac{W_{ab}}{\sqrt{|W_{ab}|^2 + (E_- - W_{aa})^2}} |\psi_0\rangle + \frac{E_- - W_{aa}}{\sqrt{|W_{ab}|^2 + (E_- - W_{ab})^2}} |\psi_0\rangle \tag{2.200}
\]

\section*{2.4.10. The very special case}

In general, we expect \(E_+ \neq E_-\), i.e. the generacy is lifted. What will happen if \(E_+ = E_-\)? From the equation

\[
E_+ = \frac{\text{tr} W \pm \sqrt{\text{tr} W^2 - 4 \text{det} W}}{2} \tag{2.201}
\]

we know that \(E_+ = E_-\) can only arise when \(\sqrt{\text{tr} W^2 - 4 \text{det} W} = 0\), i.e. \(\text{tr} W^2 - 4 \text{det} W = 0\). As shown above

\[
\text{tr} W^2 - 4 \text{det} W = (W_{aa} - W_{bb})^2 + 4 |W_{ab}|^2 \tag{2.202}
\]

Both the two terms on the r.h.s. are non-negative, and thus if we want the whole thing to be zero, we must have

\[
(W_{aa} - W_{bb})^2 = 0 \tag{2.203}
\]

and

\[
4 |W_{ab}|^2 = 0 \tag{2.204}
\]

i.e., \(W_{aa} = W_{bb}\) and \(W_{ab} = 0\).

With \(W_{aa} = W_{bb}\) and \(W_{ab} = 0\), \(W\) is actually proportional to an identity matrix.

\[
W = \begin{pmatrix} W_{aa} & 0 \\ 0 & W_{aa} \end{pmatrix} = W_{aa} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.205}
\]

This situation is highly unlikely to arise (unless there is a reason) because, in general, the \(W\) matrix has four free values to pick \(W_{aa}\), \(W_{bb}\), the real part of \(W_{ab}\) and the imaginary part of \(W_{ab}\) (note 1: \(W_{aa}\) and \(W_{bb}\) are real, so they don’t have imaginary part. note 2: \(W_{aa}\) is the complex conjugate of \(W_{ab}\), so we don’t need to consider it here as a separate degree of freedom). If you have four real values, what is the probability for these four real values to satisfy that \(W_{aa}\) matches exactly \(W_{bb}\) without any error bar, and both the real and imaginary parts of \(W_{ab}\) vanishes exactly without any error bar? Without a reason, the chance is zero. So this is a situation that we don’t need to worry much, unless there is a reason.

In most cases, such a special case arises due to symmetry. For example, time reversal symmetry tells us that there should be two degenerate states (a state and its time reversal state). Then, for \(H_0\) these two states degenerates and for \(H\), they should still be degenerate, so \(E_+ = E_-\). For that situation, it turns out that one can directly start from non-degenerate perturbation theory (no singularities will arise), although the states are degenerate. We will discuss a more generic situation later, which covers this case.

\section*{2.4.11. Review: Quantum states and quantum operators as matrices}

Once we choose a set of basis, any quantum state can be written as a vector (i.e., a N-by-1 matrix).

For a complete set of basis, \(\{|\psi_i\rangle\}\), we can write any quantum states as

\[
|\psi\rangle = \sum c_i |\psi_i\rangle \tag{2.206}
\]
where $c_i$ are complex numbers. Here, we find that if we want to describe a state, we just need to know all the coefficient $c_i$. We can write these $c_i$ as a vector

$$
\begin{pmatrix}
  c_1 \\
  c_2 \\
  c_3 \\
  \vdots
\end{pmatrix}
$$

These coefficients are

$$
c_i = \langle \psi_i | \psi \rangle
$$

To see this, we multiply $\langle \psi_j |$ for both sides of $|\psi\rangle = \sum c_i |\psi_i\rangle$

$$
\langle \psi_j | \psi \rangle = \sum c_i \langle \psi_j | \psi_i \rangle = \sum c_i \delta_{ij} = c_j
$$

**Bottom line: a quantum state is a column vector**

$$
|\psi\rangle \rightarrow \begin{pmatrix}
  c_1 \\
  c_2 \\
  c_3 \\
  \vdots
\end{pmatrix}
$$

**Conjugate states is the represented by the conjugate vector.** By definition, we know that

$$
\langle \psi \rangle = \sum c_i |\psi\rangle
$$

so, we can write all these $c_i^*$ as a row vector

$$
( c_1^* \ c_2^* \ c_3^* \ ... ) = ( \langle \psi_1 | \psi \rangle \ \langle \psi_2 | \psi \rangle \ \langle \psi_3 | \psi \rangle \ \ldots )
$$

Here, we used the fact that $\langle \psi | \psi_i \rangle$ is the complex conjugate of $\langle \psi_i | \psi \rangle$

**Inner produce of two states are product of a row vector and a column vector**

If we know two quantum states

$$
|\psi\rangle \rightarrow \begin{pmatrix}
  c_1 \\
  c_2 \\
  c_3 \\
  \vdots
\end{pmatrix}
$$

and

$$
|\phi\rangle \rightarrow \begin{pmatrix}
  d_1 \\
  d_2 \\
  d_3 \\
  \vdots
\end{pmatrix}
$$

then, we know

$$
\langle \phi | \rightarrow ( d_1^* \ d_2^* \ d_3^* \ \ldots )
$$

so

$$
\langle \phi | \psi \rangle \rightarrow ( d_1^* d_2^* d_3^* \ \ldots ) \begin{pmatrix}
  c_1 \\
  c_2 \\
  c_3 \\
  \vdots
\end{pmatrix} = d_1^* c_1 + d_2^* c_2 + d_3^* c_3 + \ldots
$$

Q: How about a quantum operator?

A: Once we choose a set of basis, a quantum operator is a matrix.

To understand this, we just need to realize that a quantum operator transforms a quantum state into a different state

$$
\hat{X} |\psi\rangle = |\phi\rangle
$$

As we have known, $|\psi\rangle$ is a column vector, and $|\phi\rangle$ is another column vector. Which object transfers a column vector to a different column
vector? We know that a matrix can do such a job

\[
\begin{pmatrix}
  x_{11} & x_{12} & x_{13} & \cdots & c_1 \\
  x_{21} & x_{22} & x_{23} & \cdots & c_2 \\
  x_{31} & x_{32} & x_{33} & \cdots & c_3 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
\end{pmatrix}
= \begin{pmatrix}
  d_1 \\
  d_2 \\
  d_3 \\
  \vdots \\
\end{pmatrix}
\]

(2.218)

So a quantum operator is really similar to a matrix. In fact, the matrix elements \(x_{ij}\) are very easy to compute

\[x_{ij} = \langle \psi_i | \hat{X} | \psi_j \rangle\]  

(2.219)

Q: How about eigenvalues and eigenstates?

A: Matrices also have eigenvalues and eigenstates

\[
\begin{pmatrix}
  x_{11} & x_{12} & x_{13} & \cdots & c_1 \\
  x_{21} & x_{22} & x_{23} & \cdots & c_2 \\
  x_{31} & x_{32} & x_{33} & \cdots & c_3 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
\end{pmatrix}
= W
\begin{pmatrix}
  c_1 \\
  c_2 \\
  c_3 \\
  \vdots \\
\end{pmatrix}
\]

(2.220)

1. The matrix of a Hermitian operator is a Hermitian matrix

2. An \(N \times N\) Hermitian matrix has \(N\) eigenvalues, each of which has an eigenvector

3. Eigenvalues of the matrix is the same as the eigenvalues of the corresponding quantum operator

4. Each eigenvector corresponds to an eigenstate, i.e. If \(\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix}\) is an eigenvector with eigenvalue \(W\), then \(|\phi\rangle = \sum c_i |\psi_i\rangle\) is an eigenstate of \(\hat{X}\) with eigenvalue \(W\).

**Final conclusion:** for a quantum system, we just need to play with matrices

**Only one problem:** these matrices are huge (\(\infty \times \infty\))

It is extremely hard to handle big matrices (say 100 million by 100 million). So this approach doesn’t make our life easier.

### 2.4.12. Degenerate perturbation theory

\[H = H_0 + \lambda H'\]

(2.221)

Using eigenstates of \(H_0\) as basis, then \(H_0\) corresponds to a diagonal matrix

\[H_0 |\psi_i\rangle = E_i^0 |\psi_i\rangle\]

(2.222)

where \(i = 1, 2, 3, \ldots\) and we request this is an orthonormal basis

\[\langle \psi_i | \psi_j \rangle = \delta_{ij}\]

(2.223)

A matrix element of the matrix is

\[\langle \psi_i | H_0 | \psi_j \rangle = \langle \psi_i | E_i^0 | \psi_j \rangle = E_i^0 \langle \psi_i | \psi_j \rangle = E_i^0 \delta_{ij}\]

(2.224)

So,

\[H_0 \rightarrow \begin{pmatrix}
  E_1^0 & 0 & 0 & \cdots \\
  0 & E_2^0 & 0 & \cdots \\
  0 & 0 & E_3^0 & \cdots \\
  \vdots & \vdots & \vdots & \ddots \\
\end{pmatrix}\]

(2.225)

This conclusion is generically true. If we use the eigenstates of an operator as our basis, then this operator is a diagonal matrix (i.e. off-diagonal terms are all zero). And along the diagonal line, we just have all the eigenvalues of this quantum operator.

\[\lambda H' \rightarrow \lambda \begin{pmatrix}
  \langle \psi_1 | H' | \psi_1 \rangle & \langle \psi_1 | H' | \psi_2 \rangle & \langle \psi_1 | H' | \psi_3 \rangle & \cdots \\
  \langle \psi_2 | H' | \psi_1 \rangle & \langle \psi_2 | H' | \psi_2 \rangle & \langle \psi_2 | H' | \psi_3 \rangle & \cdots \\
  \langle \psi_3 | H' | \psi_1 \rangle & \langle \psi_3 | H' | \psi_2 \rangle & \langle \psi_3 | H' | \psi_3 \rangle & \cdots \\
  \vdots & \vdots & \vdots & \ddots \\
\end{pmatrix}\]

(2.226)
In general, $H'$ is NOT an diagonal matrix

$$H = H_0 + \lambda H' \rightarrow \begin{pmatrix} E_0 & 0 & 0 & \ldots \\ 0 & E_2 & 0 & \ldots \\ 0 & 0 & E_3 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} + \lambda \begin{pmatrix} \langle \phi_1 | H' | \phi_1 \rangle & \langle \phi_1 | H' | \phi_2 \rangle & \langle \phi_1 | H' | \phi_3 \rangle & \ldots \\ \langle \phi_2 | H' | \phi_1 \rangle & \langle \phi_2 | H' | \phi_2 \rangle & \langle \phi_2 | H' | \phi_3 \rangle & \ldots \\ \langle \phi_3 | H' | \phi_1 \rangle & \langle \phi_3 | H' | \phi_2 \rangle & \langle \phi_3 | H' | \phi_3 \rangle & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} =$$

(2.227)

$$\begin{pmatrix} E_0 + \lambda \langle \phi_1 | H' | \phi_1 \rangle & \lambda \langle \phi_1 | H' | \phi_2 \rangle & \lambda \langle \phi_1 | H' | \phi_3 \rangle & \ldots \\ \lambda \langle \phi_2 | H' | \phi_1 \rangle & E_2 + \lambda \langle \phi_2 | H' | \phi_2 \rangle & \lambda \langle \phi_2 | H' | \phi_3 \rangle & \ldots \\ \lambda \langle \phi_3 | H' | \phi_1 \rangle & \lambda \langle \phi_3 | H' | \phi_2 \rangle & E_3 + \lambda \langle \phi_3 | H' | \phi_3 \rangle & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

This is a very large matrix and is very hard to handle in general.

However, if there are 2 degenerate states,

$$H_0 = \begin{pmatrix} \ddots & \ddots & \vdots & \vdots \\ \vdots & E_0 & 0 & \cdots \\ \vdots & 0 & E_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

(2.228)

i.e., two of the eigenvalues of $H_0$ coincides, or say two of two numbers along the diagonal line of the matrix of $H_0$ happens to be the same, the we don’t need to handle the whole big matrix, if $\lambda$ is small. Here, we can do degenerate perturbation theory, and to the first order, we can forget all other quantum states and only look at the two degenerate ones. What does this mean? Remember, that in general, a set of complete basis contains infinite number of states $|\psi_i\rangle$ with $i = 1, 2, \ldots, \infty$. As a result, the matrix of a quantum operator has dimension $\infty \times \infty$, $\langle \psi_i | \hat{X} | \psi_j \rangle$ with $i = 1, 2, \ldots, \infty$ and $j = 1, 2, \ldots, \infty$. Now, if we only limit ourselves to the two degenerate states $|\psi_a^0\rangle$ and $|\psi_b^0\rangle$, then the matrix of our quantum operator only has dimensions $2 \times 2$, because my $i$ and $j$ here can only be $a$ or $b$

$$\hat{X} \rightarrow \begin{pmatrix} \langle \psi_a^0 | \hat{X} | \psi_a^0 \rangle & \langle \psi_b^0 | \hat{X} | \psi_a^0 \rangle \\ \langle \psi_a^0 | \hat{X} | \psi_b^0 \rangle & \langle \psi_b^0 | \hat{X} | \psi_b^0 \rangle \end{pmatrix}$$

(2.229)

For $H_0$, its matrix is

$$H_0 \rightarrow \begin{pmatrix} E_0 & 0 \\ 0 & E_0 \end{pmatrix}$$

(2.230)

and for $H'$, the matrix is

$$H' \rightarrow \begin{pmatrix} \langle \psi_a^0 | H' | \psi_a^0 \rangle & \langle \psi_a^0 | H' | \psi_b^0 \rangle \\ \langle \psi_b^0 | H' | \psi_a^0 \rangle & \langle \psi_b^0 | H' | \psi_b^0 \rangle \end{pmatrix}$$

(2.231)

So our $H$ is

$$H = H_0 + \lambda H' \rightarrow \begin{pmatrix} E_0 + \lambda \langle \psi_a^0 | H' | \psi_a^0 \rangle & \lambda \langle \psi_a^0 | H' | \psi_b^0 \rangle \\ \lambda \langle \psi_b^0 | H' | \psi_a^0 \rangle & E_0 + \lambda \langle \psi_b^0 | H' | \psi_b^0 \rangle \end{pmatrix}$$

(2.232)

The eigenvalues of this matrix are $E_0 + \lambda E_a$ and $E_0 + \lambda E_b$. And the eigenvector is the same as we computed in previous section

**Bottom line:** for degenerate perturbation, we can drop all other states (with different eigenenergies), and consider a much smaller Hilbert space (only the degenerate states are considered here). Then, our Hamiltonian becomes a very small matrix, and we can diagonalize this small matrix. The eigenvalues are the eigenenergies to the first order. And the eigenvectors give us eigenwavefunctions to zeroth order.

### 2.4.13. n-fold degeneracy

If $H_0$ as n-fold degeneracy, and we want to do perturbation theory for these n degenerate states, we just ignore all other states and only keep these n states.

$$H_0 \rightarrow \begin{pmatrix} E_0 & 0 & 0 & \ldots \\ 0 & E_0 & 0 & \ldots \\ 0 & 0 & E_0 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}_{nn}$$

(2.233)
and $H'$ is a $n \times n$ matrix with matrix elements $\langle \psi_j | H' | \psi_i \rangle$, where $i = 1, \ldots, n$ and $j = 1, \ldots, n$

Then, there are two (equivalent) ways to do the calculation

Option #1: compute eigenvalues for the $n \times n$ matrix of $H'$, as $E_1 \ldots E_n$. Then the eigenenergy to the first order correction is

$$E_i = E_0 + \lambda E_i^1 + O(\lambda^2)$$

where $i = 1, 2, \ldots, n$.

Option #2: direction compute eigenvalues of the $n \times n$ matrix of $H = H_0 + \lambda H'$. You will find $n$ eigenvalue, they are

$$E_0 + \lambda E_i^1$$

where $i = 1, 2, \ldots, n$.

### 2.4.14. Nearly-degenerate perturbation theory

What if we have two states that are not totally degenerate, but nearly degenerate, i.e. two eigenstates of $H_0 | \psi_0 \rangle$ and $| \psi_0 \rangle$ has very similar energies ($E_0 - E_0^0$), but not exactly the same.

Case 1. if $\lambda H' << | E_0^0 - E_0 \rangle$, we can do non-degenerate perturbation theory

Case 2. if $\lambda H' > | E_0^0 - E_0 \rangle$ and $H' << | E_0^0 - E_0 \rangle$ for any other eigenstates of $H_0$, where $E_0^0$ represent eigenenergy of another eigenstate of $H_0$ (beyond $| \psi_0 \rangle$ and $| \psi_0 \rangle$), we can do nearly-degenerate perturbation theory.

Here, the procedure is similar to the degenerate perturbation theory, we ignore all other states and only consider $| \psi_0 \rangle$ and $| \psi_0 \rangle$. Now, every quantum operator becomes a $2 \times 2$ matrix.

$$H_0 \rightarrow \begin{pmatrix} E_0^0 & 0 \\ 0 & E_0^0 \end{pmatrix}$$

and for $H'$, the matrix is

$$H' \rightarrow \begin{pmatrix} \langle \psi_0 | H' | \psi_0 \rangle & \langle \psi_0 | H' | \psi_0 \rangle \\ \langle \psi_0 | H' | \psi_0 \rangle & \langle \psi_0 | H' | \psi_0 \rangle \end{pmatrix} = \begin{pmatrix} V_{aa} & V_{ab} \\ V_{ba} & V_{bb} \end{pmatrix}$$

Notice that $H_0$ has two different diagonal components $E_0^0 \neq E_0^0$. This is the difference between degenerate and nearly-degenerate perturbation theory. Now we consider $H', \lambda H'$, which is

$$H = H_0 + \lambda H' \rightarrow \begin{pmatrix} E_0^0 + \lambda V_{aa} & \lambda V_{ab} \\ \lambda V_{ba} & E_0^0 + \lambda V_{bb} \end{pmatrix}$$

Then, we can get eigenvalues of this matrix, and these are the eigenvalues of $H$ up to first order in perturbation theory.

Case 3. if $\lambda H' >$ too large, even larger than $| E_0^0 - E_0^0 |$ and $| E_0^0 - E_0^0 |$, then $\lambda H'$ is too large and thus cannot be considered as a perturbation and as a result, we cannot do perturbation theory anymore.

This result can be easily generalized to cases where we have more than 2 nearly-degenerate states.

### 2.4.15. Philosophy behind degenerate and nearly-degenerate perturbation theory

Assuming that $H_0$ have $n$ eigenstates who have very similar (or exactly the same) eigenenergies ($| \psi_0 \rangle, | \psi_0 \rangle \ldots$ and they all have energy near $E_0$), but all other eigenstates of $H_0$ have energies very different from these states ($| \psi_m \rangle$ has eigenenergy $E_m^0$ and $m$ is not one of the nearly degenerate states. For any $m$, we have $E_m^0$ very different from $E_0$). Then when we start our system from one (or some superposition) of these $n$ states, and then perturb the Hamiltonian by a small amount $H = H_0 + \lambda H'$ with $\lambda$ being very small. Then, because any state $| \psi_m \rangle$ has an energy much different from $E_0$, when the perturbation is small, it is (almost) impossible for the system to reach a state $| \psi_m \rangle$ from one state with energy $E_0$.

Note: for a classical system, this would be totally impossible due to energy conservation. In a classical system, if we start from a state with energy $E_0$ and then add a small amount of energy $\delta E$ to the system, the final states must have energy $E_0 + \delta E$, which would be very close to $E_0$, if $\delta E$ is small. So, it is absolutely impossible to have a final states with energy very different from $E_0$. But for a quantum system, anything is
possible (think about quantum tunneling, classically it is impossible, but for a quantum system it become possible). However, we know that in quantum mechanics, the probability for us to reach such a final states is small (although not exactly zero). Since the probability is small, to the leading order, we can ignore that probability. This is the key reason why we can ignore all those states \( |\psi_m\rangle \).

Since it is highly unlikely to reach \( |\psi_m\rangle \), we can ignore them to the leading order approximation. After we ignore all of them, our Hilbert space becomes very small, only \( n \) quantum states now. And thus our quantum operators becomes \( n \times n \) matrices. If \( n = 2 \), we can easily find the eigenvalue. If \( n = 3 \) or 4, we can get the eigenvalue (with analytic form) with a little bit of help (e.g. software like Mathematica). If \( n > 4 \) but not extremely huge), say a couple of hundreds or smaller, we can easily get the eigenvalue numerically using available software packages. If \( n = 10 \) or 100 million, it is not easy to get all the eigenvalues for the system, but we can easily get the smallest several or the largest several numerically using techniques like Lanczos algorithm.

After taking care of the \( n \times n \) matrices, we may be able to take \( |\psi_m\rangle \) back into consideration by going to higher orders in the perturbation theory.

### 2.4.16. Example: (textbook page 262)

Consider a 3D infinite cubical potential well

\[
V(x, y, z) = \begin{cases} 
0 & \text{if } 0 < x < a, \ 0 < y < a \text{ and } 0 < z < a \\
\epsilon & \text{otherwise}
\end{cases}
\]

The Hamiltonian for this system is

\[
H_0 = \frac{p^2}{2m} + V(x, y, z) = -\frac{\hbar^2}{2m} \nabla^2 + V(x, y, z)
\]

The eigenstates of \( H_0 \) are sin waves

\[
\psi_{n_x n_y n_z}(x, y, z) = \left( \frac{2}{a} \right)^{3/2} \sin \left( \frac{n_x \pi x}{a} \right) \sin \left( \frac{n_y \pi y}{a} \right) \sin \left( \frac{n_z \pi z}{a} \right)
\]

where \( n_x, n_y \) and \( n_z \) are positive integers. The eigenenergy for such a state is

\[
E_{n_x n_y n_z}^0 = \frac{\pi^2 \hbar^2}{2m a^2} (n_x^2 + n_y^2 + n_z^2)
\]

The ground state is obviously \( n_x = n_y = n_z = 1 \), which has energy \( E_{111}^0 = \frac{3 \pi^2 \hbar^2}{2ma^2} \). For simplicity, we will call this energy \( E_0^0 = \frac{3 \pi^2 \hbar^2}{2ma^2} \), where the subscript 0 represents the ground states.

There are three (degenerate) first excited states with \( n_x, n_y \) and \( n_z \) being (1, 1, 2) or (1,2,1) or (2,1,1).

\[
E_{112}^0 = E_{121}^0 = E_{211}^0 = \frac{\pi^2 \hbar^2}{2m a^2} (1 + 1 + 2^2) = \frac{3 \pi^2 \hbar^2}{2ma^2}
\]

For simplicity, we will call the energy of this first excited states \( E_1^0 \) where the subscript 1 means that this is for the first excited states. In addition, for simplicity, we define

\[
\psi_a = \psi_{112} \quad \text{and} \quad \psi_b = \psi_{121} \quad \text{and} \quad \psi_c = \psi_{211}
\]

Now, consider a perturbation

\[
H = H_0 + \lambda H'
\]

where

\[
H' = \begin{cases} 
V_0 & \text{if } 0 < x < a/2, \ 0 < y < a/2 \text{ and } 0 < z < a \\
0 & \text{otherwise}
\end{cases}
\]

For the ground state, we shall do non-degenerate perturbation theory, because there is no degeneracy, and the first order correction to the energy is
For the first excited states, there is a three-fold degeneracy, so we need to define a 3×3 matrix. For this matrix, it has three eigenvalues (all are real numbers). The equation for eigenvalues is:

\[
\text{det}(E I - W) = 0
\]

For diagonal terms

\[
W_{aa} = \langle \psi_a | H' | \psi_a \rangle = \int dx\, dy\, dz \left| \psi_{12}^0(x, y, z) \right|^2 \left| H' \right|^2 \left| \psi_{12}^0(x, y, z) \right|^2 = V_0 \int_0^{\pi/2} dx \int_0^{\pi/2} dy \int_0^{\pi/2} dz \left| \psi_{12}^0(x, y, z) \right|^2
\]

Similarly, we can show that

\[
W_{aa} = W_{bb} = W_{cc} = V_0/4
\]

For \(W_{ab}\), we have

\[
W_{ab} = \langle \psi_a | H' | \psi_b \rangle = \int dx\, dy\, dz \left| \psi_{12}^0(x, y, z) \right|^2 \left| H' \right|^2 \left| \psi_{11}^0(x, y, z) \right|^2 = V_0 \int_0^{\pi/2} dx \int_0^{\pi/2} dy \int_0^{\pi/2} dz \left| \psi_{12}^0(x, y, z) \right|^2
\]

One can show that the last integral \(\int_0^{\pi/2} dz \sin \left( \frac{\pi}{a} z \right) \sin \left( \frac{2\pi}{a} z \right)\) is zero. So \(W_{ab} = 0\). Similarly, \(W_{ab} = W_{ac} = 0\).

Finally, for \(W_{bc}\)

\[
W_{bc} = \langle \psi_b | H' | \psi_c \rangle = \int dx\, dy\, dz \left| \psi_{12}^0(x, y, z) \right|^2 \left| H' \right|^2 \left| \psi_{12}^0(x, y, z) \right|^2 = V_0 \int_0^{\pi/2} dx \int_0^{\pi/2} dy \int_0^{\pi/2} dz \left| \psi_{12}^0(x, y, z) \right|^2
\]

Thus, we have

\[
W = \begin{pmatrix}
  \frac{16}{9\pi^2} & 0 & 0 \\
  0 & \frac{16}{9\pi^2} & V_0 \\
  0 & V_0 & \frac{16}{9\pi^2}
\end{pmatrix}
\]

For this matrix, it has three eigenvalues (all are real numbers). The equation for eigenvalues is:

\[
\text{det}(E I - W) = 0
\]

And thus

\[
\left( E - \frac{V_0}{4} \right) \left( E - \frac{V_0}{4} \right)^2 - \left( \frac{16}{9\pi^2} V_0 \right)^2 = 0
\]
\[
\left(E - \frac{V_0}{4}\right)\left(E - \frac{V_0}{4} + \frac{16}{9\pi^2}V_0\right)\left(E - \frac{V_0}{4} - \frac{16}{9\pi^2}V_0\right) = 0
\] 

(2.258)

So the solutions are
\[
E_1^1 = \frac{V_0}{4} - \frac{16}{9\pi^2}V_0 = \frac{V_0}{4}\left[1 - \left(\frac{8}{3\pi}\right)^2\right]
\]

(2.259)

\[
E_2^1 = \frac{V_0}{4}
\]

(2.260)

\[
E_3^1 = \frac{V_0}{4} + \frac{16}{9\pi^2}V_0 = \frac{V_0}{4}\left[1 + \left(\frac{8}{3\pi}\right)^2\right]
\]

(2.261)

So the energies of the first excited states are
\[
E_1 = \left\{\begin{array}{l}
E_1^0 + \lambda E_1^1 + \ldots = 3 \frac{\alpha_1\beta_1}{m_0^2} + \lambda \frac{V_0}{4}\left[1 - \left(\frac{8}{3\pi}\right)^2\right] \text{ state } \#1 \\
E_2^0 + \lambda E_2^1 + \ldots = 3 \frac{\alpha_2\beta_2}{m_2^2} + \lambda \frac{V_0}{4}\left[1 - \left(\frac{8}{3\pi}\right)^2\right] \text{ state } \#2 \\
E_3^0 + \lambda E_3^1 + \ldots = 3 \frac{\alpha_3\beta_3}{m_3^2} + \lambda \frac{V_0}{4}\left[1 + \left(\frac{8}{3\pi}\right)^2\right] \text{ state } \#3
\end{array}\right.
\]

(2.262)

Now, for the eigenstates of the \( W \) matrix
\[
\begin{pmatrix}
\frac{V_0}{4} & 0 & 0 \\
0 & \frac{V_0}{4} & \frac{16}{9\pi^2}V_0 \\
0 & \frac{16}{9\pi^2}V_0 & \frac{V_0}{4}
\end{pmatrix}
\begin{pmatrix}
\alpha_1 \\
\beta_1 \\
\gamma_1
\end{pmatrix}
= E_1^1
\begin{pmatrix}
\alpha_1 \\
\beta_1 \\
\gamma_1
\end{pmatrix} = \frac{V_0}{4}\left[1 - \left(\frac{8}{3\pi}\right)^2\right]
\begin{pmatrix}
\alpha_1 \\
\beta_1 \\
\gamma_1
\end{pmatrix}
\]

(2.263)

\[
\begin{pmatrix}
V_0/4 \\
0 \\
0
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & \left(\frac{8}{3\pi}\right)^2 \\
0 & \left(\frac{8}{3\pi}\right)^2 & 1
\end{pmatrix}
\begin{pmatrix}
\alpha_1 \\
\beta_1 \\
\gamma_1
\end{pmatrix} = \frac{V_0}{4}\left[1 - \left(\frac{8}{3\pi}\right)^2\right]
\begin{pmatrix}
\alpha_1 \\
\beta_1 \\
\gamma_1
\end{pmatrix}
\]

(2.264)

By canceling \( V_0/4 \) on both sides, we get
\[
\begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & \left(\frac{8}{3\pi}\right)^2 \\
0 & \left(\frac{8}{3\pi}\right)^2 & 1
\end{pmatrix}
\begin{pmatrix}
\alpha_1 \\
\beta_1 \\
\gamma_1
\end{pmatrix} = \left[1 - \left(\frac{8}{3\pi}\right)^2\right]
\begin{pmatrix}
\alpha_1 \\
\beta_1 \\
\gamma_1
\end{pmatrix}
\]

(2.265)

It means that
\[
\alpha_1 = \left[1 - \left(\frac{8}{3\pi}\right)^2\right]\alpha_1
\]

(2.266)

\[
\beta_1 + \left(\frac{8}{3\pi}\right)^2\gamma_1 = \left[1 - \left(\frac{8}{3\pi}\right)^2\right]\beta_1
\]

(2.267)

\[
\left(\frac{8}{3\pi}\right)^2\beta_1 + \gamma_1 = \beta_1
\]

(2.268)

The first equation means \( \alpha_1 = 0 \) and the last two means \( \beta_1 = -\gamma_1 \). In addition, we know that normalization condition requires \( |\alpha|^2 + |\beta|^2 + |\gamma|^2 = 1 \), so
\[
\begin{pmatrix}
\alpha_1 \\
\beta_1 \\
\gamma_1
\end{pmatrix} = \begin{pmatrix}
0 \\
\frac{1}{\sqrt{2}} \\
-\frac{1}{\sqrt{2}}
\end{pmatrix}
\]

(2.269)

So,
In particular, if \( \langle \psi \rangle \) basis is extremely low (we will in general have \( A \neq \psi \)). As we have shown above, the key in degenerate perturbation theory is to find a good set of basis, such that \( \psi = \psi_a + \psi_b + \psi_c = \psi_{211} = \left( \frac{2}{a} \right)^{3/2} \sin \left( \frac{\pi}{a} x \right) \sin \left( \frac{2\pi}{a} y \right) \sin \left( \frac{\pi}{a} z \right) \) (2.270)

Using the same approach, we find that

\[
\begin{pmatrix}
\alpha_2 \\ \beta_2 \\ \gamma_2
\end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}
\]

and thus

\[
\psi(x, y, z) = \alpha_2 \psi_a + \beta_2 \psi_b + \gamma_2 \psi_c = \psi_{211} = \left( \frac{2}{a} \right)^{3/2} \sin \left( \frac{\pi}{a} x \right) \sin \left( \frac{\pi}{a} y \right) \sin \left( \frac{\pi}{a} z \right)
\]

Finally, for the third eigenstate, we can use the same method to show that

\[
\begin{pmatrix}
\alpha_3 \\ \beta_3 \\ \gamma_3
\end{pmatrix} = \begin{pmatrix} 0 \\ -1 \sqrt{2} \\ 1 \sqrt{2} \end{pmatrix}
\]

and thus

\[
\psi(x, y, z) = \alpha_3 \psi_a + \beta_3 \psi_b + \gamma_3 \psi_c = \psi_{211} = \left( \frac{2}{a} \right)^{3/2} \sin \left( \frac{\pi}{a} x \right) \sin \left( \frac{\pi}{a} y \right) \sin \left( \frac{\pi}{a} z \right)
\]

2.4.17. A small trick for finding \( \psi_1 \) and \( \psi_2 \)

We use two-fold degenerate here as an example, but the conclusion here can be easily generalized.

As we have shown above, the key in degenerate perturbation theory is to find a good set of basis, such that \( \langle \psi \rangle \) basis is extremely low (we will in general have \( A \neq \psi \)). In the most general situation, we state from a set of states \( \{ \psi_a \} \) and \( \{ \psi_b \} \). If it is a good set already (i.e., \( \langle \psi_a \rangle \ H \ | \psi_b \rangle = 0 \)), we don't need to find another basis. We can just use them and

\[
E = E_0 + \lambda \langle \psi_a \rangle \ H \ | \psi_a \rangle + O(\lambda^2)
\]

and

\[
E = E_0 + \lambda \langle \psi_b \rangle \ H \ | \psi_b \rangle + O(\lambda^2)
\]

In general, we would be so lucky. i.e., if we just randomly choose a a set of states \( \{ \psi_a \} \) and \( \{ \psi_b \} \), the chance for this basis to be a good set of basis is extremely low (we will in general have \( \langle \psi_a \rangle \ H \ | \psi_b \rangle \neq 0 \)). Is there a way to help us pick \( \{ \psi_a \} \) and \( \{ \psi_b \} \)? The answer is yes, for some cases.

If there is another quantum operator \( \hat{A} \), which compute with both \( H_0 \) and \( H' \), then we can use the common eigenstates of \( \hat{A} \) and \( H_0 \) as the basis for \( H_0 \)

\[
H_0 \ | \psi_a \rangle = E_0 \ | \psi_a \rangle
\]

\[
H_0 \ | \psi_b \rangle = E_0 \ | \psi_b \rangle
\]

\[
\hat{A} \ | \psi_a \rangle = A_a \ | \psi_a \rangle
\]

\[
\hat{A} \ | \psi_b \rangle = A_b \ | \psi_b \rangle
\]

In particular, if \( A_a \neq A_b \), then \( \{ \psi_a \} \) and \( \{ \psi_b \} \) are already a good set of basis.

To see this, we just need to prove that \( \langle \psi_a \rangle \ | H' \ | \psi_b \rangle = 0 \)

Because we have assumed that \( \hat{A} \ H' = \hat{A} \ H' = 0 \),
\[ 0 = \langle \psi_0^0 | \hat{A} H' - H' \hat{A} | \psi_0^0 \rangle = \langle \psi_0^0 | \hat{A} H' | \psi_0^0 \rangle - \langle \psi_0^0 | H' \hat{A} | \psi_0^0 \rangle = A_a \langle \psi_0^0 | H' | \psi_0^0 \rangle - A_b \langle \psi_0^0 | H' | \psi_0^0 \rangle = (A_a - A_b) \langle \psi_0^0 | H' | \psi_0^0 \rangle \] (2.281)

If \( A_a \neq A_b \), this equation means that \( \langle \psi_0^0 | H' | \psi_0^0 \rangle = 0 \).

For this situation, although we have a degeneracy, one can just do non-degenerate perturbation for \( | \psi_a^0 \rangle \) and \( | \psi_b^0 \rangle \) (separately) and there will be no singularities at all.

## 2.5. the fine structure of a hydrogen atom

### 2.5.1. Relativistic correction

In QMI, we solved an ideal model for a hydrogen atom (i.e. a particle in \( 1/r \) potential). In a real hydrogen atom, that model missed some of the physics, and one of them is relativistic effects.

Q: what is the energy of a particle, if the particle is moving at speed \( v \) and the rest mass \( m \).

\[ E = M c^2 = \frac{m}{\sqrt{1 - \frac{v^2}{c^2}}} \] (2.282)

Q: what is the momentum of a particle, if the particle is moving at speed \( v \) and the rest mass \( m \).

\[ p = M v = \frac{m}{\sqrt{1 - \frac{v^2}{c^2}}} v \] (2.283)

As a result,

\[ E = \sqrt{p^2 c^2 + m^2 c^4} \] (2.284)

To prove this relation, we start from the r.h.s.,

\[ \sqrt{p^2 c^2 + m^2 c^4} = \sqrt{\frac{m^2 v^2 c^2}{1 - \frac{v^2}{c^2}} + \frac{m^2 c^4 (1 - \frac{v^2}{c^2})}{1 - \frac{v^2}{c^2}}} = \sqrt{\frac{m^2 v^2 c^2 + m^2 c^4 - m^2 c^2 v^2}{1 - \frac{v^2}{c^2}}} = \sqrt{\frac{m^2 c^4}{1 - \frac{v^2}{c^2}}} = \frac{m}{\sqrt{1 - \frac{v^2}{c^2}}} c^2 = E \] (2.285)

This relation between \( E \) and \( p \) is an very important relation for relativistic physics!

Q: what is kinetic energy?

A: First, find the energy of a particle when it is not moving \( p = 0 \). Then we measure the energy again when it is moving (with momentum \( p \)). The energy difference between them is the kinetic energy for this particle.

\[ T = E - m c^2 = \sqrt{p^2 c^2 + m^2 c^4} - m c^2 = m c \left( \sqrt{1 + \left( \frac{p}{m c} \right)^2} - 1 \right) \] (2.286)

When particle is moving at low velocity (\( v << c \), \( p << m c \), and thus \( p/m c << 1 \). As a result, we can use the following expansion

\[ \sqrt{1 + x} = 1 + \frac{x}{2} - \frac{x^2}{8} + \ldots \] (2.287)

i.e.,

\[ \sqrt{1 + x} - 1 = \frac{x}{2} - \frac{x^2}{8} + \ldots \] (2.288)
So,

\[
T = mc^2 \left( \frac{1}{2} \left( \frac{p}{mc} \right)^2 - \frac{1}{8} \left( \frac{p}{mc} \right)^4 + \ldots \right) = \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \ldots \tag{2.289}
\]

The first term here is the kinetic energy in classical mechanics. In relativistic physics, the kinetic energy is NOT just \(p^2/2m\). Instead, we have a lot of corrections. These corrections are small if a particle is moving at low speed. There, we can treat them as perturbation

\[
H' = -\frac{p^4}{8m^3c^2} \tag{2.290}
\]

**NOTE:** this treatment is NOT the rigorous way to combine special relativity with quantum mechanics, because this treatment has one major flaw. At larger \(p\) or small \(m\) (e.g. consider a very light particle), the series will diverge. This problem comes from the fact that we used square root in the definite of the Hamiltonian. Square root is NOT an analytic function near small \(x\), and thus will cause trouble (to see this, think about \(f(x) = \sqrt{x}\), one can easily show that for the first order derivative, \(x=0\) is infinite \(\lim_{x \to 0} f'(x) \to \infty\)). The correct way to do it is to use a matrix. Notice that for a matrix, square root arises naturally (e.g., the eigenvalue of \(\begin{pmatrix} mc^2 & p c \\ p c & -mc^2 \end{pmatrix}\) are \(\pm \sqrt{p^2c^2 + m^2c^4}\). We get square root without having any square root in the matrix). The person who figured this out is Dirac and this is Dirac's theory for relativistic fermions.

If we ignore higher order terms, our hydrogen atom should follow this Hamiltonian

\[
H = H_0 + H' \tag{2.291}
\]

where \(H_0 = \frac{p^2}{2m} + V(r)\). With the perturbation \(H'\), the energy of a hydrogen atom will be different from what we computed early on. How large is the difference? This question can be answered by the perturbation theory.

We have already known the energy spectrum of \(H_0\),

\[
E_n^0 = -\frac{13.6 \text{ eV}}{n^2} \quad \text{with} \quad n = 1, 2, 3, \ldots \tag{2.292}
\]

More precisely,

\[
E_n^0 = -\frac{1}{n^2} \frac{m}{2\hbar^2} \left( \frac{e^2}{4\pi\epsilon_0} \right)^2 \tag{2.293}
\]

We often define Bohr radius \(a\) as

\[
a = \frac{\hbar^2}{4\pi\epsilon_0} \frac{4\pi\epsilon_0}{m} e^2 \tag{2.294}
\]

And then,

\[
E_n^0 = -\frac{1}{n^2} \frac{1}{2} \left( \frac{e^2}{4\pi\epsilon_0 a} \right) \tag{2.295}
\]

For \(E_n\), there are \(n^2\) degenerate quantum states (ignore spin at this moment) \(\psi_{n l m}\) where \(l\) is the angular momentum quantum number \(l = 0, 1, 2, \ldots n - 1\) and \(m\) is the quantum number for \(L_z\) and \(m = -l, -l + 1, \ldots 0, \ldots, l - 1, l\)

\[
L_z^2 \psi_{n l m} = \hbar^2 (l + 1) \psi_{n l m} \tag{2.296}
\]

\[
L_z \psi_{n l m} = \hbar m \psi_{n l m} \tag{2.297}
\]

For, \(n = 1\) there is no degeneracy, and we can do non-degenerate perturbation theory. For any \(n > 1\), there are \(n\) degenerate states, and thus we should do degenerate perturbation theory. However, we are very lucky here. We don’t need to worry about degenerate perturbation theory, because \(\psi_{n l m}\) is already a good set of basis:

\[
\langle \psi_{n l m} | H' | \psi_{n l' m} \rangle = 0 \tag{2.298}
\]

if \(l \neq l'\) or \(m \neq m'\).
This is because both \( H_0 \) and \( H' \) commute with \( L^2 \). And we can also show that both \( H_0 \) and \( H' \) commute with \( L_z \). Here, \( L^2 \) and \( L_z \) serve as the \( \hat{A} \) operator that we defined in the previous section. For a fixed \( n \), because the degenerate states all have different eigenvalues for \( L^2 \) and \( L_z \) (different \( l \) and \( m \)), \( \langle \psi_{nlm} | H' | \psi_{nlm} \rangle = 0 \). So we don’t need to choose any other basis and can start with non-degenerate perturbation.

The correction to the energy is (to the first order)

\[
E_{nlm} = \langle \psi_{nlm} | H' | \psi_{nlm} \rangle = -\frac{1}{8m^3c^2} \langle \psi_{nlm} | p^2 | \psi_{nlm} \rangle = -\frac{1}{8m^3c^2} \langle \psi_{nlm} | p^2 | \psi_{nlm} \rangle = -\frac{1}{8m^3c^2} \langle \psi_{nlm} | p^2 | \psi_{nlm} \rangle
\]  

(2.299)

For \( | \psi_{nlm} \rangle \), we know that

\[
H_0 | \psi_{nlm} \rangle = E_{nlm} | \psi_{nlm} \rangle
\]  

(2.300)

\[
\left( \frac{\hat{\mathbf{p}}}{2m} + \hat{V} \right) | \psi_{nlm} \rangle = E_{nlm} | \psi_{nlm} \rangle
\]  

(2.301)

\[
\left( \frac{\hat{\mathbf{p}}}{2m} \right) | \psi_{nlm} \rangle = (E_{nlm} - V) | \psi_{nlm} \rangle
\]  

(2.302)

\[
p^2 | \psi_{nlm} \rangle = 2m(E_{nlm} - V) | \psi_{nlm} \rangle
\]  

(2.303)

\[
p^2 | \psi_{nlm} \rangle = 2m(E_{nlm} - V) | \psi_{nlm} \rangle
\]  

(2.304)

The conjugate of this equation gives

\[
\langle \psi_{nlm} | p^2 = \langle \psi_{nlm} | 2m(E_{nlm} - V)
\]  

(2.305)

So,

\[
\langle \psi_{nlm} | p^2 | \psi_{nlm} \rangle = \langle \psi_{nlm} | 2m(E_{nlm} - V) | \psi_{nlm} \rangle = 4m^2 \langle \psi_{nlm} | (E_{nlm} - V)^2 | \psi_{nlm} \rangle
\]  

(2.306)

Here, \( V = -\frac{e^2}{4\pi\epsilon_0 r} \) and

\[
\left( \frac{\hat{\mathbf{p}}}{2m} \right) | \psi_{nlm} \rangle = (E_{nlm} - V) | \psi_{nlm} \rangle
\]  

(2.307)

Without going into details, we will just show the results here

\[
\left( \frac{\hat{\mathbf{p}}}{2m} \right) | \psi_{nlm} \rangle = \frac{1}{n^2a}
\]  

(2.308)

\[
\left( \frac{\hat{\mathbf{p}}}{2m} \right) | \psi_{nlm} \rangle = \frac{1}{n^2(l + \frac{1}{2})a^2}
\]  

(2.309)

Thus,

\[
E_{nlm} = -\frac{1}{8m^3c^2} \langle \psi_{nlm} | p^2 | \psi_{nlm} \rangle = -\frac{1}{8m^3c^2} 4m^2 \left[ (E_{nlm})^2 + E_{nlm}^2 \frac{e^2}{2\pi\epsilon_0 n^2a} + \left( \frac{e^2}{4\pi\epsilon_0 a} \right)^2 \frac{1}{n^2(l + \frac{1}{2})a^2} \right]
\]  

(2.310)

\[
E_{nlm} = -\frac{1}{8m^3c^2} 4m^2 \left[ (E_{nlm})^2 + 2E_{nlm} \frac{e^2}{2\pi\epsilon_0 n^2a} + \left( \frac{e^2}{4\pi\epsilon_0 a} \right)^2 \frac{1}{n^2(l + \frac{1}{2})a^2} \right]
\]  

(2.311)

As we have shown early on

\[
E_{nlm} = -\frac{1}{n^2} \frac{1}{2}\left( \frac{e^2}{4\pi\epsilon_0 a} \right)
\]  

(2.311)
\[ -2n^2 E_n^0 = \frac{e^2}{4\pi \varepsilon_0 a} \]  \hspace{1cm} (2.312)

\[ E_{nlm} = -\frac{1}{2m c^2} \left[ \left( E_n^0 \right)^2 + \frac{2 E_n^0}{n^2} \frac{e^2}{4\pi \varepsilon_0 a} + \frac{1}{n^3(l + \frac{1}{2})} \left( \frac{e^2}{4\pi \varepsilon_0 a} \right)^2 \right] = \]

\[ -\frac{1}{2m c^2} \left( E_n^0 \right)^2 - 2 E_n^0 \frac{e^2}{n^2} - 2 E_n^0 + \frac{1}{n^3(l + \frac{1}{2})} \left( -2n^2 E_n^0 \right)^2 \] = \left( E_n^0 \right)^2 \left( \frac{4n}{2m c^2} \left( l + \frac{1}{2} \right) - 3 \right) \hspace{1cm} (2.313)

So, the eigen-energy in a H atom shall be

\[ E_{nlm} = E_n^0 + E_{nlm}^1 + \ldots = E_n^0 \left( \frac{4n}{2m c^2} \left( l + \frac{1}{2} \right) - 3 \right) + \ldots \hspace{1cm} (2.314) \]

The zeroth order term

\[ E_n^0 = -\frac{1}{2} \frac{m}{n^2} \frac{e^2}{\hbar^2} \left( \frac{e^2}{4\pi \varepsilon_0 \hbar c} \right)^2 \hspace{1cm} (2.315) \]

it is proportional to

\[ E_n^0 \propto \frac{m}{\hbar^2} \left( \frac{e^2}{4\pi \varepsilon_0 \hbar c} \right)^2 \left( \frac{e^2}{4\pi \varepsilon_0 \hbar c} \right) = \left( \frac{e^2}{4\pi \varepsilon_0 \hbar c} \right)^2 m c^2 \hspace{1cm} (2.316) \]

The prefactor \( \frac{e^2}{4\pi \varepsilon_0 \hbar c} \) is a very important physics constant, known as the fine structure constant.

\[ \alpha = \frac{e^2}{4\pi \varepsilon_0 \hbar c} \approx \frac{1}{137.036} \hspace{1cm} (2.317) \]

So,

\[ E_n^0 \propto \alpha^2 m c^2 \hspace{1cm} (2.318) \]

The first order term

\[ E_{nlm}^1 \propto \frac{E_n^0}{m c^2} \propto \alpha^4 m^2 c^2 \hspace{1cm} (2.319) \]

If we compare the first order and zeroth order term,

\[ \frac{E_{nlm}^1}{E_n^0} \propto \frac{\alpha^4 m^2 c^2}{\alpha^2 m^2 c^2} = \alpha^2 \approx \left( \frac{1}{137.036} \right)^2 \approx \frac{1}{10000} \hspace{1cm} (2.320) \]

So indeed, the perturbation theory works, i.e. higher order term is much smaller than the leading order. (remember that Taylor expansions only converge when the small parameter \( \lambda \) is small enough. Here, our small parameter is \( \alpha \), which is smaller than 1%).

**Relativistic correction is indeed a small correction in a H atom**

NOTE: the fine structure constant is one of the most important physics constant. It is dimensionless. It involves special relativity (contains the speed of light). It involves quantum mechanics (having \( \hbar \) in its definition) and it also involves E&M (having \( \varepsilon_0 \) and \( e \)). In this section, we showed that it is so lucky for us that for a H atom, because \( \alpha \) is small, our relativistic correction is indeed small and thus we can do perturbation theory. In QFT (QED), small \( \alpha \) means that interactions between particles (quantum electron-dynamics) is a small perturbation. To the leading order, we can treat a particle as a free particle, and then add E&M interactions as a perturbation. Because the small parameter \( \alpha \) is so small, in QED, our perturbation theory converge very fast. First order perturbation gives us an accuracy of the order \( \alpha^4 \sim 10^{-2} \). Second order perturbation increases the accuracy to \( \alpha^2 \sim 10^{-4} \). By going to 5th order in perturbation, we can get an accuracy of the order \( \alpha^5 \sim 10^{-10} \). This is the reason why QED is such a successful theory.

**2.5.2. Spin-orbit coupling**
In QM I, we treat the spin of an electron as an independent quantity, independent from the orbit angular momentum. For example in a hydrogen atom, for any eigenwavefunction $\psi_{nlm}(x, y, z)$, it actually means two degenerate states: (1) one electron with wavefunction $\psi_{nlm}(x, y, z) \text{ and } \text{spin up}$ and (2) one electron with wavefunction $\psi_{nlm}(x, y, z) \text{ and } \text{spin down}$. This conclusion remains the same after we take into account the relativistic correction (now the eigenenergy depends on both $n$ and $l$, but still for every eigenwavefunction, it means two degenerate state when we taken into account the spins).

In this section, we will consider one more effect, which was ignored previously and this effect will tell us the spin of an electron and its orbital motion are coupled together.

**Warning:** you may find that the derivation in this section very disturbing. Because for multiple times, after some derivations, we will say the following without providing much justification, “by the way, this result is in fact not quite right, and we will need to throw in an extra factor of 2 to get the correct answer.” The reason for these extra factors of 2 is because this section is NOT treating spin-orbit coupling in the rigorous and correct way, which requires Dirac’s equation. Instead, what we are trying to do here is to use various tricks trying to recover Dirac’s finally conclusion without using Dirac’s equation. These tricks (they are not rigorous at all) get some part of the story right, but in many cases, they lead to wrong results. Because we already know the right answer from Dirac’s equation, whenever we find that these tricks fail to get the correct answer, we will correct it by adding some extra factor. Within our deviation, these extra factors look totally unreasonable and weird, but if one start from Dirac’s equation, the results are all very natural and straightforward. *Bottom line, please don’t take these derivations very seriously, because they are not supposed to give (fully) correct description after all. But the physics, at the end of the day, is correct.*

### Magnetic dipole of an electron

If a charge particle moves in circles, it creates circular current, and the circular current will result in a magnetic dipole moment (according to E&M). To see this, we use a simple model to demonstrate this physics. Assuming that we have a ring, and there is a charged particle (with charge $q$) moving around the ring. The dipole moment is

$$\vec{\mu} = \frac{1}{2} q \vec{r} \times \vec{v}$$

(2.321)

where $q$ is the charge of the particle. $\vec{r}$ and $\vec{v}$ are the location and velocity of the particle.

$$\vec{\mu} = \frac{1}{2} q \vec{r} \times \vec{v} = \frac{q}{2m} (m \vec{r} \times \vec{v}) = \frac{q}{2m} \vec{L}$$

(2.322)

where $\vec{L} = m \vec{r} \times \vec{v}$ is the angular momentum.

Similarly, if we have a spinning charged particle, the angular momentum from the spin will also result in a magnetic dipole. The dipole moment from spins is also proportional to the angular momentum of the spin, but with an extra factor known as the $g$-factor

$$\vec{\mu} = g \frac{q}{2m} \vec{S}$$

(2.323)

The charge of an electron is $-e$ (negative charge), so

$$\vec{\mu} = -g \frac{e}{2m} \vec{S}$$

(2.324)

and $g$ is a number, whose value is really close to 2. **In this course, we will say that $g = 2$ for simplicity**, but in reality, $g = 2.00231930436182$. In Dirac theory, $g$ is exactly 2. The reason the real value of $g$ is a little bit larger than 2 is due to interactions between electrons and photons (light), which wasn’t considered in Dirac’s equation.

**Note:** in many cases, people absorb the minus sign into the definition of $g$,

$$\vec{\mu} = g \frac{e}{2m} \vec{S}$$

(2.325)

where $g = -2$. But no matter what convention one adopts,

$$\vec{\mu} = -\frac{e}{m} \vec{S}$$

(2.326)

We can define the Bohr magneton, which is a fundamental physics constant
\[\mu_B = \frac{e\hbar}{2m} = 9.27400968 \times 10^{-24} \text{ J/T}\] (2.327)

and

\[\bar{\mu} = g\frac{e\hbar}{2m} = \frac{g e \hbar}{2m} = g \mu_B \frac{\hbar}{h}\] (2.328)

For an electron, the magnetic dipole is \(\pm \mu_B\). We demonstrate this by considering the dipole moment along the z axis:

\[\mu_z = g \mu_B S_z \hbar = g \mu_B S_z \hbar\] (2.329)

The spin operator \(S_z\) has eigenvalues \(\pm \hbar/2\), and \(g = -2\). For an eigenstate of \(S_z\),

\[\mu_z = \begin{cases} -\mu_B & \text{if } S_z \text{ eigenvalue is } +\hbar/2, \text{ i.e. spin up} \\ +\mu_B & \text{if } S_z \text{ eigenvalue is } -\hbar/2, \text{ i.e. spin down} \end{cases}\] (2.330)

Effective B field from the nucleon

If we stand on an electron (using the electron as our reference frame), we will find that the nucleon is moving around us in a circle. Because the nucleon has positive charge \(+e\), when it moves around us, it generates a circular current and thus leads to a magnetic field. According to the "Biot–Savart law" in E&M, the B field generated by a wire with current \(I\) is

\[\vec{B} = \frac{\mu_0 I}{4\pi} \int d\vec{l} \times \vec{r} \over r^3\] (2.331)

where \(d\vec{l}\) is a small section of the wire and the direction is parallel to the wire, \(\vec{r}\) is the distance between the wire and the place at which we want to measure the B field. For a circular motion of the nucleon, the wire here is a circle and we want to know the B field at the center of the circle.

\[B = \frac{\mu_0 I}{4\pi} \int_0^{2\pi} \frac{e r d\theta r}{r^3} = \frac{\mu_0 I}{4\pi} \int_0^{2\pi} \frac{e}{2r} = \frac{\mu_0 I}{2r}\] (2.332)

The current \(I\) here is

\[I = \frac{e}{T}\] (2.333)

where \(e\) is (the absolute value of) the charge of an electron (remember that the nucleon in a hydrogen atom is \(+e\)). \(T\) is the time it takes for the nucleon to go around a circle.

\[I = \frac{e}{2\pi/\omega} = \frac{e\omega}{2\pi}\] (2.334)

Here, \(\omega\) is the angular velocity. Notice that the angular velocity here is the same as the angular velocity of the electron \(\omega\) (in the rest frame).

\[I = \frac{e\omega}{2\pi} = \frac{eL}{2\pi m r^2}\] (2.335)

So

\[B = \frac{\mu_0 I}{2r} = \frac{\mu_0 e L}{2\pi m r^2} = \frac{e L}{4\pi \epsilon_0 m r^3}\] (2.336)

Notice that \(e_0 \mu_0 = 1/c^2\) and in addition, it is easy to realize that \(\vec{B} \parallel \vec{L}\), so we get

\[\vec{B} = \epsilon_0 \mu_0 \frac{e L}{4\pi \epsilon_0 m r^3} = \frac{e}{4\pi \epsilon_0 m c^2 r^3} \vec{L}\] (2.337)

Magnetic dipole in an B field

We proved above that using the frame of the electron, the electron feels a B field, which is generated by the nucleon.
\[ \vec{B} = \frac{e}{4\pi\varepsilon_0} \frac{\vec{L}}{mc^2 r^3} \]  

(2.338)

and the electron has a magnetic dipole

\[ \vec{\mu} = -\frac{e}{m} \hat{S} \]  

(2.339)

We have a dipole in a \( B \) field, we shall have energy

\[ H' = -\vec{\mu} \cdot \vec{B} = \frac{e^2}{4\pi\varepsilon_0} \frac{1}{m^2 c^2 r^3} \hat{S} \cdot \vec{L} \]  

(2.340)

Here, our naive tricks miss a factor of 1/2 in comparison to the correct result (from Dirac’s equation). The right result should be

\[ H_{SO} = \frac{e^2}{8\pi\varepsilon_0} \frac{1}{m^2 c^2 r^3} \hat{S} \cdot \vec{L} \]  

(2.341)

Therefore, we shall add one extra term to the Hamiltonian

\[ H = H_0 + H_{SO} \]  

(2.342)

Here, \( H_0 \) is what we learned in QM1. And \( H_{SO} \) is this new term. Here, we treat \( H_0 \) as unperturbed Hamiltonian, and treat \( H_{SO} \) as a small perturbation.

**Basis without \( H_{SO} \)**

For the commutation relations for the orbital angular momentum, we know that

\[ [L_x, L_y] = i\hbar L_z \]  

(2.343)

\[ [L_y, L_z] = i\hbar L_x \]  

(2.344)

\[ [L_z, L_x] = i\hbar L_y \]  

(2.345)

or we can write the same formular as

\[ [L_i, L_j] = i\hbar \varepsilon_{i,j,k} L_k \]  

(2.346)

where \( \varepsilon_{i,j,k} \) is the Levi-Civita symbol.

For \( L \), we know that we can define the operator \( L^2 \)

\[ L^2 = L_x^2 + L_y^2 + L_z^2 \]  

(2.347)

And we know that it commute with \( L_z \), \([L^2, L_z] = 0\). As a result, we cannot measure all the three components of the angular momentum due to the uncertainty principle. But, we can measure \( L^2 \) and \( L_z \) at the same time, by defining common eigenstates for these two operators

\[ L^2 \psi_{lm}(x, y, z) = \hbar(l + 1) \psi_{lm}(x, y, z) \]  

(2.348)

\[ L_z \psi_{lm}(x, y, z) = m \hbar \psi_{lm}(x, y, z) \]  

(2.349)

Here, \( l \) is a non-negative integer, \( l = 0, 1, 2, \ldots \), and \( m \) is an integer between \( +l \) and \( -l \). The eigenwavefunctions is very easy to write down in spherical coordinates

\[ \psi_{lm}(r, \theta, \phi) = R(r) Y_{lm}(\theta, \phi) \]  

(2.350)

where \( R(r) \) is an arbitrary function of \( r \) (the function doesn’t depend on \( \theta \) or \( \phi \)), and \( Y_{lm}(\theta, \phi) \) are a set of special functions known as the spherical harmonics.

For spins, we have the same commutation relation,

\[ [S_x, S_y] = i\hbar S_z \]  

(2.351)

\[ [S_y, S_z] = i\hbar S_x \]  

(2.352)

\[ [S_z, S_x] = i\hbar S_y \]  

(2.353)
And again, we can define
\[
S^2 = S_x^2 + S_y^2 + S_z^2
\]  
(2.354)

And same as above, \([S^2, S_z] = 0\), so we can measure \(S^2\) and \(S_z\) at the same time.

\[
S^2 \left| s, m \right> = s(s + 1) \hbar^2 \left| s, m \right>
\]  
(2.355)

\[
S_z \left| s, m \right> = m \hbar \left| s, m \right>
\]  
(2.356)

where \(s\) is a non-negative integer or half-integer, \(s = 0, 1/2, 1, 3/2, \ldots\) Once \(s\) is determined, \(m = -s, -s + 1, \ldots, s - 1, s\). For electrons, \(s = 1/2\), and thus \(m = -1/2\) or \(+1/2\).

In addition, we know that \(\vec{S}\) and \(L\) commute with each other,

\[
[\vec{S}, \vec{L}] = 0
\]  
(2.357)

Without spin-orbit coupling (i.e. the unperturbed Hamiltonian \(H_0\)), we can easily prove that \([H_0, \vec{S}] = [H_0, \vec{L}] = 0\), as a result, we find that the following operators commute with one another, \(H_0, L^2, L_z, S^2, S_z\). So we can request our quantum states to be common eigenstates of all these operators: \(|n, l, m, s, s_z\rangle\)

\[
H_0 \left| n, l, m, s, s_z \right> = -\frac{13.6 \text{ eV}}{n^2} \left| n, l, m, s, s_z \right>
\]  
(2.358)

\[
L_z \left| E_n, l, m, s, s_z \right> = l(l + 1) \hbar^2 \left| n, l, m, s, s_z \right>
\]  
(2.359)

\[
L_z \left| E_n, l, m, s, s_z \right> = m \hbar \left| n, l, m, s, s_z \right>
\]  
(2.360)

\[
S^2 \left| E_n, l, m, s, s_z \right> = \frac{3}{4} \hbar^2 \left| n, l, m, s, s_z \right>
\]  
(2.361)

\[
S_z \left| E_n, l, m, s, s_z \right> = s \hbar \left| n, l, m, s, s_z \right>
\]  
(2.362)

where \(s_z = +1/2\) or \(-1/2\). For an electron, we know that \(s\) is always \(1/2\), so we don’t really need to write it out: \(|n, l, m, s_z\rangle\). Compare to the results without spins (\(\psi_{\text{bas}}\)), the only thing we get here is an extra index \(s_z = \pm 1/2\). This quantum number tells me whether my spin is pointing up or down. At the end of the day, we didn’t get anything beyond what we have already known, except that we now need to specify whether the spin of the electron is up or down.

With SO coupling, the basis described above is NOT a good option, because \(|n, l, m, s_z\rangle \neq |n', l', m', s_z'\rangle \neq 0\), i.e. to do degenerate perturbation theory, we will need a new basis.

**Basis with \(H_{SO}\)**

To get the proper basis, we can go throw the derivation that we demonstrated for degenerate perturbation theory. Here, instead, we will use a trick to get the correct basis direction. The trick is what we proved early on. We know that if we can find a quantum operator, \(\hat{A}\), which commutes with both \(H_0\) and the perturbation \(H_{SO}\), we can use common eigenstates of \(\hat{A}\) and \(H_0\) as a set of basis. If in this set, every state has a different eigenvalue for \(\hat{A}\), then it is a good state for degenerate perturbation theory.

In the previous section (relativistic correction), we used \(L^2\) and \(L_z\) to serve as the \(\hat{A}\) operator. Here, after taking into account spins and for the perturbation \(H_{SO}\), we will need to use \(L^2, S^2, F\) and \(J_z\) as \(\hat{A}\).

If an electron have both orbit and spin angular momenta, we can add them up to get the total angular momentum

\[
\hat{J} = \vec{L} + \vec{S}
\]  
(2.363)

or equivalently,

\[
J_z = L_z + S_z
\]  
(2.364)

\[
J_y = L_y + S_y
\]  
(2.365)

\[
J_x = L_x + S_x
\]  
(2.366)
For $J_0$, we have the same commutation relation
\begin{equation}
[J_+, J_-] = i \hbar J_0
\end{equation}
\begin{equation}
[J_+, J_0] = i \hbar J_-
\end{equation}
\begin{equation}
[J_-, J_0] = i \hbar J_+
\end{equation}
And we can also define $J^2$ as
\begin{equation}
J^2 = J_+ J_- + J_0^2
\end{equation}
Same as $L$ and $S$, we know that $[J^2, J_0] = 0$. So we can measure $J^2$ and $J_0$ at the same time
\begin{equation}
J^2 \{ j, m \} = j(j + 1) \hbar^2 \{ s, m \}
\end{equation}
\begin{equation}
J_0 \{ j, m \} = m \hbar \{ s, m \}
\end{equation}
where $j$ is an integer or half-integer, $j = 0, 1/2, 1, 3/2, ...$ Once $j$ is determined, $m = -s, -s + 1, ..., s - 1, s$. For electrons, $s = 1/2$, and thus $m = -1/2$ or $1/2$.

If we have a particle with spin quantum number $s$ and orbit angular momentum quantum number $l$, then $j = l + s, l + s - 1, ... |l - s|$. NOTE: $j$ cannot be negative. For spin $s = 1/2$, this means that $j = l - 1/2$ or $j = l + 1/2$ for $l \geq 1$. And $j = 1/2$ if $l = 0$.

- If we put an electron on an $s$-wave state $(l = 0)$, the total angular momentum $j = 1/2$
- If we put an electron on an $p$-wave state $(l = 1)$, the total angular momentum $j = 1/2$ or $3/2$
- If we put an electron on an $d$-wave state $(l = 2)$, the total angular momentum $j = 3/2$ or $5/2$
- If we put an electron on an $f$-wave state $(l = 3)$, the total angular momentum $j = 5/2$ or $7/2$
- ...

In our homework, we will show that $J^2, J_0, L^2, S^2$ commute with $H_{SO}$. It is also straightforward to see that $J^2, J_0, L^2, S^2$ all commute with $H_0$, so we can use them as our $\hat{A}$ operator. In addition, it is also easy to verify that these four operators commute with each other, so we can define common eigenstates for $H_0, J^2, J_0, L^2, S^2$ and using this common eigenstates as our basis for perturbation theory
\begin{equation}
H_0 | n, l, s, j, j_z \rangle = -\frac{13.6 \text{ eV}}{n^2} | n, l, s, j, j_z \rangle
\end{equation}
\begin{equation}
L^2 \{ n, l, s, j, j_z \} = l(l + 1) \hbar^2 \{ n, l, s, j, j_z \}
\end{equation}
\begin{equation}
S^2 \{ n, l, s, j, j_z \} = s(s + 1) \hbar^2 \{ n, l, s, j, j_z \} = \frac{3}{4} \hbar^2 \{ n, l, s, j, j_z \}
\end{equation}
\begin{equation}
J^2 \{ n, l, s, j, j_z \} = j(j + 1) \hbar^2 \{ n, l, s, j, j_z \}
\end{equation}
\begin{equation}
J_0 \{ n, l, s, j, j_z \} = j \hbar \{ n, l, s, j, j_z \}
\end{equation}
Notice that
\begin{equation}
J^2 = \vec{J} \cdot \vec{J} = (\vec{L} + \vec{S}) \cdot (\vec{L} + \vec{S}) = \vec{L} \cdot \vec{L} + \vec{S} \cdot \vec{S} + 2 \vec{S} \cdot \vec{L} = L^2 + S^2 + 2 \vec{S} \cdot \vec{L}
\end{equation}
As a result,
\begin{equation}
\vec{S} \cdot \vec{L} = \frac{J^2 - L^2 - S^2}{2}
\end{equation}
So we can write our perturbation as
\begin{equation}
H_{SO} = \frac{e^2}{8 \pi \epsilon_0} \frac{1}{m^2 c^2 r^3} \vec{S} \cdot \vec{L} = \frac{e^2}{8 \pi \epsilon_0} \frac{1}{m^2 c^2 r^3} \frac{J^2 - L^2 - S^2}{2}
\end{equation}
The first order perturbation theory
\begin{equation}
E_{nljsjz}^{1} = (n, l, s, j, j_z \rangle \langle H_{SO} | n, l, s, j, j_z \rangle = \frac{e^2}{8 \pi \epsilon_0} \frac{1}{2 m^2 c^2} \frac{J^2 - L^2 - S^2}{r^3} \{ n, l, s, j, j_z \rangle \langle H_{SO} | n, l, s, j, j_z \rangle
\end{equation}
\[ \frac{e^2}{8 \pi \epsilon_0 \frac{1}{2 m^2 c^2}} \left<n, l, s, j, j_z\right| \frac{j(j + 1) - h(l + 1) - s(s + 1)}{r^3} \left| n, l, s, j, j_z\right> = \]

\[ \frac{e^2 \hbar^2}{16 \pi \epsilon_0} \left<n, l, s, j, j_z\right| \frac{j(j + 1) - h(l + 1) - s(s + 1)}{m^2 c^2} \left| n, l, s, j, j_z\right> = \]

The average value for \(1/r^3\) is known for the unperturbed Hamiltonian

\[ \left<n, l, s, j, j_z\right| \frac{1}{r^3} \left| n, l, s, j, j_z\right> = \int \frac{d^3 r}{r^3} \psi_{n,l}^\ast (r, \theta, \phi) \frac{1}{r^3} \psi_{n,l}(r, \theta, \phi) = \frac{1}{l(l + 1/2) (l + 1) n^3 a^3} \]  

(2.382)

So

\[ E_{n,l,j,j_z} = \frac{c^2 \hbar^2}{16 \pi \epsilon_0} \left<n, l, s, j, j_z\right| \frac{j(j + 1) - h(l + 1) - s(s + 1)}{m^2 c^2} \left| n, l, s, j, j_z\right> = \]

\[ \frac{c^2 \hbar^2}{16 \pi \epsilon_0} \left<n, l, s, j, j_z\right| \frac{j(j + 1) - h(l + 1) - s(s + 1)}{m^2 c^2} \left| n, l, s, j, j_z\right> = \]

\[ \frac{c^2 \hbar^2}{16 \pi \epsilon_0} \left<n, l, s, j, j_z\right| \frac{j(j + 1) - h(l + 1) - s(s + 1)}{m^2 c^2} \left| n, l, s, j, j_z\right> = \]

\[ E_{n,l,j,j_z} = \frac{1}{l(l + 1/2) (l + 1) n^3 a^3} \]

(2.383)

Remember, the Bohr radius is

\[ a = \frac{\hbar^2}{m c^2} \]

(2.384)

and the eigenenergies for the unperturbed Hamiltonian is

\[ E_n = -\frac{1}{n^2} \left( \frac{c^2}{8 \pi \epsilon_0 \frac{1}{2 m^2 c^2}} \right) \]

(2.385)

\[ E_{n,l,j,j_z} = \frac{c^2 \hbar^2}{16 \pi \epsilon_0} \left<n, l, s, j, j_z\right| \frac{j(j + 1) - h(l + 1) - s(s + 1)}{m^2 c^2} \left| n, l, s, j, j_z\right> = \]

\[ \frac{c^2 \hbar^2}{16 \pi \epsilon_0} \left<n, l, s, j, j_z\right| \frac{j(j + 1) - h(l + 1) - s(s + 1)}{m^2 c^2} \left| n, l, s, j, j_z\right> = \]

\[ \frac{c^2 \hbar^2}{16 \pi \epsilon_0} \left<n, l, s, j, j_z\right| \frac{j(j + 1) - h(l + 1) - s(s + 1)}{m^2 c^2} \left| n, l, s, j, j_z\right> = \]

(2.386)

Recall that in the previous section, we find that the relativistic correction is (at the first order)

\[ -\left( \frac{E_n^0}{2 m c^2} \right)^2 \frac{4 n}{l + \frac{1}{2}} + 3 \]

(2.387)

If we combine both effects together, to the first order, the energy is

\[ E_{n,l,j,j_z} = E_n^0 + \left( \frac{E_n^0}{2 m c^2} \right)^2 \left( \frac{2}{l(l + 1/2) (l + 1)} \left( j + 1 \right) - \frac{3}{4} n - \frac{4 n}{l + \frac{1}{2}} + 3 \right) \]

(2.388)

Notice that \(j = l + 1/2\) or \(l - 1/2\), so we have \(l = j - 1/2\) or \(j + 1/2\). For \(l = -1/2\), we find that

\[ E_{n,l,j,j_z} = E_n^0 + \left( \frac{E_n^0}{2 m c^2} \right)^2 \left( \frac{4 n}{j + 1/2} - 3 \right) \]

(2.389)

for \(l = j + 1/2\), we find exactly the same result

\[ E_{n,l,j,j_z} = E_n^0 + \left( \frac{E_n^0}{2 m c^2} \right)^2 \left( \frac{4 n}{j + 1/2} - 3 \right) \]

(2.390)

So we conclude, no matter what, we have
After taking into account both SO coupling and relativistic correction, we find that the energy of a quantum state only depends on \( n \) and \( j \),

\[
E_{n,j} = E_n^0 - \frac{(E_n^0)^2}{2m c^2} \left( \frac{4n}{j + 1/2} - 3 \right) + \ldots
\]

(2.391)

This is the fine structure correction in a hydrogen atom.

According to this formula, the fine structure correction always reduces the energy of a state by a very small fraction (the correction is \( \alpha^2 \sim 10^{-4} \), which is a 0.01% change). The smaller the \( j \) is the bigger this correction is. So s-wave states (with \( l = 0 \) get the largest modification). For states with \( l > 0 \), e.g. p-wave, d-wave, etc., they split into two different energy levels (with \( j = l - 1/2 \) and \( j = l + 1/2 \), and the former has lower energy than the latter).

NOTE: the fine structure correction can also be written as

\[
E_{n,j} = E_n^0 - \frac{(E_n^0)^2}{2m c^2} \left( \frac{4n}{j + 1/2} - 3 \right) + \ldots
\]

(2.392)

because

\[
E_n^0 = -\frac{1}{2n^2} m c^2 \left( \frac{e^2}{4 \pi \epsilon_0 h c} \right)^2 = -\frac{a^2}{2n^2} m c^2
\]

(2.393)

where \( \alpha \) is the fine structure constant

\[
\alpha = \frac{e^2}{4 \pi \epsilon_0 h c} = \frac{1}{137.036}
\]

(2.394)

we can rewrite the formula as

\[
E_{n,j} = E_n^0 \left[ 1 + \frac{\alpha^2}{4n^2} \left( \frac{4n}{j + 1/2} - 3 \right) + \ldots \right] = E_n^0 \left[ 1 + \frac{\alpha^2}{n^2} \left( \frac{n}{j + 1/2} - \frac{3}{4} \right) + \ldots \right] = -\frac{13.6 \text{ eV}}{n^2} \left[ 1 + \frac{\alpha^2}{n^2} \left( \frac{n}{j + 1/2} - \frac{3}{4} \right) + \ldots \right]
\]

(2.395)

\[
H_2 = -\vec{\mu}_L \cdot \vec{B}
\]

(2.396)

2.6. The Zeeman effect

In the previous section, we found that after considering relativistic effects (i.e., fine structure), the eigenenergies in a hydrogen atom only depend on the quantum numbers \( n \) and \( j \). In particular, the energy is independent of \( j_z \).

For a fixed \( j, j_z = -j, -j + 1, \ldots, j - 1, j \), all have the same energy, i.e. \( 2j + 1 \)-fold degeneracy.

In this section, we will show that in the presence of an external \( B \) field, these \( 2j + 1 \)-fold degeneracy will be lifted.

In a magnetic field, the energy of a magnetic dipole is \( E = -\vec{\mu} \cdot \vec{B} \). So for an atom

\[
H_2 = -\vec{\mu}_L \cdot \vec{B}
\]

(2.397)

where \( \vec{\mu}_L \) is the magnetic dipole moment from orbit motion

\[
\vec{\mu}_L = -\frac{e}{2m} \vec{L} = -\mu_B \frac{\vec{L}}{\hbar}
\]

(2.398)

and \( \vec{\mu}_S \) is the magnetic dipole moment from electron spin

\[
\vec{\mu}_S = -2\times \frac{e}{2m} \vec{S} = -\frac{e}{m} \vec{S} = -2\mu_B \frac{\vec{S}}{\hbar}
\]

(2.399)

where \( \mu_B = \frac{e\hbar}{2m} = 5.788 \times 10^{-5} \text{ eV/T} \) is Bohr magneton. Here, \( \vec{L} \) and \( \vec{S} \) are angular momenta from orbit motion and electron spin respectively.
The eigenstates of this Hamiltonian is the same as the eigenstates of
\[ H \]
When
\[ H \] for the Hamiltonian above, the key difficulty lies in the fact that the last two terms, \( H_{SO} \) and \( H_z \), do not commute with each other. For \( H_z \), we must know \( L_z \) and \( S_z \). However, we have learned early on \( H_{SO} \) doesn’t commute with \( L_z \) and \( S_z \) (\( H_{SO} \) commutes with \( j_z \), but not with \( S_z \) or \( L_z \), as we showed in our homework). So, we cannot measure \( H_{SO} \) with \( L_z \) and \( S_z \) at the same time, but \( H_z \) needs information about \( L_z \) and \( S_z \). This is the confliction.

**NOTE:** this problem comes from the g factor for electrons. \( \mu = -\mu_B \frac{B}{\hbar} \) and \( \mu = -2 \mu_B \frac{\delta}{\hbar} \), the prefactor for them are DIFFERENT! (differ by a factor of 2, i.e., the g-factor). If there is no this extra factor \( g = 2 \), things would be very easy. There, \( H_z' = \mu_B B \frac{L_z - 2 S_z}{\hbar} \), so we only need \( J_z \). But unfortunately, \( H_z' \) is not proportional to \( J_z \).

### 2.6.2. Strong field

When \( H_z' \gg H_{SO}' \), we can treat \( H_{SO}' \) and \( H_z' \) (they two are comparable as we learned in the previous section) as perturbation, and thus our unperturbed Hamiltonian is
\[ H_0 + H_z' \]

The eigenstates of this Hamiltonian is the same as the eigenstates of \( H_0 \): \( | n, l, m_l, m_s \rangle \). Here, \( | n, l, m_s \rangle \) are the eigenwavefunctions that we learned in QM I. Here, we add back the spin \( S_z \) quantum state \( m_s \).

\[ L^2 | n, l, m_l, m_s \rangle = \hbar (l + 1) | n, l, m_l, m_s \rangle \]
\[ S_z | n, l, m_l, m_s \rangle = m_s | n, l, m_l, m_s \rangle \]
\[ L_z | n, l, m_l, m_s \rangle = m_s | n, l, m_l, m_s \rangle \]
\[ H_0 | n, l, m_l, m_s \rangle = -\frac{13.6 \text{ eV}}{n^2} | n, l, m_l, m_s \rangle \]
\[ H_z' | n, l, m_l, m_s \rangle = \frac{\mu_B B}{\hbar} \left( L_z + 2 S_z \right) | n, l, m_l, m_s \rangle = \frac{\mu_B B}{\hbar} L_z | n, l, m_l, m_s \rangle + 2 \frac{\mu_B B}{\hbar} S_z | n, l, m_l, m_s \rangle = \frac{\mu_B B}{\hbar} \left( h m_l | n, l, m_l, m_s \rangle + 2 \frac{\mu_B B}{\hbar} m_s | n, l, m_l, m_s \rangle = \mu_B B (m_l + 2 m_s) | n, l, m_l, m_s \rangle \right) \]

So our zeroth order eigenenergy is
\[ E_{\epsilon,l,m_l,m_s} = -\frac{13.6 \text{ eV}}{n^2} + \mu_B B (m_l + 2 m_s) \]

At \( B = 0 \), we know that energy is independent of \( m_l \) and \( m_s \), i.e., all quantum states are degenerate with (at least \( 2 \times (2 l + 1) \)-fold degenerate). For finite \( B \) however, these states splits.
Example: if we consider states \( n = 2 \) and \( l = 1 \) (first excited states with orbit angular moment quantum number \( l = 1 \)). There, \( m_l = -1, 0, +1 \) and \( m_s = \frac{1}{2} \) or \( -\frac{1}{2} \). At \( B = 0 \), all these six states are degenerate \((E = -13.6/4 = -3.4 \text{ eV})\). In the presence of strong \( B \) field,

\[
E_{n,l,m_l} = \begin{cases} 
-3.4 \text{ eV} + 2 \mu_B B & m_l = +1 \text{ and } m_s = +1/2 \\
-3.4 \text{ eV} + \mu_B B & m_l = 0 \text{ and } m_s = +1/2 \\
-3.4 \text{ eV} & m_l = -1 \text{ and } m_s = +1/2 \quad \text{or} \quad m_l = +1 \text{ and } m_s = -1/2 \\
-3.4 \text{ eV} - \mu_B B & m_l = 0 \text{ and } m_s = -1/2 \\
-3.4 \text{ eV} - 2 \mu_B B & m_l = -1 \text{ and } m_s = -1/2 
\end{cases}
\]

(2.411)

Now, we consider \( H_{SO}^* \) and \( H_{l}^* \). Because we assumed that they are much smaller than \( H_0 \) and \( H_{l} \), we treat them as perturbation and compute the first order correction to the eigenenergy

\[
E_{n,l,m_l,m_s} = \langle n, l, m_s, m_l | H_{l}^* + H_{SO}^* | n, l, m_s, m_l \rangle
\]

(2.412)

The realistic correction is same as what we learned before

\[
\langle n, l, m_s, m_l | H_{l}^* | n, l, m_s, m_l \rangle = -\frac{(E_0^0)^2}{2 m c^2} \left( \frac{4 n}{l + \frac{1}{2}} - 3 \right)
\]

(2.413)

Because

\[
E_0^0 = -\frac{1}{2 n^2} \frac{m}{\hbar^2} \left( \frac{e^2}{4 \pi \varepsilon_0} \right)^2 \approx -\frac{13.6}{n^2} \text{ eV}
\]

(2.414)

\[
\alpha = \frac{e^2}{4 \pi \varepsilon_0 \hbar c} \approx \frac{1}{137.036}
\]

(2.415)

we know that

\[
\frac{E_0^0}{m c^2} = -\frac{1}{2 n^2} \frac{1}{\hbar^2} \left( \frac{e^2}{4 \pi \varepsilon_0} \right)^2 = -\frac{1}{2 n^2} \alpha^2
\]

(2.416)

so

\[
\langle n, l, m_s, m_l | H_{l}^* | n, l, m_s, m_l \rangle = E_0^0 \frac{E_0^0}{2 m c^2} \left( \frac{4 n}{l + \frac{1}{2}} - 3 \right) = -\frac{E_0^0}{2} \frac{\alpha^2}{2 n^2} \left( \frac{4 n}{l + \frac{1}{2}} - 3 \right) = -13.6 \text{ eV} \frac{\alpha^2}{n^2} \left( \frac{n}{l + \frac{1}{2}} - \frac{3}{4} \right)
\]

(2.417)

For spin-orbit coupling,

\[
\langle n, l, m_s, m_l | H_{SO}^* | n, l, m_s, m_l \rangle = \left( n, l, m_s, m_l \right) \hat{S} \cdot \hat{L} \left| n, l, m_s, m_l \right> = \frac{\alpha^2}{8 \pi \varepsilon_0 m^2 c^2 \hbar^3} \left( n, l, m_s, m_l \right) \left( \hat{S} \cdot \hat{L} \right) \left| n, l, m_s, m_l \right>
\]

(2.418)

Notice that in the zeroth order wavefunctions, \( \langle n, l, 1, m_z \rangle \), spin and orbit angular momenta are independent of each other, so

\[
\left( n, l, m_s, m_l \right) \hat{S} \cdot \hat{L} \left| n, l, m_s, m_l \right> = \left( \hat{S} \cdot \hat{L} \right) \left( n, l, m_s, m_l \right) = \left( S_x \right) \left( L_x \right) + \left( S_y \right) \left( L_y \right) + \left( S_z \right) \left( L_z \right)
\]

(2.419)

In QM 1, we learned that for eigenstates of \( L^2 \) and \( L_z \), \( L_x \left( L_x \right) = \left( L_x \right) \). And similarly, \( \left( S_x \right) = \left( S_y \right) = 0 \). And thus

\[
\left( n, l, m_s, m_l \right) \hat{S} \cdot \hat{L} \left| n, l, m_s, m_l \right> = \left( \hat{S} \cdot \hat{L} \right) \left( n, l, m_s, m_l \right) = m_s \hbar = m_s m_l \hbar
\]

(2.420)

As a result,

\[
\langle n, l, m_s, m_l | H_{SO}^* | n, l, m_s, m_l \rangle = \frac{\alpha^2}{8 \pi \varepsilon_0 m^2 c^2 \hbar^3} \left( n, l, m_s, m_l \right) \hat{S} \cdot \hat{L} \left( n, l, m_s, m_l \right) = \frac{\alpha^2}{8 \pi \varepsilon_0 m^2 c^2 \hbar^3} m_s m_l \hbar \left( n, l, m_s, m_l \right) \left( \frac{1}{r^3} \right) \left( n, l, m_s, m_l \right)
\]

(2.421)

and

\[
\langle n, l, s, j_z | \left( \frac{1}{r^3} \right) | n, l, s, j_z \rangle = \int d^3 r \psi_{s,j_za}(r, \theta, \phi) \frac{1}{r^3} \psi_{s,j_za}(r, \theta, \phi) = \frac{1}{l(l + 1/2)(l + 1) n^3 a^3}
\]

(2.422)
where \( a \) is the Bohr radius \( a = \frac{\hbar^2}{m} \frac{4\pi \epsilon_0}{c^2} \)

So

\[
\langle n, l, m_s, m_l \mid H_{SO}' \mid n, l, m_l, m_s \rangle = \frac{e^2}{8\pi \epsilon_0} \frac{m^2 c^2}{\hbar^2} \frac{1}{l(l+1/2)(l+1)n^3} \frac{1}{(l+1/2)(l+1)n^3} = \frac{1}{2} \left( \frac{e^2}{4\pi \epsilon_0 \hbar} \right)^2 \frac{m}{h^2} \frac{m_s m_l}{l(l+1/2)(l+1)n^3}
\]

(2.423)

Because

\[
E_{n}^0 = -\frac{m}{2n^2} \frac{1}{\hbar^2} \left( \frac{e^2}{4\pi \epsilon_0} \right)^2 = -\frac{13.6}{n^2} \text{eV}
\]

(2.424)

\[
\alpha = \frac{e^2}{4\pi \epsilon_0 \hbar c} \approx \frac{1}{137.036}
\]

(2.425)

we can rewrite the formula as

\[
\langle n, l, m_s, m_l \mid H_{SO}' \mid n, l, m_l, m_s \rangle = 13.6 \text{eV} \alpha^2 \frac{m_s m_l}{l(l+1/2)(l+1)n^3}
\]

(2.426)

So our first order correction is

\[
E_{a,l,m_l;m_s} = \langle n, l, m_l, m_s \mid H' \mid n, l, m_l, m_s \rangle = -13.6 \text{eV} \alpha^2 \frac{n}{l(l+1/2)(l+1)n^3} + 13.6 \text{eV} \alpha^2 \frac{m_s m_l}{l(l+1/2)(l+1)n^3}
\]

(2.427)

\[
\frac{13.6 \text{eV}}{n^3} \alpha^2 \frac{3}{4n} - \frac{1}{l(l+1/2)(l+1)} = \frac{13.6 \text{eV}}{n^3} \alpha^2 \frac{3}{4n} - \frac{[(l+1) - m_s m_l]}{l(l+1/2)(l+1)}
\]

So

\[
E_{a,l,m_l;m_s} = E_{a,l,m_l;m_s}^0 + E_{a,l,m_l;m_s}' = -\frac{13.6 \text{eV}}{n^2} + \mu_B \mu_B (m_l + 2 m_s) + \frac{13.6 \text{eV}}{n^3} \alpha^2 \frac{3}{4n} - \frac{[(l+1) - m_s m_l]}{l(l+1/2)(l+1)}
\]

(2.428)

Bottom line: at very strong field (second term much larger than the last one), the energy splittings between the levels are proportional to \( B \) and the slope is proportional to \( \mu_B (m_l + 2m_s) \). The eigenstates are (almost) \( | n, l, m_l, m_s \rangle \), where \( m_s \) and \( m_l \) are good quantum numbers. (we should arrange the states according to the orbit and spin angular moment, NOT the total angular momentum \( j \)).

2.6.3. Weak field

When \( H_z << H_{SO}' \), we can treat \( H_z \) as a small perturbation. The zeroth order Hamiltonian (i.e. ignoring \( H_z \)) is what we studied in the previous section. There, we know that eigenstates are \( | n, j, l, s, m_j \rangle \) and the eigenenergy is

\[
E_{a,j} = E_{a,0} \left[ 1 + \frac{\alpha^2}{n^2} \left( \frac{4n}{j+1/2} - 3 \right) + \ldots \right] = E_{a,0} \left[ 1 + \frac{\alpha^2}{n^2} \left( \frac{n}{j+1/2} - 3 \right) + \ldots \right] = -\frac{13.6 \text{eV}}{n^2} \left[ 1 + \frac{\alpha^2}{n^2} \left( \frac{n}{j+1/2} - 3 \right) + \ldots \right]
\]

(2.429)

In first order perturbation theory, the energy correction is

\[
E_{a,j,l,m_l} = \langle n, j, l, s, m_j \mid H_z \mid n, j, l, s, m_j \rangle = \mu_B \frac{\hbar}{\hbar} \left[ \hat{B} \cdot \hat{l} + 2 \hat{S} \cdot \hat{B} \right] | n, l, s, m_j \rangle
\]

(2.430)

The key is to compute expectation values of \( \hat{L} \) and \( \hat{S} \) for eigenstates of \( \hat{J}^2 \) and \( \hat{J}_z \). Here, we use the fact that

\[
\langle \hat{L} \rangle \parallel \langle \hat{S} \rangle \parallel \langle \hat{J} \rangle
\]

(2.431)

So
\( \langle \vec{L} \rangle = \left\langle \vec{L} \cdot \vec{J} \right\rangle \) \hfill (2.432)

Here,
\[ \langle \vec{B} \cdot \vec{L} \rangle = \left\langle \vec{L} \cdot \vec{J} \frac{B \cdot \vec{J}}{\mathcal{P}} \right\rangle = \left\langle n, j, l, s, m_J \mid \vec{L} \cdot \vec{J} \frac{B \cdot \vec{J}}{\mathcal{P}^2} \mid n, j, l, s, m_J \right\rangle \] \hfill (2.433)

For \( B/z \), we have \( B \cdot \vec{J} = B J_z \)
\[ \langle \vec{B} \cdot \vec{L} \rangle = \left\langle n, j, l, s, m_J \mid \vec{L} \cdot \vec{J} \frac{B J_z}{\mathcal{P}} \right\rangle = \left\langle n, j, l, s, m_J \right\rangle \] \hfill (2.434)

Here, we use the fact that
\[ \mathcal{P} \mid n, j, l, s, m_J \rangle = m_J \hbar \mid n, j, l, s, m_J \rangle \] \hfill (2.435)
\[ \mathcal{P} \mid n, j, l, s, m_J \rangle = j(j + 1) \hbar^2 \mid n, j, l, s, m_J \rangle \] \hfill (2.436)

so
\[ \langle \vec{B} \cdot \vec{L} \rangle = \left\langle n, j, l, s, m_J \mid \vec{L} \cdot \vec{J} \frac{B m_J}{j(j + 1) \hbar} \right\rangle = \frac{B m_J}{j(j + 1) \hbar} \left\langle n, j, l, s, m_J \right\rangle \] \hfill (2.437)

For \( \vec{L} \cdot \vec{J} \), we use the fact that
\[ \vec{S} = \vec{j} - \vec{l} \]
\[ \vec{S} \cdot \vec{S} = (\vec{j} - \vec{l}) \cdot (\vec{j} - \vec{l}) = \vec{j} \cdot \vec{j} + \vec{l} \cdot \vec{l} - 2 \vec{J} \cdot \vec{J} \] \hfill (2.438)

Thus,
\[ \vec{L} \cdot \vec{J} = \frac{\vec{j} \cdot \vec{j} + \vec{l} \cdot \vec{l} - \vec{S} \cdot \vec{S}}{2} \] \hfill (2.439)

So,
\[ \left\langle n, j, l, s, m_j \mid \vec{L} \cdot \vec{J} \right\rangle = \frac{\vec{j} \cdot \vec{j} + \vec{l} \cdot \vec{l} - \vec{S} \cdot \vec{S}}{2} \left\langle n, j, l, s, m_j \right\rangle \]
\[ = \left\langle n, j, l, s, m_j \right\rangle \frac{\hbar^2 (j(j + 1) + l(l + 1) - s^2)}{2} \]
\[ = \frac{\hbar^2}{2} [j(j + 1) + l(l + 1) - s] \] \hfill (2.440)

we know that \( s = 1/2 \) for electrons,
\[ \left\langle n, j, l, s, m_j \mid \vec{L} \cdot \vec{J} \right\rangle = \frac{\hbar^2}{2} [j(j + 1) + l(l + 1) - 3/4] \] \hfill (2.441)

Therefore,
\[ \langle \vec{B} \cdot \vec{L} \rangle = \frac{B m_J}{j(j + 1) \hbar} \left\langle n, j, l, s, m_J \mid \vec{L} \cdot \vec{J} \right\rangle = \frac{B m_J}{j(j + 1) \hbar} \frac{\hbar^2}{2} [j(j + 1) + l(l + 1) - 3/4] = \frac{B m_J}{j(j + 1) \hbar} [j(j + 1) + l(l + 1) - 3/4] \] \hfill (2.442)

Similarly, we can prove that
\[ \langle \vec{B} \cdot \vec{S} \rangle = \frac{B m_J}{j(j + 1) \hbar} \left\langle n, j, l, s, m_J \mid \vec{S} \cdot \vec{J} \right\rangle = \frac{B m_J}{j(j + 1) \hbar} \frac{\hbar^2}{2} [j(j + 1) + l(l + 1) + 3/4] = \frac{B m_J}{j(j + 1) \hbar} [j(j + 1) + l(l + 1) + 3/4] \] \hfill (2.443)
And thus
\[ E_{n,j,l,m_j} = \langle n, j, l, s, m_j | H^* | n, j, l, s, m_j \rangle = \]
\[ = \frac{\mu_B}{h} \left\{ \begin{array}{l}
\frac{B m_j}{2 j(j+1)} [j(j+1) + l(l+1) - 3/4] + 2 \frac{B m_j}{2 j(j+1)} [j(j+1) - l(l+1) + 3/4] \\
\end{array} \right\} = \]
\[ = \frac{\mu_B}{h} \left\{ \begin{array}{l}
B m_j [j(j+1) - l(l+1) + 3/4] = \mu_B B m_j \frac{1}{2} [j(j+1) - l(l+1) + 3/4] \\
\end{array} \right\} = (2.445) \]

We can define an atomic g-factor,
\[ g_j = \frac{1}{2 j(j+1)} [3 j(j+1) - l(l+1) + 3/4] = \frac{3}{2} - \frac{l(l+1)}{2 j(j+1)} = \frac{3}{2} \frac{(l+\frac{1}{2})(l-\frac{1}{2})}{2 j(j+1)} \]
\[ (2.446) \]

If \( j = l + 1/2 \), we find that
\[ g_j = \frac{3}{2} - \frac{(l+\frac{1}{2})(l-\frac{1}{2})}{2 j(j+1)} = \frac{3}{2} - \frac{(l+\frac{1}{2})(l+\frac{1}{2})}{2 j(j+1)} = \frac{3}{2} - \frac{l+\frac{1}{2} - 1/2}{2 j(j+1)} = \frac{3}{2} - \frac{l}{2 j(j+1)} = \frac{3}{2} - \frac{1}{2} \frac{l+1}{2} = 1 - \frac{1}{2(l+1)} \]
\[ (2.447) \]

With this \( g_j \) factor,
\[ E_{n,j,l,m_j} = \mu_B B m_j \]
\[ (2.449) \]

The energy correction (first order), is proportional to \( m_j \) (total angular momentum along the field).

Total energy:
\[ E_{n,j} = -\frac{13.6 \text{ eV}}{n^2} \left[ 1 + \frac{\alpha^2}{n^2} \left( \frac{n}{j+1/2} - \frac{3}{4} \right) \right] + \mu_B B m_j \]
\[ (2.450) \]

### 2.6.4. Intermediate-field

\[ H = H_0 + H^*_0 + H_{SO} + H_z^* \]

When \( H_{SO}^* + H_z^* \), we should treat the last three terms as perturbation. Here, we can treat the problem using degenerate perturbation theory. For \( H_0 \), we consider \( n = 2 \) states (first excited states). In QM I, we learned that there are 4 degenerate states: one s-wave state with \( l = 0 \) and three p-wave states \((l = 1, m_l = -1, 0, 1)\). If we consider spins, there are \( 4 \times 2 = 8 \) states. In first order degenerate perturbation theory, we can ignore all other states except \( n = 2 \) states, and only focus on these 8 states. So the perturbation Hamiltonian is now an 8x8 matrix, which was shown in textbook (page 248). The eigenvalues of this matrix give us the first order corrections in energy.

### 2.7. Summary

**Objective:** compute eigenvalues for the Hamiltonian
\[ \hat{H} | \phi_n \rangle = E_n | \phi_n \rangle \]

**Key assumption:**
\[ \hat{H} = \hat{H}_0 + \lambda \hat{H}^* \]
\[ (2.452) \]

where the second term \( \lambda \hat{H}^* \) is dramatically smaller than the first part.
2.7.1. nondegenerate perturbation theory

Step 1: solve for eigenstates for $\hat{H}_0$
\[
\hat{H}_0 \left| \phi_n^0 \right\rangle = E_n^0 \left| \phi_n^0 \right\rangle
\] (2.454)

If the state that we consider has no degeneracy, we use nondegenerate perturbation theory

Step 2: first order correction
\[
E_n^1 = \langle \phi_n^0 \left| \hat{H}' \right| \phi_n^0 \rangle
\] (2.455)

Step 3: second order correction
\[
E_n^2 = \sum_{\mu > n} \langle \phi_n^0 \left| \hat{H}' \right| \phi_\mu^0 \rangle \frac{1}{E_n^0 - E_\mu^0} \langle \phi_\mu^0 \left| \hat{H}' \right| \phi_n^0 \rangle = \sum_{\mu > n} \left| \langle \phi_n^0 \left| \hat{H}' \right| \phi_\mu^0 \rangle \right|^2 \frac{E_n^0 - E_\mu^0}{E_n^0 - E_\mu^0}
\] (2.456)

Step 4: eigenenergy
\[
E_n = E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \ldots
\] (2.457)

Wave functions:
\[
\left| \psi_n \right\rangle = \left| \phi_n^0 \right\rangle + \lambda \sum_{\mu > n} \left| \phi_\mu^0 \right\rangle \frac{1}{E_n^0 - E_\mu^0} \langle \phi_\mu^0 \left| \hat{H}' \right| \phi_n^0 \rangle + \ldots
\] (2.458)

2.7.2. degenerate perturbation theory

Step 1: solve for eigenstates for $\hat{H}_0$
\[
\hat{H}_0 \left| \phi_n^0 \right\rangle = E_n^0 \left| \phi_n^0 \right\rangle
\] (2.459)

If the state that we consider has degeneracy (i.e. there is at least one other state has the same eigenenergy), we use degenerate perturbation theory:
\[
\hat{H}_0 \left| \psi_n^0 \right\rangle = E_n^0 \left| \psi_n^0 \right\rangle
\] (2.460)

and
\[
\hat{H}_0 \left| \psi_n^0 \right\rangle = E_n^0 \left| \psi_n^0 \right\rangle
\] (2.461)

Step 1: Create a $n \times n$ matrix (if there is an $n$-fold degeneracy),
\[
W = \begin{pmatrix}
\langle \phi_n^0 \left| \hat{H}' \right| \phi_n^0 \rangle & \langle \phi_n^0 \left| \hat{H}' \right| \phi_\mu^0 \rangle \\
\langle \phi_\mu^0 \left| \hat{H}' \right| \phi_n^0 \rangle & \langle \phi_\mu^0 \left| \hat{H}' \right| \phi_\mu^0 \rangle
\end{pmatrix}
\] (2.462)

Step 2: The eigenvalues of the matrix is the first order correction
\[
E_1 = E_n^0 + \lambda E_n + O(\lambda^2)
\] (2.463)
\[
E_2 = E_n^0 + \lambda E_n + O(\lambda^2)
\] (2.464)

Wavefunctions: eigenvectors
\[
\begin{pmatrix}
W_{aa} & W_{ab} \\
W_{ba} & W_{bb}
\end{pmatrix}
\begin{pmatrix}
\alpha_1 \\
\beta_1
\end{pmatrix} = E_n^0 \begin{pmatrix}
\alpha_1 \\
\beta_1
\end{pmatrix}
\] (2.465)

and
\[
\begin{pmatrix}
W_{aa} & W_{ab} \\
W_{ba} & W_{bb}
\end{pmatrix}
\begin{pmatrix}
\alpha_2 \\
\beta_2
\end{pmatrix} = E_n^0 \begin{pmatrix}
\alpha_2 \\
\beta_2
\end{pmatrix}
\] (2.466)

where $E_+$ and $E_-$ are the two eigenvalues.
\[
\left| \psi_1^0 \right\rangle = \alpha_1 \left| \phi_n^0 \right\rangle + \beta_1 \left| \phi_n^0 \right\rangle
\] (2.467)
\[ |\psi_0^0\rangle = \alpha_2 |\psi_0^0\rangle + \beta_2 |\psi_0^0\rangle \]  

(2.468)

2.7.3. Perturbation theory in matrix formula (example: homework 2.3)

This case usually uses eigenstates of \( H_0 \) as basis,
\[ \hat{H}_0 |\psi_n^0\rangle = E_n^0 |\psi_n^0\rangle \]  

(2.469)

With one complete set of basis, we can write an operator as a matrix
\[ (H_0)_{mn} = \langle \psi_0^m | \hat{H}_0 | \psi_0^n \rangle \]  

(2.470)

If we use eigenstates of \( H_0 \) as basis, the matrix is diagonal and the diagonal components are eigenvalues of the \( H_0 \)
\[ H_0 \rightarrow \begin{pmatrix} E_1^0 & 0 & 0 & \ldots \\ 0 & E_2^0 & 0 & \ldots \\ 0 & 0 & E_3^0 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \]  

(2.471)

Using the same basis, we can write \( H' \) as a matrix
\[ \lambda H' \rightarrow \lambda \begin{pmatrix} \langle \psi_1 | H' | \psi_1 \rangle & \langle \psi_1 | H' | \psi_2 \rangle & \langle \psi_1 | H' | \psi_3 \rangle & \ldots \\ \langle \psi_2 | H' | \psi_1 \rangle & \langle \psi_2 | H' | \psi_2 \rangle & \langle \psi_2 | H' | \psi_3 \rangle & \ldots \\ \langle \psi_3 | H' | \psi_1 \rangle & \langle \psi_3 | H' | \psi_2 \rangle & \langle \psi_3 | H' | \psi_3 \rangle & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \]  

(2.472)

In many cases, we only need keep a small number of states (e.g. only the tree states with lowest energy)
\[ H_0 = \begin{pmatrix} E_1^0 & 0 & 0 \\ 0 & E_2^0 & 0 \\ 0 & 0 & E_3^0 \end{pmatrix} \]  

(2.473)

and
\[ \lambda H' = \lambda \begin{pmatrix} \langle \psi_1 | H' | \psi_1 \rangle & \langle \psi_1 | H' | \psi_2 \rangle & \langle \psi_1 | H' | \psi_3 \rangle \\ \langle \psi_2 | H' | \psi_1 \rangle & \langle \psi_2 | H' | \psi_2 \rangle & \langle \psi_2 | H' | \psi_3 \rangle \\ \langle \psi_3 | H' | \psi_1 \rangle & \langle \psi_3 | H' | \psi_2 \rangle & \langle \psi_3 | H' | \psi_3 \rangle \end{pmatrix} \]  

(2.474)

Objective: compute eigenstates for the \( H \) matrix
\[ H = \begin{pmatrix} E_1^0 & 0 & 0 \\ 0 & E_2^0 & 0 \\ 0 & 0 & E_3^0 \end{pmatrix} + \lambda \begin{pmatrix} \langle \psi_1 | H' | \psi_1 \rangle & \langle \psi_1 | H' | \psi_2 \rangle & \langle \psi_1 | H' | \psi_3 \rangle \\ \langle \psi_2 | H' | \psi_1 \rangle & \langle \psi_2 | H' | \psi_2 \rangle & \langle \psi_2 | H' | \psi_3 \rangle \\ \langle \psi_3 | H' | \psi_1 \rangle & \langle \psi_3 | H' | \psi_2 \rangle & \langle \psi_3 | H' | \psi_3 \rangle \end{pmatrix} \]  

(2.475)

Assumption: the second matrix is much smaller than the first

Nondegenerate perturbation:
Among the three unperturbed eigenvalues, \( E_1^0 \), \( E_2^0 \) and \( E_3^0 \), if one of them is different from the other two (e.g. \( E_1^0 \) is different), then we can use nondegenerate perturbation theory.

First order perturbation:
\[ E_1^1 = \langle \psi_1 | H' | \psi_1 \rangle \]  

(2.476)

Notice that it is just one element in the second matrix in \( H \).

Second order perturbation:
\[ E_1^2 = \sum_{m=0} \frac{\langle \psi_0^0 | H' | \psi_m^0 \rangle}{E_1^0 - E_m^0} \frac{1}{\langle \psi_0^0 | H' | \psi_0^0 \rangle} \langle \psi_0^0 | H' | \psi_0^0 \rangle = \frac{\langle \psi_1^0 | H' | \psi_2^0 \rangle \langle \psi_2^0 | H' | \psi_1^0 \rangle}{E_1^0 - E_2^0} + \frac{\langle \psi_1^0 | H' | \psi_3^0 \rangle \langle \psi_3^0 | H' | \psi_1^0 \rangle}{E_1^0 - E_3^0} \]  

(2.477)

Notice that the denominator are just elements from the first matrix in \( H \) and the numerators are from the second matrix.

Degenerate perturbation:
Among the three unperturbed eigenvalues, \( E_1^0 \), \( E_2^0 \) and \( E_3^0 \), if two (or more) of them are identical (e.g. \( E_2^0 \) and \( E_3^0 \) have the same value, then we can use degenerate perturbation theory.
Step one: create the $W$ matrix

$$W = \begin{pmatrix} \langle \psi_2^0 | H^* | \psi_2^0 \rangle & \langle \psi_2^0 | H^* | \psi_3^0 \rangle \\ \langle \psi_3^0 | H^* | \psi_2^0 \rangle & \langle \psi_3^0 | H^* | \psi_3^0 \rangle \end{pmatrix}$$

(2.478)

Step two: compute eigenvalues of $W$, which are the first order correction
Variational Method

3.1. Basic idea

If we are trying to find the ground-state energy for a quantum system, we can utilize the following fact: the ground state has the lowest possible energy for the Hamiltonian (by definition).

\[ E_g \leq \langle \psi | H | \psi \rangle = \langle H \rangle \]

Therefore, by finding the wavefunction \( |\psi\rangle \) that minimizes \( \langle \psi | H | \psi \rangle \), we find the ground state and the expectational value \( \langle \psi | H | \psi \rangle \) is the ground-state energy.

In principle, here we need to compare all possible wavefunctions and to find the smallest possible value for \( \langle \psi | H | \psi \rangle \).

In reality, for a complex Hamiltonian, finding the true minimum is often very challenging (even numerically).

In practice, we do the second best thing. Instead of considering all possible wavefunctions, we first make a "guess" and only consider a small family of wavefunctions. Then among these wavefunctions, we find the one that minimizes \( \langle \psi | H | \psi \rangle \). This method is known as the "variational method"

3.2. Comments:

1. Exact vs. approximate
   - If one compare all possible wavefunctions and find the one that minimizes \( \langle \psi | H | \psi \rangle \), then this is an exact solution (i.e. one finds the exact ground state).
   - However, in reality, most of the cases, we cannot find exact solution, so we only compared some of the wavefunctions (not all possible ones). This is an approximation, and thus the results we found would be “approximate” ones, instead of exact ones.

2. Overestimate \( E_g \)
   - The variational method always overestimate the ground-state energy (i.e. the approximation ground state energy that we get in this method is always larger, or equal, to the true ground state energy). In other words, the method produces an upper bound for the ground-state energy. By considering more and more wavefunctions in the minimization procedure, this upper bound is getting closer and closer to the global minimum, which is the real \( E_g \).

3. How accurate is this approximate method?
   - It depends on whether we made a “good guess” or not about the ground-state wavefunction. A “good guess” here means whether the wavefunctions that we considered here is “close” enough to the true ground state or not.
   - How to make a “good guess”? There is no general rule here, and it relies on good physics intuitions (see examples below).

4. Ground state energy vs ground state wavefunctions
The variational method usually produces good estimation for ground-state energy. But the ground-state wavefunctions obtained in the same calculation is often much less accurate.

In other words, the (approximate) ground-state energy that we get here is often not so bad. But the (approximate) ground state wavefunction is often not close to the real one at all.

Proof: Assume that the true ground state is $|\psi_0\rangle$. However, we didn’t guess the exact ground state in our variational method calculations. Instead, we got a wavefunction that slightly off. $|\psi\rangle = |\psi_0\rangle + \lambda |\psi'\rangle$ where $\lambda$ is a small quantity, and $|\psi'\rangle$ is superpositions of excited states that are orthogonal to $|\psi_0\rangle$, i.e., $\langle\psi' | \psi_0\rangle = 0$. For this "approximate" wavefunction, the approximate ground state energy we get is

$$E_g' = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} = \left( E_g + \lambda^2 \langle \psi' | H | \psi' \rangle \right) \left( 1 - \lambda^2 \langle \psi' | \psi' \rangle + \ldots \right) = E_g + \lambda^2 \langle \psi' | H | \psi' \rangle - E_g \langle \psi' | \psi' \rangle + \ldots$$

(3.2)

(3.3)

The error bar for the ground state energy is proportional to $\lambda^2$

$$|E_g' - E_g| = \lambda^2 \langle \psi' | H | \psi' \rangle - E_g \langle \psi' | \psi' \rangle + \ldots$$

(3.4)

The error bar for the ground state wavefunction is proportional to $\lambda$

$$|\psi' - |\psi_0\rangle| = \lambda \langle \psi' | H | \psi' \rangle - E_g \langle \psi' | \psi' \rangle + \ldots$$

(3.5)

For small $\lambda$, $\lambda^2 << \lambda$. And thus, the ground state energy has much better accuracy (smaller error bar) than the ground state wavefunction.

5. Another important application for the variational principle:

- Assumes that we have two different (approximate) methods to compute the ground state wavefunctions, and we get two answers $|\psi_1\rangle$ and $|\psi_2\rangle$. Which one of the is better (more accurate)?

- One easy way to determine is to use the variational principle and computes to quantities

$$E_{g1} = \langle \psi_1 | H | \psi_1 \rangle$$

(3.6)

and

$$E_{g2} = \langle \psi_2 | H | \psi_2 \rangle$$

(3.7)

If $E_{g1} < E_{g2}$, the first method is (presumably) better. If $E_{g1} > E_{g2}$, then the second method is probability better.

### 3.3. Example: Helium atom

Consider a helium atom, which has a nucleon with charge $+2e$ and two electrons. The wavefunction is $\psi(\vec{r}_1, \vec{r}_2)$, where $\vec{r}_1$ and $\vec{r}_2$ are the coordinates of the two electrons respectively. Hamiltonian is

$$H = \frac{-\hbar^2}{2m} \left( \nabla_1^2 + \nabla_2^2 \right) - \frac{2e^2}{4\pi\epsilon_0} \frac{1}{r_1} - \frac{2e^2}{4\pi\epsilon_0} \frac{1}{r_2} + \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\vec{r}_1 - \vec{r}_2|}$$

(3.8)

Here, the first term is the kinetic energy of the two electrons. The last three terms are the Coulomb interactions between the nucleon and electron #1, nucleon and electron #2, and between the two electrons respectively.

Experiments told us that the ground state energy of an helium atom is

$$E_g = -78.975 \text{ eV}$$

(3.9)

Now, let’s try to find the ground state energy using variational methods:
3.3.1. First guess: Maybe the ground-state wavefunction is very similar to a hydrogen atom?

If there is no interactions between the two electrons, the Hamiltonian can be separate into two parts

\[ H_{\text{no-ee-interactions}} = \left( -\frac{\hbar^2}{2m} \nabla^2 - \frac{2e^2}{4\pi \epsilon_0} \frac{1}{r_1} \right) + \left( -\frac{\hbar^2}{2m} \nabla^2 - \frac{2e^2}{4\pi \epsilon_0} \frac{1}{r_2} \right) \]  

(3.10)

In other words, the system becomes two (decoupled) hydrogen-atom problems. For electron #1, its Hamiltonian is

\[ H_1 = -\frac{\hbar^2}{2m} \nabla^2 - \frac{2e^2}{4\pi \epsilon_0} \frac{1}{r_1} \]  

(3.11)

and for electron #2, it is

\[ H_2 = -\frac{\hbar^2}{2m} \nabla^2 - \frac{2e^2}{4\pi \epsilon_0} \frac{1}{r_2} \]  

(3.12)

Each of these two Hamiltonian is a hydrogen atom Hamiltonian, but the nucleon charge is now doubled. Remember, the typical hydrogen atom Hamiltonian looks like

\[ H_{\text{hydrogen}} = -\frac{\hbar^2}{2m} \nabla^2 - \frac{e^2}{4\pi \epsilon_0} \frac{1}{r} \]  

(3.13)

The second term has \( e^2 \) in the numerator, but there it is \( 2e^2 \), because the nucleon of a helium atom has charge +2e.

For a hydrogen atom with nucleon charge \( +e \), the ground state wavefunction is

\[ \psi_{100}(r) = \frac{1}{\sqrt{\pi a^3}} e^{-r/a} \]  

(3.14)

where \( a \) is the Bohr radius,

\[ a = \frac{4\pi \epsilon_0 \hbar^2}{m e^2} = 0.529 \times 10^{-10} \text{ m} \]  

(3.15)

The ground state energy is

\[ E_1 = -\frac{m}{2\hbar^2} \left( \frac{e^2}{4\pi \epsilon_0} \right)^2 = -13.6 \text{ eV} \]  

(3.16)

Now, for a hydrogen like atom with nucleon charge \( Z \times e \) (and one electron), the Hamiltonian is

\[ H_{\text{charge} \ Z} = -\frac{\hbar^2}{2m} \nabla^2 - \frac{Ze^2}{4\pi \epsilon_0} \frac{1}{r} \]  

(3.17)

If we compare it with a real hydrogen problem, the only thing we did here is to change \( e^2 \) to \( Ze^2 \). And thus for this charge Ze hydrogen like atom, the solution would be almost the same as an ordinary hydrogen atom, except that we need to change all \( e^2 \) into \( Ze^2 \). As a result, our Bohr radius now becomes

\[ a^* = \frac{4\pi \epsilon_0 \hbar^2}{Z m e^2} = \frac{a}{Z} = \frac{0.529 \times 10^{-10} \text{ m}}{Z} \]  

(3.18)

And thus the ground state wavefunction is now

\[ \psi_{100}(r) = \frac{1}{\sqrt{\pi a^*^3}} e^{-r/a^*} = \frac{1}{\sqrt{\pi a^3 / Z^3}} e^{-Z^{2/3} r / a^*} \]  

(3.19)

The ground state energy is

\[ E_1^* = -\frac{m}{2\hbar^2} \left( \frac{Ze^2}{4\pi \epsilon_0} \right)^2 = Z^2 \left[ -\frac{m}{2\hbar^2} \left( \frac{e^2}{4\pi \epsilon_0} \right)^2 \right] = Z^2 E_1 = -13.6 Z^2 \text{ eV} \]  

(3.20)

If we ignore the interactions between electrons, the ground state wavefunction would be
\( \psi_0(r_1, r_2) = \psi_{100}(r_1) \psi_{100}(r_2) \)

where

\[
\psi_{100}(r) = \frac{1}{\sqrt{\pi a^3/Z^2}} \ e^{-2r/a} = \frac{1}{\sqrt{\pi a^3/2^3}} \ e^{-2r/\sqrt{2}}
\]

(3.22)

In other words,

\[
\psi_0(r_1, r_2) = \psi_{100}(r_1) \psi_{100}(r_2) = \frac{8}{\pi a^3} e^{-2(r_1+r_2)/\sqrt{2}}
\]

(3.23)

This wavefunction is NOT the ground state of a helium atom, because we ignored the interaction between electrons. But let’s use it as a guess wavefunction for the ground state (which is an approximation) and see what is the energy we get.

\[
(H) = \langle \psi_0(r_1, r_2) | H | \psi_0(r_1, r_2) \rangle
\]

(3.24)

Here, we first rewrite the Hamiltonian as three parts:

\[
H = H_1 + H_2 + H_{ee}
\]

(3.25)

where

\[
H_1 = -\frac{\hbar^2}{2m} \nabla_1^2 - \frac{2e^2}{4\pi \epsilon_0} \frac{1}{r_1}
\]

(3.26)

and

\[
H_2 = -\frac{\hbar^2}{2m} \nabla_2^2 - \frac{2e^2}{4\pi \epsilon_0} \frac{1}{r_2}
\]

(3.27)

and

\[
H_{ee} = e^2 \frac{1}{4\pi \epsilon_0} \frac{1}{|\vec{r}_1 - \vec{r}_2|}
\]

(3.28)

As a result,

\[
(H) = \langle \psi_0(r_1, r_2) | H | \psi_0(r_1, r_2) \rangle = \langle \psi_0(r_1, r_2) | H_1 + H_2 + H_{ee} | \psi_0(r_1, r_2) \rangle
\]

\[
= \langle \psi_0(r_1, r_2) | H_1 | \psi_0(r_1, r_2) \rangle + \langle \psi_0(r_1, r_2) | H_2 | \psi_0(r_1, r_2) \rangle + \langle \psi_0(r_1, r_2) | H_{ee} | \psi_0(r_1, r_2) \rangle
\]

(3.29)

The first term is very simply

\[
\langle \psi_0(r_1, r_2) | H_1 | \psi_0(r_1, r_2) \rangle = \int d\vec{r}_1 \int d\vec{r}_2 \psi_{100}^*(r_1) \psi_{100}^*(r_2) H_1 \psi_{100}(r_1) \psi_{100}(r_2)
\]

(3.30)

Because, \( H_1 \) only depends on \( r_1 \) and is independent of \( r_2 \), we know that

\[
\langle \psi_0(r_1, r_2) | H_1 | \psi_0(r_1, r_2) \rangle = \int d\vec{r}_1 \int d\vec{r}_2 \psi_{100}^*(r_1) \psi_{100}^*(r_2) H_1 \psi_{100}(r_1) \psi_{100}(r_2)
\]

\[
= \int d\vec{r}_1 \psi_{100}^*(r_1) H_1 \psi_{100}(r_1) \int d\vec{r}_2 \psi_{100}^*(r_2) \psi_{100}(r_2)
\]

(3.31)

The second integral is identity, because it is the normalization condition of a wavefunction

\[
\int d\vec{r}_2 \psi_{100}^*(r_2) \psi_{100}(r_2) = 1
\]

(3.32)

The first integral is also easily to compute, because \( \psi_{100}(r_1) \) is the ground state of \( H_1 \), and thus the integral is the ground state energy for \( H_1 \), which we already know

\[
E_{100}^* = -13.6 \ Z eV = -13.6 \ eV + 2e^2 = -54.4 \ eV
\]

(3.33)

So we find that

\[
\langle \psi_0(r_1, r_2) | H_1 | \psi_0(r_1, r_2) \rangle = -54.4 \ eV
\]

(3.34)
Similarly, we can show that
\[
\langle \psi_0(r_1, r_2) | H_2 | \psi_0(r_1, r_2) \rangle = -54.4 \text{ eV}
\] (3.35)

So,
\[
(H) = \langle \psi_0(r_1, r_2) | H | \psi_0(r_1, r_2) \rangle \\
= \langle \psi_0(r_1, r_2) | H_1 | \psi_0(r_1, r_2) \rangle + \langle \psi_0(r_1, r_2) | H_2 | \psi_0(r_1, r_2) \rangle + \langle \psi_0(r_1, r_2) | H_{ac} | \psi_0(r_1, r_2) \rangle \\
= -54.4 \text{ eV} - 54.4 \text{ eV} + \langle \psi_0(r_1, r_2) | H_{ac} | \psi_0(r_1, r_2) \rangle \\
\approx -109 \text{ eV} + \langle \psi_0(r_1, r_2) | H_{ac} | \psi_0(r_1, r_2) \rangle
\] (3.36)
To compute the last term, we evaluate the following integral
\[
\langle \psi_0(r_1, r_2) | H_{ac} | \psi_0(r_1, r_2) \rangle = \int d\vec{r}_1 \int d\vec{r}_2 \psi_0^*(r_1, r_2) H_{ac} \psi_0(r_1, r_2)
\]
\[
= \int d\vec{r}_1 \int d\vec{r}_2 \frac{8}{\pi a^3} e^{-\frac{r_1^2}{4a^2}} \frac{1}{|\vec{r}_1 - \vec{r}_2|} \frac{8}{\pi a^3} e^{-\frac{r_2^2}{4a^2}}
\]
\[
= \frac{64}{\pi^2 a^6} \int d\vec{r}_1 \int d\vec{r}_2 \frac{e^{-\frac{r_1^2}{4a^2}} e^{-\frac{r_2^2}{4a^2}}}{|\vec{r}_1 - \vec{r}_2|}
\] (3.37)
We first do the integral for \( r_2 \),
\[
\int d\vec{r}_2 \frac{e^{-\frac{r_1^2}{4a^2}} e^{-\frac{r_2^2}{4a^2}}}{|\vec{r}_1 - \vec{r}_2|}
\]
(3.38)
For this integral, without loss of generality, we can assume \( \vec{r}_1 \) is along the \( z \) direction, and thus we can convert the integral into spherical coordinates
\[
\int d\vec{r}_2 \frac{e^{-\frac{r_1^2}{4a^2}} e^{-\frac{r_2^2}{4a^2}}}{|\vec{r}_1 - \vec{r}_2|} = \int_0^{\pi} \int_0^{\infty} r_2^2 \sin \theta d\theta \int_0^{2\pi} d\phi \frac{e^{-\frac{r_1^2}{4a^2}} e^{-\frac{r_2^2}{4a^2}}}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta}}
\] (3.39)
Here, \( \int_0^{2\pi} d\phi = 2\pi \), because nothing depends on \( \phi \) in the integral.
\[
\int d\vec{r}_2 \frac{e^{-\frac{r_1^2}{4a^2}} e^{-\frac{r_2^2}{4a^2}}}{|\vec{r}_1 - \vec{r}_2|} = \int_0^{\infty} \int_0^{\pi} r_2^2 \sin \theta d\theta \int_0^{2\pi} d\phi \frac{2\pi e^{-\frac{r_1^2}{4a^2}} e^{-\frac{r_2^2}{4a^2}}}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta}}
\] (3.40)
For the \( \theta \) integral, we can define \( x = \cos \theta \), as a result,
\[
dx = d\cos \theta = -\sin \theta d\theta
\] (3.41)
and thus
\[
\int_0^{\pi} \sin \theta d\theta = -\int_{\cos \pi}^{\cos 0} dx = -\int_{-1}^{1} dx = \int_{-1}^{1} dx
\] (3.42)
And therefore,
\[
\int d\vec{r}_2 \frac{e^{-\frac{r_1^2}{4a^2}} e^{-\frac{r_2^2}{4a^2}}}{|\vec{r}_1 - \vec{r}_2|} = \int_0^{\infty} \int_0^{\pi} \int_{-1}^{1} dx \frac{2\pi e^{-\frac{r_1^2}{4a^2}} e^{-\frac{r_2^2}{4a^2}}}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 x}} = \int_0^{\infty} 2\pi \frac{e^{-\frac{r_1^2}{4a^2}} e^{-\frac{r_2^2}{4a^2}}}{\sqrt{2r_1 r_2}} dr_2 \int_{-1}^{1} \frac{dx}{\sqrt{\frac{r_1^2 r_2^2}{2r_1 r_2} - x}}
\] (3.43)
We know that
\[
\int dx \frac{1}{\sqrt{a-x}} = -2 \sqrt{a-x}
\] (3.44)
So
\[
\int d\vec{r}_2 \frac{e^{-\frac{r_1^2}{4a^2}} e^{-\frac{r_2^2}{4a^2}}}{|\vec{r}_1 - \vec{r}_2|} = 2 \sqrt{a - \frac{r_1^2 r_2^2}{2r_1 r_2}}
\]
\[
\int_{r_1}^{r_2} \frac{dx}{\sqrt{\frac{r_1^2 + r_2^2}{2 r_1 r_2} - x}} = -2 \sqrt{\frac{r_1^2 + r_2^2}{2 r_1 r_2} - x} \bigg|_{x=x_0}^{x=x_1} \bigg| = -2 \sqrt{\frac{r_1^2 + r_2^2}{2 r_1 r_2} - 1} + 2 \sqrt{\frac{r_1^2 + r_2^2}{2 r_1 r_2} + 1}
\]

(3.45)

and thus

\[
\int d^2 r \ e^{-\frac{e}{2}(r_1^2 + r_2^2)} \bigg| \frac{1}{r_1 - r_2} \bigg| \ = \int_0^{\infty} \frac{2 \pi e^{-\frac{e}{2}(r_1^2 + r_2^2)}}{2 r_1 r_2} r_2^2 dr_2 \int_{r_1}^{r_2} \frac{dx}{\sqrt{\frac{r_1^2 + r_2^2}{2 r_1 r_2} - x}}
\]

\[
= \int_0^{\infty} \frac{4 \pi e^{-\frac{e}{2}(r_1^2 + r_2^2)}}{2 r_1 r_2} \left( \frac{r_1^2 + r_2^2}{2 r_1 r_2} + 1 - \frac{r_1^2 + r_2^2}{2 r_1 r_2} - 1 \right) r_2^2 dr_2
\]

\[
(3.46)
\]

If \( r_1 > r_2, r_1 + r_2 = |r_1 - r_2| = r_1 + r_2 - r_1 - r_2 = 2 r_2 \)

If \( r_2 > r_1, r_1 + r_2 - |r_1 - r_2| = r_1 + r_2 + r_1 - r_2 = 2 r_1 \)

\[
\int d^2 r \ e^{-\frac{e}{2}(r_1^2 + r_2^2)} \bigg| \frac{1}{r_1 - r_2} \bigg| \ = \int_0^{\infty} \frac{2 \pi e^{-\frac{e}{2}(r_1^2 + r_2^2)}}{r_1 r_2} (r_1 + r_2 - |r_1 - r_2|) r_2^2
\]

\[
= \int_0^{\infty} \frac{2 \pi e^{-\frac{e}{2}(r_1^2 + r_2^2)}}{r_1 r_2} (r_1 + r_2 - |r_1 - r_2|) r_2^2 + \int_0^{\infty} \frac{2 \pi e^{-\frac{e}{2}(r_1^2 + r_2^2)}}{r_1 r_2} (r_1 + r_2 - |r_1 - r_2|) r_2^2
\]

\[
= \int_0^{\infty} \frac{2 \pi e^{-\frac{e}{2}(r_1^2 + r_2^2)}}{r_1 r_2} (2 r_2)^2 + \int_0^{\infty} \frac{2 \pi e^{-\frac{e}{2}(r_1^2 + r_2^2)}}{r_1 r_2} (2 r_1)^2
\]

\[
= \int_0^{\infty} \frac{4 \pi e^{-\frac{e}{2}(r_1^2 + r_2^2)}}{r_1} r_2^2 + \int_0^{\infty} \frac{4 \pi e^{-\frac{e}{2}(r_1^2 + r_2^2)}}{r_2} r_2^2
\]

\[
= \frac{4 \pi e^{-\frac{\psi}{4}}}{r_1} \int_0^{\infty} dr_2 e^{\frac{\psi}{4} r_2^2} + 4 \pi e^{-\frac{\psi}{4}} \int_0^{\infty} dr_2 4 \pi e^{-\frac{\psi}{4} r_2^2}
\]

\[
= \frac{4 \pi e^{-\frac{\psi}{4}}}{r_1} \left[ \frac{a^3 - \frac{1}{32} a e^{-\frac{\psi}{4}} (a^2 + 4 a r_1 + 8 r_1^2)}{16} \right] + 4 \pi e^{-\frac{\psi}{4}} \left[ \frac{a^3 e^{-\frac{\psi}{4}} (a + 4 r_1)}{16} \right]
\]

\[
= \frac{\pi a^3 e^{-\frac{\psi}{4}}}{8 r_1} \left[ 1 - e^{-\frac{\psi}{4}} \left( 1 + \frac{2 r_1}{a} \right) \right]
\]

Now, we go back to \( \langle \phi_0(r_1, r_2) | H_{\text{loc}} | \phi_0(r_1, r_2) \rangle \)

\[
\langle \phi_0(r_1, r_2) | H_{\text{loc}} | \phi_0(r_1, r_2) \rangle = \frac{64}{\pi^2 a^6} \frac{e^2}{4 \pi \epsilon_0} \int d\vec{r}_1 \int d\vec{r}_2 \ e^{-\frac{\psi}{4}(r_1^2 + r_2^2)} \bigg| \frac{1}{r_1 - r_2} \bigg|
\]

\[
= \frac{64}{\pi^2 a^6} \frac{e^2}{4 \pi \epsilon_0} \int d\vec{r}_1 \pi a^3 e^{-\frac{\psi}{4} r_1^2} \left[ 1 - e^{-\frac{\psi}{4}} \left( 1 + \frac{2 r_1}{a} \right) \right]
\]

\[
= \frac{8}{\pi a^3} \frac{e^2}{4 \pi \epsilon_0} \int r_1^2 dr_1 \int d\Omega \ e^{-\frac{\psi}{4} r_1^2} \left[ 1 - e^{-\frac{\psi}{4}} \left( 1 + \frac{2 r_1}{a} \right) \right]
\]

(3.48)
In the previous section, we assumed that the wavefunction to be

\[ \psi(r_1, r_2) = \psi_{100}(r_1) \psi_{100}(r_2) \]  

where

\[ \psi_{100}(r) = \frac{1}{\sqrt{\pi a^3 / Z^2}} e^{-Zr/a} \]  

We specifically assumed that the wavefunction \( \psi_{100}(r) \) are ground state of a hydrogen like atom with \( Z = 2 \). Now, let’s consider \( Z \) as a adjustable parameter. And try to find which value of \( Z \) give us the lowest energy. Now our wavefunction is

\[ \psi(r_1, r_2) = \frac{Z^3}{\pi a^3 e^{-Zr/a}} \]  

This structure of wavefunctions include the wavefunctions that we used above, and in addition, we adds other possible wavefunctions into our consideration, and thus the results must be better than what we did above.

Now, we rewrite the Hamiltonian into three terms:

\[ H = \left( -\frac{\hbar^2}{2m} \nabla^2 - \frac{Z e^2}{4\pi \epsilon_0 r_1} \right) + \left( -\frac{\hbar^2}{2m} \nabla^2 - \frac{Z e^2}{4\pi \epsilon_0 r_2} \right) + \left( \frac{e^2}{4\pi \epsilon_0} \right) \left( \frac{1}{\tilde{r}_1} - \frac{1}{\tilde{r}_2} \right) \]  

The first two terms are both the Hamiltonian of a hydrogen-like atom whose nucleon has change \( Z \times e \). Because our wavefunction is the ground state of such a Hamiltonian,

\[ \langle \psi | -\frac{\hbar^2}{2m} \nabla^2 - \frac{Z e^2}{4\pi \epsilon_0 r_1} | \psi \rangle = \langle \psi | -\frac{\hbar^2}{2m} \nabla^2 - \frac{Z e^2}{4\pi \epsilon_0 r_2} | \psi \rangle = E_1 = -13.6 Z^2 \text{eV} \]  

As a result,

\[ \langle \psi | H | \psi \rangle = \langle \psi | -\frac{\hbar^2}{2m} \nabla^2 - \frac{Z e^2}{4\pi \epsilon_0 r_1} | \psi \rangle + \langle \psi | -\frac{\hbar^2}{2m} \nabla^2 - \frac{Z e^2}{4\pi \epsilon_0 r_2} | \psi \rangle + \langle \psi | (Z - 2) \epsilon^2 \frac{1}{4\pi \epsilon_0 r_1} | \psi \rangle + \langle \psi | (Z - 2) \epsilon^2 \frac{1}{4\pi \epsilon_0 r_2} | \psi \rangle \]  

Notice that \( E_1 \) is slightly larger than the real value. And try to find which value of \( Z \) give us the lowest energy. Now our wavefunction is

\[ \psi(r_1, r_2) = \frac{Z^3}{\pi a^3 e^{-Zr/a}} \]  

This structure of wavefunctions include the wavefunctions that we used above, and in addition, we adds other possible wavefunctions into our consideration, and thus the results must be better than what we did above.
\[
= -27.2 Z^2 \text{ eV} + \psi 1 \frac{(Z-2) e^2}{4 \pi \epsilon_0 r_1} + \frac{(Z-2) e^2}{4 \pi \epsilon_0 r_2} + \frac{e^2}{4 \pi \epsilon_0} \frac{1}{|\vec{r}_1 - \vec{r}_2|} \psi
\]

\[
= -27.2 Z^2 \text{ eV} + \psi 1 \frac{(Z-2) e^2}{4 \pi \epsilon_0} \left( \psi 1 \frac{1}{r_1} \right) + \frac{(Z-2) e^2}{4 \pi \epsilon_0} \left( \psi 1 \frac{1}{r_2} \right) + \frac{e^2}{4 \pi \epsilon_0} \left( \psi 1 \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right)
\]

Here,
\[
\left( \psi 1 \frac{1}{r_1} \right) = \int dr_1^1 \int dr_2^2 \psi_0^*(r_1, r_2) \frac{1}{r_1} \psi_0(r_1, r_2)
\]

\[
\left( \psi 1 \frac{1}{r_1} \right) = \int dr_1^1 \int dr_2^2 \psi_{100}^*(r_1) \psi_{100}(r_2) \frac{1}{r_1} \psi_{100}(r_1, r_2) = \int dr_1^1 \psi_{100}^*(r_1) \frac{1}{r_1} \psi_{100}(r_1) \int dr_2^2 \psi_{100}^*(r_2) \psi_{100}(r_2)
\]

\[
\text{The second integral is 1, because it is the normalization condition of a wavefunction.}
\]

\[
\left( \psi 1 \frac{1}{r_1} \right) = \frac{4 \pi}{\pi a^3 Z^3} \int_0^{\infty} r_1^2 dr_1 \frac{1}{r_1} e^{-2Z r_1} = \frac{4 Z^3}{a^3} \int_0^{\infty} dr_1 r_1 e^{-2Z r_1} = \frac{4 Z^3}{a^3} \frac{a^2}{4 Z^2} = \frac{Z}{a}
\]

Similarly, we can show that \( \langle \psi 1 \frac{1}{r_1} \rangle = \frac{Z}{a} \).

So we have,
\[
\langle \psi H \psi \rangle = -27.2 Z^2 \text{ eV} + \frac{Z (Z-2) e^2}{4 \pi \epsilon_0 a} + \frac{e^2}{4 \pi \epsilon_0} \left( \psi 1 \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right)
\]

We know from the study on hydrogen atom,
\[
E_1 = -\frac{e^2}{2} \frac{1}{4 \pi \epsilon_0 a} = -13.6 \text{ eV}
\]

So,
\[
\langle \psi H \psi \rangle = -27.2 Z^2 \text{ eV} + \frac{Z (Z-2) e^2}{4 \pi \epsilon_0 a} + \frac{e^2}{4 \pi \epsilon_0} \left( \psi 1 \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right)
\]

\[
= -27.2 Z^2 \text{ eV} + 4 Z (Z-2) [13.6 \text{ eV}] + \frac{e^2}{4 \pi \epsilon_0} \left( \psi 1 \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right)
\]

\[
= \left[ 4 Z (Z-2) - 2 Z^2 \right] \times 13.6 \text{ eV} + \frac{e^2}{4 \pi \epsilon_0} \left( \psi 1 \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right)
\]

\[
= [2 Z^2 - 8 Z] \times 13.6 \text{ eV} + \frac{e^2}{4 \pi \epsilon_0} \left( \psi 1 \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right)
\]

Now, we come to the last term, \( \left( \psi 1 \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right) \), the calculation here is pretty much the same as what we did in the previous section, so we skip the details
\[
\frac{e^2}{4 \pi \epsilon_0} \left( \psi 1 \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right) = -\frac{5 Z}{4} E_1 = -\frac{5 Z}{4} (-13.6 \text{ eV}) = \frac{5 Z}{4} \times 13.6 \text{ eV}
\]

So, we get
\[
\langle \psi H \psi \rangle = [2 Z^2 - 8 Z] \times 13.6 \text{ eV} + \frac{5 Z}{4} \times 13.6 \text{ eV} = [2 Z^2 - \frac{27}{4} Z] \times 13.6 \text{ eV}
\]

To get the ground state, we want to minimize the energy \( \langle \psi H \psi \rangle \). Here, it means that we want find the value for Z that minimizes the
To find the extreme of a function, we request the first order derivative to be zero

$$\frac{d}{dZ}\left[2Z^2 - \frac{27}{4}Z\right] = 4Z - \frac{27}{4} = 0$$  \hspace{1cm} (3.65)

And the solution is

$$Z = \frac{27}{16} \approx 1.69$$  \hspace{1cm} (3.66)

This means that for the wavefunctions that we considered,

$$\psi_0(r_1, r_2) = \frac{Z^3}{\pi a^3} e^{-Z \frac{r_1 + r_2}{a}}$$  \hspace{1cm} (3.67)

when \(Z = \frac{27}{16} \approx 1.69\), the energy reaches its minimum value, and the minimum value is

$$\langle \psi \mid H \mid \psi \rangle = \left[2Z^2 - \frac{27}{4}Z\right] \times 13.6 \text{ eV} = \left[2 \times \left(\frac{27}{16}\right)^2 - \frac{27}{4} \times \frac{27}{16}\right] \times 13.6 \text{ eV} = \frac{729}{128} \times 13.6 \text{ eV} = -77.5 \text{ eV}$$  \hspace{1cm} (3.68)

The experimental value for ground state energy is \(E_g = -78.975 \text{ eV}\). Here, our estimation is very close to the real value. The error bar is about 1% or 2%

**Comment 1:** considering that we are assuming a very simple wavefunction, this accuracy is amazingly good.

**Comment 2:** If one tries more complicated wavefunctions, the accuracy can be further improved.

**Comment 3:** as we proved early on, the variation methods always overestimate the ground state energy. Here, indeed, we get a larger energy than the real value.

### 3.4. Example: Hydrogen molecule ion

Consider a hydrogen molecule ion, \(H_2^+\). Here, we have two hydrogen nucleons (separated by distance \(R\)), and one electron moving around both of the nucleons. The Hamiltonian is

$$H = -\frac{\hbar^2}{2m} \nabla^2 - \frac{e^2}{4\pi \epsilon_0} \frac{1}{r_1} - \frac{e^2}{4\pi \epsilon_0} \frac{1}{r_2} + \frac{e^2}{4\pi \epsilon_0} \frac{1}{R}$$  \hspace{1cm} (3.69)

Here, the first term is the kinetic energy of the electron. The second and the third terms are potential energy between nucleons and the electron. The last term is the potential energy between two nucleons. The textbook missed the last term in the Hamiltonian, but added it back at the end of the day.

For a guess wavefunction, we can use the following one

$$\psi(r_1, r_2) = A[\psi_{100}(r_1) + \psi_{100}(r_2)]$$  \hspace{1cm} (3.70)

where \(\psi_{100}(r_1)\) is the ground-state wavefunction for a hydrogen atom around the first nucleon. \(\psi_{100}(r_2)\) is the ground-state wavefunction for a hydrogen atom around the second nucleon (remember that the two nucleon is separated by the distance \(R\)). \(A\) is a normalization factor, to make sure that \(\langle \psi \mid \psi \rangle = 1\).

Here, the separation between to nucleons, \(R\), is a unknown number. This is our tuning parameter when we try to minimize \(\langle \psi \mid H \mid \psi \rangle\), i.e. by minimizing \(\langle \psi \mid H \mid \psi \rangle\), we will find the value of \(R\) that has the lowest energy. And this value of \(R\) is our theoretical prediction for the spatial separation between the two nucleons in a hydrogen molecule ion.

I will skip the detailed calculation, which can be found in the textbook.

$$\langle \psi \mid H \mid \psi \rangle = E_1 \times \left(1 + 2 \frac{\frac{a}{R} - \left(1 + \frac{a}{R}\right) e^{-2Ra} + \frac{1}{a} \left(\frac{a}{R}\right)^{-a} e^{-Ra}}{1 + e^{-Ra} \left[1 + \left(\frac{a}{R}\right) + \frac{1}{2} \left(\frac{a}{R}\right)^2\right] - 2 \frac{a}{R}}\right)$$  \hspace{1cm} (3.71)

where \(E_1 = -13.6 \text{ eV}\) and \(a\) is the Bohr radius \(a = \frac{4\pi \epsilon_0 \hbar^2}{m e^2} = 0.529 \times 10^{-10} \text{ m}\).
Now, we define $x = R/a$,

$$
\langle \psi | H | \psi \rangle = -13.6 \text{ eV} \times \left[ 1 + 2 \frac{\frac{1}{x} - \left(1 + \frac{1}{x}\right) e^{-2x} + (1 + x) e^{-x}}{1 + e^{-x}(1 + x + \frac{1}{3} x^2)} - \frac{2}{x} \right]
$$

$$
= -13.6 \text{ eV} \times \left[ 1 + 2 \frac{1 - (x + 1) e^{-2x} + x(1 + x) e^{-x}}{1 + e^{-x}(1 + x + \frac{1}{3} x^2)} - \frac{2}{x} \right]
$$

$$
= -13.6 \text{ eV} \times \left[ 1 + 2 \frac{1 + e^{-x}(1 + x + \frac{1}{3} x^2)}{1 + e^{-x}(1 + x + \frac{1}{3} x^2)} - \frac{2}{x} \right]
$$

$$
= -13.6 \text{ eV} \times \left[ 1 + \frac{2}{x} e^{-x} \left( x^2 - 1 - \frac{1}{3} x^2 \right) e^{-x} \right]
$$

$$
= -13.6 \text{ eV} \times \left[ 1 - \frac{2}{x} e^{-x} \left( 1 + \frac{1}{3} x^2 \right) \right]
$$

(3.72)

If we plot $\langle \psi | H | \psi \rangle$ as a function of $x$, we get this figure:

The energy minimum is reached at $x \approx 2.4$, which means that the separation between the two protons is $R = x a = 2.4 \times 0.529 \times 10^{-10} \text{ m} = 1.27 \times 10^{-10} \text{ m}$.

The experimental measurement shows that $R = 1.06 \times 10^{-10} \text{ m}$. The value rough estimation that we did, produces reasonably good results.
The WKB approximation
Minimal coupling and Berry’s phase

Q: For charged quantum particles (e.g. electrons), if we apply an E&M field (E or B), the particle will feel the field. How do we consider the effect E and B fields in quantum mechanics?
A: Minimal coupling

### 5.1. Minimal coupling

#### 5.1.1. review on E&M (vector potential and gauge transformations)

In E&M, we learned that we can describe E and B fields using “potentials”: scalar potential $\phi$ and vector potential $A$.

$$\vec{E} = -\nabla \phi - \frac{\partial \vec{A}}{\partial t}$$  \hspace{1cm} (5.1)

and

$$\vec{B} = \nabla \times \vec{A}$$  \hspace{1cm} (5.2)

In addition, we learned that the potentials are NOT uniquely defined. In fact, we can change the scalar and vector potential in the following way, which would NOT change $E$ and $B$ fields and thus all physics would remain the same.

$$\phi(\vec{r}) \rightarrow \phi'(\vec{r}) = \phi(\vec{r}) - \frac{\partial \Lambda(\vec{r}, t)}{\partial t}$$  \hspace{1cm} (5.3)

$$\vec{A}(\vec{r}) \rightarrow \vec{A}'(\vec{r}) = \vec{A}(\vec{r}) + \nabla \Lambda(\vec{r}, t)$$  \hspace{1cm} (5.4)

where $\Lambda(\vec{r}, t)$ is any arbitrary function of $\vec{r}$ and $t$.

To prove that these new potentials $\phi'$ and $\vec{A}'$ gives exactly the same $E$ and $B$ fields as $\phi$ and $\vec{A}$, we compute the corresponding field strength

$$\vec{E}' = -\nabla \phi' - \frac{\partial \vec{A}'}{\partial t} = -\nabla \left( \phi - \frac{\partial \Lambda(\vec{r}, t)}{\partial t} \right) - \frac{\partial \left( \vec{A} + \nabla \Lambda \right)}{\partial t} = -\nabla \phi + \nabla \left( \frac{\partial \Lambda}{\partial t} - \frac{\partial \vec{A}}{\partial t} - \frac{\partial \nabla \Lambda}{\partial t} \right) - \frac{\partial \vec{A}}{\partial t} + \nabla \left( \frac{\partial \Lambda}{\partial t} - \frac{\partial \nabla \Lambda}{\partial t} \right)$$  \hspace{1cm} (5.5)

The last two terms are identical (with opposite sign), so they cancel.

$$\vec{E}' = -\nabla \phi - \frac{\partial \vec{A}}{\partial t} = -\nabla \phi + \nabla \left( \frac{\partial \vec{A}}{\partial t} + \nabla \Lambda \right) = \vec{E}$$  \hspace{1cm} (5.6)

So, the $E$ field remains the same when we change $\phi$ and $A$ to $\phi'$ and $A'$.

For the B field,

$$\vec{B}' = \nabla \times \vec{A}' = \nabla \times \left[ \vec{A} + \nabla \Lambda \right] = \nabla \times \vec{A} + \nabla \times \nabla \Lambda$$  \hspace{1cm} (5.7)
The last term contains $\nabla \times \nabla$, which is zero (The cross product for a vector with itself is always zero).

$$\vec{B}' = \nabla \times \vec{A} + \nabla \nabla \Lambda = \nabla \times \vec{A} = \vec{B}$$ (5.8)

So, the $B$ field remains the same when we change $\varphi$ and $A$ to $\varphi'$ and $A'$.

As a result, we can freely choose whether we want to use $\varphi$ and $A$ as our potentials or $\varphi'$ and $A'$. This “free choice” is known as a gauge choice. Using a different gauge will NOT change any physics.

The transformation mentioned above is known as a gauge transformation

$$\varphi \rightarrow \varphi' = \varphi - \frac{\partial \Lambda}{\partial t}$$ (5.9)

$$\vec{A} \rightarrow \vec{A}' = \vec{A} + \nabla \Lambda$$ (5.10)

Comment: scalar and vector potentials $\varphi$ and $A$ are NOT measurable quantities. In E&M we can only measure $E$ and $B$. We cannot measure $\varphi$ and $A$.

This is because we can freely choose a different gauge, which changes the value of $\varphi$ and $A$. On the other hand, for any measurable quantities, we can measure its value, and thus it must have a unique value. Since $\varphi$ and $A$ do not have unique values, they cannot be measurable quantities.

In fact, we get a more general statement along the same line of thinking: any gauge-dependent quantity (i.e. a quantity whose value changes when we choose a different gauge), can NOT be a measurable quantity.

5.1.2. Gauge transformation in quantum mechanics

In classical physics, gauge transformation only changes potentials. Matters (e.g. motion of particles) don’t care about gauge choice at all. This is because in classical physics, the motion of a particle (or any other objects) are described by measurable quantities (e.g. velocity, momentum, mass, etc). For measurable quantities, we know they can NOT change under gauge transformations.

However, in quantum, the story is different. Because in quantum mechanics, matters are described by a non-measurable quantity: i.e., the wavefunction $\psi(\vec{r}, t)$. The wavefunction is a complex function. The absolute value of the wavefunction is a measurable quantity (i.e. the square of the absolute value is the probability to find a particle at this point). However, the phase of a wavefunction is NOT measurable.

As a result, (at least in principle), the phase of a wavefunction could depends on the gauge choice. In fact, this is exactly what happens in quantum mechanics.

$$\varphi \rightarrow \varphi' = \varphi - \frac{\partial \Lambda}{\partial t}$$ (5.11)

$$\vec{A} \rightarrow \vec{A}' = \vec{A} + \nabla \Lambda$$ (5.12)

$$\psi \rightarrow \psi' = \psi e^{\frac{i q \Lambda}{\hbar}}$$ (5.13)

where $q$ is the charge of the particle.

Comment: for charge neutral particles, $q = 0$, the wavefunction doesn’t change, when we switch from $\varphi$ and $A$ to $\varphi'$ and $A'$. But for any charged particles, the complex phase of the wavefunction must change, when we switch to a different gauge choice.

To prove why we need to change the wavefunction when we change to a different gauge, we shall look at the Schrodinger equation. For any physics law, it should be invariant (i.e. remains the same) when we change to a different gauge (because different gauge choices describes exactly the same physics). As will be shown below, for the Schrodinger equation (which is one fundamental physics law), if we do not change the wavefunction, it would NOT remain the same after we switch to a different gauge. And the only way to ensure that the Schrodinger equation doesn’t change is to change the wavefunction according to the transformation describe above.

5.1.3. Minimal coupling

Now, we try to write down the Schrodinger’s equation for a charged particle moving in $E$ and $B$ fields. The way to couple the motion of a charged particle with $E$&$B$ fields is known as the minimal coupling.

We start from charge neutral particles with $q = 0$, whose Schrodinger equation is
\[ i \hbar \frac{\partial \psi}{\partial t} = \left[ \frac{\hat{p}^2}{2m} + V(r) \right] \psi \] 

(5.14)

where \( \hat{p} = -i \hbar \nabla \) is the momentum quantum operator.

Now, if we want to describe the motion of a charged particle, we only need to do the following

\[-i \hbar \nabla \rightarrow -i \hbar \nabla - q \hat{A} \] 

(5.15)

In addition, we need to do one more thing

\[ i \hbar \frac{\partial}{\partial t} \rightarrow i \hbar \frac{\partial}{\partial t} - q \varphi \] 

(5.16)

In summary, what we need to do are two things:

\[ i \hbar \frac{\partial}{\partial t} \rightarrow i \hbar \frac{\partial}{\partial t} - q \varphi \] 

(5.17)

and

\[-i \hbar \nabla \rightarrow -i \hbar \nabla - q \hat{A} \] 

(5.18)

More precisely, this are four formulas, instead of two, because the last formula is a vector formula with three components, \( x, y \) and \( z \)

\[ i \hbar \frac{\partial}{\partial t} \rightarrow i \hbar \frac{\partial}{\partial t} - q \varphi \] 

(5.19)

\[-i \hbar \partial_x \rightarrow -i \hbar \partial_x - q A_x \] 

(5.20)

\[-i \hbar \partial_y \rightarrow -i \hbar \partial_y - q A_y \] 

(5.21)

\[-i \hbar \partial_z \rightarrow -i \hbar \partial_z - q A_z \] 

(5.22)

When we have four formulas (one related with time and three with space), special relativity told us that we can often combine them together as one four-dimensional vector formula.

In special relativity, we can combine \( \varphi \) and \( \hat{A} \) together as a four-dimensional vector, known has the electromagnetic four-potential

\[ A^\mu = \left[ \frac{\varphi}{c}, A_x, A_y, A_z \right] \] 

(5.23)

In the same time, we can group \( \frac{\partial}{\partial t} \) and \( \nabla \) together as a four-dimensional vector, known as the four-gradient

\[ \partial^\mu = \left[ \frac{1}{c} \frac{\partial}{\partial t}, -\partial_x, -\partial_y, -\partial_z \right] \] 

(5.24)

As a result, when we study charged particles, we can start from the Schrodinger equation for a charge neutral particle. Whenever, we see a (time or spatial) derivative, we do the following substitution.

\[ i \hbar \omega^\mu \rightarrow i \hbar \omega^\mu - q A^\mu \] 

(5.25)

Now, let’s try the minimal coupling on the following Schrodinger equation

\[ i \hbar \frac{\partial}{\partial t} \psi = \left[ i \hbar \nabla \right]^2 + V(r) \] 

(5.26)

After we do the minimal substitution, we get

\[ \left( i \hbar \frac{\partial}{\partial t} - q \varphi \right) \psi = \left[ i \hbar \nabla - q A \right]^2 + V(r) \psi \] 

(5.27)

\[ i \hbar \frac{\partial}{\partial t} \psi - q \varphi \psi = \left[ i \hbar \nabla - q A \right]^2 + V(r) \psi \] 

(5.28)

we can move the second term on the l.h.s. to the r.h.s.
\[ i \hbar \frac{\partial}{\partial t} \psi = \left[ \frac{(-i \hbar \nabla - q A)^2}{2m} + V(r) \right] \psi + q \varphi \psi \] (5.29)

\[ i \hbar \frac{\partial}{\partial t} \psi = \left[ \frac{(-i \hbar \nabla - q A)^2}{2m} + V(r) + q \varphi \right] \psi \] (5.30)

And know, we can define a new Hamiltonian

\[ H = \frac{(-i \hbar \nabla - q A)^2}{2m} + V(r) + q \varphi \] (5.31)

The minimal coupling told us that the Hamiltonian has two modifications:

1. For the kinetic energy, the momentum is not longer just a derivative. We get an extra part \( qA \). This part is new, and it is the first time that we learned it (unless your theoretical mechanics class discussed similar topics for Hamiltonian mechanics).

2. For the potential energy, we get an additional contribution \( q\varphi \). This part is NOT new and NOT surprising at all. For static electrics, scalar potential \( \varphi \) is just the “electric potential”. If we have a charge particle with charge \( q \) moving in the presence of electric potential \( \varphi \), it gets a potential energy \( q\varphi \).

5.1.4. Verify the gauge transformation

Now we prove that the Schrodinger equation with minimal coupling is invariant under the gauge transformation

\[ \varphi \rightarrow \varphi' = \varphi - \frac{\partial \Lambda}{\partial t} \] (5.32)

\[ \vec{A} \rightarrow \vec{A}' = \vec{A} + \nabla \Lambda \] (5.33)

\[ \psi \rightarrow \psi' = \psi e^{\frac{iq}{\hbar} \Lambda} \] (5.34)

We know that

\[ \left( i \hbar \frac{\partial}{\partial t} - q \varphi \right) \psi = \left[ \frac{(-i \hbar \nabla - q A)^2}{2m} + V(r) \right] \psi \] (5.35)

Now, if we change to a different gauge, \( \varphi' \), \( \varphi' \) and \( A' \), the Schrodinger equation shall remain the same.

We start from the l.h.s.. After the gauge transformation, it becomes

\[ \left( i \hbar \frac{\partial}{\partial t} - q \varphi' \right) \psi' = \left( i \hbar \frac{\partial}{\partial t} - q \varphi' \right) \psi' \] (5.36)

For the first term

\[ i \hbar \frac{\partial}{\partial t} \psi' = i \hbar \frac{\partial}{\partial t} \psi e^{\frac{iq}{\hbar} \Lambda} = i \hbar \frac{\partial}{\partial t} \psi + i \hbar \psi \frac{\partial}{\partial t} e^{\frac{iq}{\hbar} \Lambda} = i \hbar \frac{\partial}{\partial t} \psi + i \hbar \psi e^{\frac{iq}{\hbar} \Lambda} \frac{\partial}{\partial t} \psi - q \psi e^{\frac{iq}{\hbar} \Lambda} \frac{\partial}{\partial t} \psi \]

\[ = e^{\frac{iq}{\hbar} \Lambda} \left[ i \hbar \frac{\partial}{\partial t} \psi - q \hbar \frac{\partial}{\partial t} \psi \right] \] (5.37)

As a result,

\[ \left( i \hbar \frac{\partial}{\partial t} - q \varphi' \right) \psi' = \left( i \hbar \frac{\partial}{\partial t} - q \varphi' \psi' \right) = e^{\frac{iq}{\hbar} \Lambda} \left[ i \hbar \frac{\partial}{\partial t} \psi - q \frac{\partial}{\partial t} \psi \right] - q \varphi' e^{\frac{iq}{\hbar} \Lambda} \psi \]

\[ = e^{\frac{iq}{\hbar} \Lambda} \left[ i \hbar \frac{\partial}{\partial t} \psi - q \frac{\partial}{\partial t} \psi - q \varphi \psi \right] \]

\[ = e^{\frac{iq}{\hbar} \Lambda} \left[ i \hbar \frac{\partial}{\partial t} \psi - q \psi \right] \] (5.38)

Now, we look a the right hand side. After the gauge transformation, it becomes,
Because we know that and the r.h.s. becomes Here, we proved that Let’s start from \( \text{ⅇ}^\text{i} \) which is the old l.h.s. with an extra factor Therefore, For the first term, Let’s start from \( \text{ⅇ}^\text{i} \) which is the old r.h.s., with an extra factor Above, we proved that when we move \( \text{ⅇ}^\text{i} \) \( \text{ⅇ}^\text{i} \) \( \text{ⅇ}^\text{i} \) \( \text{ⅇ}^\text{i} \) which is the old r.h.s., with an extra factor \( \text{ⅇ}^\text{i} \) because we know that and the r.h.s. becomes Therefore, we immediately know that
\[
\left(i \hbar \frac{\partial}{\partial t} - q \varphi \right) \psi' = \left[ \frac{(-i \hbar \nabla - q A')^2}{2 m} + V(r) \right] \psi'
\]

(5.49)

i.e., using the new gauge, the Schrödinger equation takes exactly the same form (i.e. the physics law remains the same).

### 5.2. Example: Landau levels and quantum Hall effect

#### 5.2.1. 2D electron gas

Now, we consider free electrons moving in 2D.

\[
H = \frac{(-i \hbar \nabla)^2}{2 m} = \frac{(-i \hbar \partial_x)^2}{2 m} + \frac{(-i \hbar \partial_y)^2}{2 m}
\]

(5.50)

Here, we have no potential energy, only kinetic energy. The eigenwavefunctions and eigenenergy for this Hamiltonian are very simple.

The eigen-wavefunctions are 2D plane waves

\[
\psi = e^{i k_x x + i k_y y}
\]

(5.51)

and the corresponding eigenenergy is

\[
\epsilon = \frac{P_x^2}{2 m} + \frac{P_y^2}{2 m} = \frac{\hbar^2 k_x^2}{2 m} + \frac{\hbar^2 k_y^2}{2 m} = \frac{\hbar^2}{2 m} (k_x^2 + k_y^2)
\]

(5.52)

Key conclusion: The eigenenergy can be any positive real number.

#### 5.2.2. 2D electron gas + uniform magnetic field

Now we apply a uniform B field (B is along the z-direction). As will shown below, the eigen-energy will now become discrete energy levels (similar to a harmonic oscillator). Here, eigen-energy can only some discrete values.

For an uniform B field along z, we can represent it using the following scalar and vector potential

\[
\varphi = 0
\]

(5.53)

and

\[
A = (0, B x, 0)
\]

(5.54)

It is easy to verify that

\[
\mathbf{E} = -\nabla \varphi - \frac{\partial A}{\partial t} = 0
\]

(5.55)

and

\[
\mathbf{B} = \nabla \times \mathbf{A} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & B x & 0 \end{vmatrix} = (\partial_y B x) \hat{z} = \mathbf{B} \hat{z}
\]

(5.56)

Note: This choice of \( \varphi \) and \( A \) are not the unique choice. We can choose any gauge, but the final result would be the same, since our Schrodinger equation doesn’t depend on gauge choices.

Now the Hamiltonian changes into

\[
H = \frac{(-i \hbar \partial_x)^2}{2 m} + \frac{(-i \hbar \partial_y)^2}{2 m} = \frac{(-i \hbar \partial_x)^2}{2 m} + \frac{(-i \hbar \partial_y - q B x)^2}{2 m}
\]

(5.57)

The static Schrodinger equation is

\[
\frac{(-i \hbar \partial_x)^2}{2 m} \psi(x, y) + \frac{(-i \hbar \partial_y - q B x)^2}{2 m} \psi(x, y) = \epsilon \psi(x, y)
\]

(5.58)

First, we notice that
\[ [H, -i \hbar \partial_y] = 0 \]  
(5.59)

This is because \( H \) doesn’t contain \( y \), so it commutes with \(-i \hbar \partial_y\).

This commutator implies that we can find common eigenstates of \( H \) and \(-i \hbar \partial_y\):

\[
\frac{-i \hbar}{2m} \psi(x, y) + \frac{(i \hbar \partial_y - q B x)^2}{2m} \psi(x, y) = \epsilon \psi(x, y)
\]

(5.60)

and

\[-i \hbar \partial_y \psi(x, y) = k_y \psi(x, y)\]

(5.61)

The solution of the second equation is easy to find

\[ \psi(x, y) = f(x) e^{i k_y y} \]

(5.62)

If we put this solution back to the first equation, we get

\[
\frac{-i \hbar}{2m} \left[ f(x) e^{i k_y y} \right] + \frac{(i \hbar \partial_y - q B x)^2}{2m} \left[ f(x) e^{i k_y y} \right] = \epsilon \left[ f(x) e^{i k_y y} \right]
\]

(5.63)

We multiply both sides with \( e^{-i k_y y} \)

\[
\frac{-i \hbar}{2m} f(x) + \frac{(i \hbar \partial_y - q B x)^2}{2m} f(x) = \epsilon f(x)
\]

(5.64)

In the second term on the l.h.s., we cannot simply cancel \( e^{-i k_y y} \) and \( e^{i k_y y} \). This is become one of them is before \( \partial_y \), while the other is after \( \partial_y \).

We know that we cannot switch \( \partial_y \) and \( e^{i k_y y} \), because \( \partial_y \) does not commute with \( y \).

Here, we utilize the fact that \(-i \hbar \partial_y e^{i k_y y} = k_y \ h e^{i k_y y}\)

\[
\frac{-i \hbar}{2m} f(x) + \frac{(i \hbar k_y - q B x)^2}{2m} f(x) = \epsilon f(x)
\]

(5.66)

\[
\frac{-i \hbar}{2m} f(x) + \frac{(i \hbar k_y - q B x)^2}{2m} f(x) = \epsilon f(x)
\]

(5.67)

\[
\left[ \frac{-i \hbar}{2m} + \frac{(i \hbar k_y - q B x)^2}{2m} \right] f(x) = \epsilon f(x)
\]

(5.68)

\[
\left[ \frac{-i \hbar}{2m} + \frac{q^2 B^2 (x - \frac{k_y}{q B})^2}{2m} \right] f(x) = \epsilon f(x)
\]

(5.69)

How, we define

\[ x_0 = \frac{h k_y}{q B} \]

(5.70)

and

\[ k = \frac{q^2 B^2}{m} \]

(5.71)

\[
\left[ \frac{-i \hbar}{2m} + \frac{k}{2} (x - x_0)^2 \right] f(x) = \epsilon f(x)
\]

(5.72)

This equation is the static Schrodinger equation for a harmonic oscillator (with equilibrium position at \( x = x_0 \)). For a harmonic oscillator, we know that the eigenenergy is

\[ \epsilon_n = \hbar \omega (n + 1/2) \]

(5.73)
where \( \omega = \sqrt{\frac{k}{m}} \)

Here, our \( k \) is \( \frac{q^2 B^2}{m} \), so we found that
\[
\omega = \sqrt{\frac{k}{m}} = \sqrt{\frac{q^2 B^2}{m^2}} = \frac{q B}{m}
\] (5.74)

This frequency is known as the “cyclotron frequency” and the eigenenergies are
\[
\epsilon_n = \hbar \omega (n + 1/2) = \frac{q h B}{m} \left( n + \frac{1}{2} \right)
\] (5.75)

The eigenenergy can no longer take arbitrary positive values. Only some discrete values are not allowed. Each of these discrete energy level is known as a Landau level.

### 5.2.3. cyclotron frequency

The cyclotron frequency has a simple physical meaning. In classical mechanics, if we have a charged particle moving in a uniform B field, the trajectory is a circle. This is because the Lorentz force is always perpendicular to velocity, which means that the acceleration is always perpendicular to the velocity, i.e. a circular motion.

If the velocity of the particle is \( v \), the Lorentz force is
\[
F = q v B
\] (5.76)

For a circular motion, the acceleration is
\[
a = \frac{v^2}{r}
\] (5.77)

We know that \( F = m a \), and thus
\[
q v B = m \frac{v^2}{r}
\] (5.78)
\[
\frac{q B}{m} = \frac{v}{r}
\] (5.79)

We found that neither \( v \) nor \( r \) is a universal number. However, their ratio is independent of \( v \) or \( r \). Once we fixed the B field, because \( q/m \) can never change for a particle (ignore relativistic effects), \( v/r \) is always a constant.

Q: what is \( v/r \) for a circular motion?

A: it is the angular frequency \( \omega \)
\[
\omega = \frac{q B}{m}
\] (5.80)

This is the cyclotron frequency.

Quantum mechanics told us, once we have a periodic motion with angular momentum \( \omega \), if it is a harmonic oscillator, we shall expect eigenenergy \( \epsilon = h \omega (n + 1/2) \). This is exactly what happened here for Landau levels.

### 5.3. Example: particles moving around a solenoid.

Consider a charged particle moving around a solenoid.
For an infinite long solenoid, there is $B$ field inside the solenoid, but the $B$ field outside the cylinder is zero. Now consider a charged particle with charge $q$. We put the particle on a ring, which circles around the solenoid, i.e. the particle can only move on this circular ring.

By changing the electric current that passes through the solenoid, we can change the field inside the solenoid, but the $B$ field outside will always remain zero. Now, we ask the question whether the motion of the charged particle will change when we change the $B$ field inside the solenoid.

In classical mechanics, the motion of the particle does NOT depend on the solenoid, because the $B$ field is always zero for any point on the ring. In classical mechanics, if $B = 0$, it means that the particle cannot see the $B$ field and thus the solenoid doesn’t matter at all, as far as the particle is concerned.

However, in quantum mechanics, this is not the case. It turns out that the particle does know the magnetic field inside the solenoid, although the field strength for any point on the ring is zero.

5.3.1. vector potential

In quantum mechanics, the magnetic field couples to the motion of a particle via vector potential $\mathbf{A}$. This is very different from classical mechanics, where what matters is $\mathbf{B}$, instead of $\mathbf{A}$. If the motion of the particle only depends on $\mathbf{B}$, then the solenoid doesn’t matter, because $B = 0$ for the ring (this is what happens in classical mechanics). If the motion of the particle relies on $\mathbf{A}$, things would be different, because when $B = 0$, $\mathbf{A}$ may not be zero.

Q: what is $\mathbf{A}$ for a solenoid

A: It is not zero on the ring, if the solenoid has a nonzero current.

Consider the following integral along the ring:

$$\oint d\mathbf{r} \cdot \mathbf{A} = \int \int d\sigma \left( \nabla \times \mathbf{A} \right) = \int \int d\sigma B = \Phi$$

(5.81)

Here, we used Stokes' theorem and the fact that the curl of $\mathbf{A}$ is $\mathbf{B}$. $\Phi$ is the magnetic flux inside the solenoid.

Conclusion: the integral $\oint d\mathbf{r} \cdot \mathbf{A}$ equals to the magnetic flux inside the solenoid. Although $B = 0$ outside the solenoid, $\mathbf{A}$ is NOT zero (if $\mathbf{A} = 0$, the integral would be zero, which is impossible, because we know that the magnetic flux is nonzero).

We can choose the gauge such that

$$\mathbf{A} = \Phi \frac{\hat{\varphi}}{2\pi r}$$

(5.82)

Here, we use the cylindrical coordinate, and $\hat{\varphi}$ is the unit vector that is tangential to the ring.

It is easy to verify that

$$\mathbf{A} = \Phi \frac{\hat{\varphi}}{2\pi r} (-\sin \varphi, \cos \varphi, 0) = \Phi \frac{\hat{\varphi}}{2\pi r} \left( \frac{y}{r}, \frac{x}{r}, 0 \right) = \Phi \frac{\hat{\varphi}}{2\pi} \left( \frac{y}{r^2 + y^2}, \frac{x}{r^2 + y^2}, 0 \right)$$

(5.83)

5.3.2. Hamiltonian
\[
H = \frac{(-i \hbar \partial_x - q A_x)^2}{2m} + \frac{(-i \hbar \partial_y - q A_y)^2}{2m} + \frac{(-i \hbar \partial_z - q A_z)^2}{2m}
\]
\[
= \frac{(-i \hbar \partial_x + \frac{q \Phi}{2\pi \hbar x^2 + y^2})^2}{2m} + \frac{(-i \hbar \partial_y - \frac{q \Phi}{2\pi \hbar x^2 + y^2})^2}{2m} + \frac{(-i \hbar \partial_z)^2}{2m}
\]
\[
= -\frac{\hbar^2}{2m} \left( \frac{\partial_x^2 + \partial_y^2 + \partial_z^2}{2\pi \hbar x^2 + y^2} \right)^2 + \frac{\hbar^2}{2m} \left( \frac{q \Phi}{2\pi \hbar x^2 + y^2} \right)^2 x^2 + y^2 + \frac{\hbar^2}{2m} \left( \frac{q \Phi}{2\pi \hbar x^2 + y^2} \right)^2 \left( \frac{1}{2\pi \hbar x^2 + y^2} \right)^2
\]

Start with the first term:
\[
\left( \partial_x + i \frac{q \Phi}{2\pi \hbar x^2 + y^2} \right)^2 \psi(x, y, z) = \frac{y^2}{(x^2 + y^2)^2}
\]

Similarly, the second term is
\[
\left( \partial_y - \frac{q \Phi}{2\pi \hbar x^2 + y^2} \right)^2 \psi(x, y, z) = \frac{y^2}{(x^2 + y^2)^2}
\]

As a result,
\[
H = -\frac{\hbar^2 \left( \partial_x^2 + \partial_y^2 + \partial_z^2 \right)}{2m} + \frac{\hbar^2}{2m} \left( \frac{q \Phi}{2\pi \hbar x^2 + y^2} \right)^2 x^2 + y^2 + \frac{\hbar^2}{2m} \left( \frac{q \Phi}{2\pi \hbar x^2 + y^2} \right)^2 \left( \frac{1}{2\pi \hbar x^2 + y^2} \right)^2
\]

Convert to the cylindrical coordinate,
\[
\nabla^2 \psi(r, \phi, z) = \frac{1}{r} \frac{\partial}{\partial r} \left[ r \frac{\partial \psi(r, \phi, z)}{\partial r} \right] + \frac{1}{r^2} \frac{\partial^2 \psi(r, \phi, z)}{\partial \phi^2} + \frac{\partial^2 \psi(r, \phi, z)}{\partial z^2}
\]

Because the particle can only move along the ring, \( r \) and \( z \) are fixed and thus the wavefunction only depends on \( \phi \), independent of \( r \) and \( z \). \( \psi(r, \phi, z) = \psi(\phi) \). Therefore, \( \partial_z \psi = \partial_z \psi = 0 \)
\[ \nabla^2 \psi = \frac{1}{r} \frac{\partial}{\partial r} \left[ r \frac{\partial \psi}{\partial r} \right] + \frac{1}{r^2} \frac{\partial^2 \psi}{\partial \varphi^2} + \frac{\partial^2 \psi}{\partial z^2} = \frac{1}{r^2} \frac{\partial^2 \psi}{\partial r^2} - \frac{1}{r^2} \frac{\partial^2 \psi}{\partial \varphi^2} = \frac{1}{r^2} \frac{d^2 \psi}{d\varphi^2} \]  

(5.90)

For the second term,

\[ (y \partial_x - x \partial_y) \psi(x, y, z) = -\vec{z} \cdot (x, y, 0) \times (\partial_x, \partial_y, \partial_z) \psi(x, y, z) = -\vec{z} \cdot \vec{r} \times \nabla \psi(x, y, z) \]

(5.91)

We know that

\[ \vec{r} = r \hat{\varphi} \]

(5.92)

\[ \vec{r} \times \nabla \psi(x, y, z) = r \hat{\varphi} \times \left( \frac{\partial \psi}{\partial r} \hat{r} + \frac{1}{r} \frac{\partial \psi}{\partial \varphi} \hat{\varphi} + \frac{\partial \psi}{\partial z} \hat{z} \right) = r \hat{\varphi} \times \left( \frac{1}{r} \frac{\partial \psi}{\partial \varphi} \hat{\varphi} + r \hat{r} \times \frac{\partial \psi}{\partial z} \hat{z} \right) \]

(5.93)

\[ \frac{\partial \psi}{\partial \varphi} = -\frac{\partial \psi}{\partial z} \]

(5.94)

\[ (y \partial_x - x \partial_y) \psi(x, y, z) = -\vec{z} \cdot \vec{r} \times \nabla \psi(x, y, z) = -\frac{\partial \psi}{\partial \varphi} = -\frac{d \psi}{d \varphi} \]

(5.95)

Therefore,

\[ H \psi(x, y, z) = -\frac{\hbar^2}{2m} \nabla^2 \psi(x, y, z) - i \frac{\hbar q \Phi}{2\pi m r^2} (y \partial_x - x \partial_y) \psi(x, y, z) + \frac{1}{2m} \left( \frac{q \Phi}{2\pi r} \right)^2 \psi(x, y, z) \]

(5.96)

\[ = -\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{d^2 \psi}{d\varphi^2} + i \frac{\hbar q \Phi}{2\pi m r^2} \frac{d \psi}{d \varphi} + \frac{1}{2m} \left( \frac{q \Phi}{2\pi r} \right)^2 \psi \]

(5.97)

The Schrodinger equation

\[ H \psi(\varphi) = E \psi(\varphi) \]

(5.98)

So we have

\[ \frac{\hbar^2}{2m} \frac{1}{r^2} \frac{d^2 \psi}{d\varphi^2} + i \frac{\hbar q \Phi}{2\pi m r^2} \frac{d \psi}{d \varphi} + \frac{1}{2m} \left( \frac{q \Phi}{2\pi r} \right)^2 \psi = E \psi(\varphi) \]

(5.99)

\[ \frac{d^2 \psi}{d\varphi^2} + \frac{i q \Phi}{\pi \hbar} \frac{d \psi}{d \varphi} + \left( \frac{q \Phi}{2\pi \hbar} \right)^2 \psi = -\frac{2m E r^2}{\hbar^2} \psi \]

(5.100)

Define \( \beta = \frac{q \Phi}{2\pi R_0} \) and \( \epsilon = \frac{2m E R_0^2}{\hbar^2} - \beta^2 \)

\[ \frac{d^2 \psi}{d\varphi^2} + \frac{2i \beta}{\pi \hbar} \frac{d \psi}{d \varphi} + \epsilon \psi = 0 \]

(5.101)

So,

\[ \frac{d^2 \psi}{d\varphi^2} - 2i \beta \frac{d \psi}{d \varphi} + \epsilon \psi = 0 \]

(5.102)

The solution for this second order differential equation is

\[ \psi(\varphi) = A e^{\lambda \varphi} \]

(5.103)

\[ -\lambda^2 \psi + 2\lambda \beta \psi + \epsilon \psi = 0 \]

(5.104)

\[ -\lambda^2 \psi + 2\lambda \beta \psi + \epsilon \psi = 0 \]

(5.105)
\[ \lambda^2 - 2\beta \lambda - \epsilon = 0 \]  

So we have,

\[ \lambda = \beta \pm \sqrt{\beta^2 + \epsilon} = \frac{q \Phi}{2\pi \hbar} \pm \sqrt{\frac{2mE r^2}{\hbar^2}} = \frac{q \Phi}{2\pi \hbar} \pm \frac{r}{\hbar} \sqrt{2mE} \]  

(5.107)

On a ring, \( \psi(\varphi) = \psi(\varphi + 2\pi) \),

\[ \psi(\varphi) = A e^{i\lambda \varphi} \]  

(5.108)

\[ \psi(\varphi + 2\pi) = A e^{i\lambda (\varphi + 2\pi)} = A e^{i\lambda \varphi} e^{2\pi i \lambda} \]  

(5.109)

So, we have

\[ e^{2\pi i \lambda} = 1 \]  

(5.110)

\[ \lambda = n \]  

(5.111)

where \( n \) is an integer. As a result,

\[ \lambda = \beta \pm \sqrt{\beta^2 + \epsilon} = \frac{q \Phi}{2\pi \hbar} \pm \sqrt{\frac{2mE r^2}{\hbar^2}} = \frac{q \Phi}{2\pi \hbar} \pm \frac{r}{\hbar} \sqrt{2mE} = n \]  

(5.112)

\[ E_n = \frac{\hbar^2}{2mr^2} \left( n - \frac{q \Phi}{2\pi \hbar} \right)^2 \]  

(5.113)

In the absence of magnetic field, \( \Phi = 0 \),

\[ E_n = \frac{\hbar^2}{2mr^2} n^2 \]  

(5.114)

In the presence of magnetic field, \( \Phi \neq 0 \)

\[ E_n = \frac{\hbar^2}{2mr^2} \left( n - \frac{q \Phi}{2\pi \hbar} \right)^2 \]  

(5.115)

The eigenenergy depends on \( \Phi \). Although \( B = 0 \) for the ring, the quantum motion of the particle DOES know the magnetic field inside the ring.

5.3.3. Eigenenergy

Here, we consider a few cases:

(a) No magnetic field

If \( \Phi = 0 \), we have one ground state with \( n = 0 \), its energy is \( E_0 = 0 \). For excited states, \( n \neq 0 \), they are all two-fold degenerate, \( E_n = E_{-n} = \frac{\hbar^2}{2mr^2} n^2 \)

(b) Integer flux

If \( \Phi = \frac{2\pi \hbar}{q} m \), i.e., \( \frac{q \Phi}{2\pi \hbar} \) is an integer \( m \)

\[ E_n = \frac{\hbar^2}{2mr^2} (n - m)^2 \]  

(5.116)

Here, we redefine \( n - m \) as our known integer \( n \)

\[ E_n = \frac{\hbar^2}{2mr^2} n^2 \]  

(5.117)

It is the same as the case with \( \Phi = 0 \).

Key conclusion, for \( \Phi = \frac{2\pi \hbar}{q} m \), no matter what the value of \( m \), they have the same eigenenergies as the case of \( B = 0 \).

(c) Half integer flux
If $\Phi = \frac{2\pi\hbar}{q}$ ($m + 1/2$), i.e., $\frac{q\Phi}{2\pi\hbar}$ is a half integer $m + 1/2$

$$E_n = \frac{\hbar^2}{2 m r^2} \left( n - m - \frac{1}{2} \right)^2$$

(5.118)

Here, we redefine $n - m$ as our known integer $n$

$$E_n = \frac{\hbar^2}{2 m r^2} \left( n - \frac{1}{2} \right)^2$$

(5.119)

Here, any quantum state (including the ground state) is two-fold degenerate. The ground states are:

$$E_1 = E_0 = -\frac{\hbar^2}{8 m r^2}$$

(5.120)

First excited states,

$$E_2 = E_{-1} = \frac{\hbar^2}{2 m r^2} \left( 2 - \frac{1}{2} \right)^2 = \frac{9 \hbar^2}{8 m r^2}$$

(5.121)

Second excited states

$$E_3 = E_{-2} = \frac{\hbar^2}{2 m r^2} \left( 3 - \frac{1}{2} \right)^2 = \frac{25 \hbar^2}{8 m r^2}$$

(5.122)

(d) Non-integer, non-half integer flux

If $\frac{q\Phi}{2\pi\hbar}$ is neither an integer nor a half integer,

$$E_n = \frac{\hbar^2}{2 m r^2} \left( n - \frac{q\Phi}{2\pi\hbar} \right)^2$$

(5.123)

None of the eigenstates is degenerate. $E_n \neq E_m$ is $n \neq m$.

5.3.4. Eigenwavefunction

Q: $\Phi = 0$ and $\Phi = \frac{2\pi\hbar}{q} m$ have the same eigenenergies. But do they have the same eigenwavefunctions?

A: No

For $\Phi = 0$,

$$E_n = \frac{\hbar^2}{2 m r^2} \left( n - \frac{q\Phi}{2\pi\hbar} \right)^2 = \frac{\hbar^2 n^2}{2 m r^2}$$

(5.124)

The corresponding wavefunctions are

$$\psi_n = A e^{i n \varphi}$$

(5.125)

The ground state is $n = 0$. Eigenenergy $E_0 = 0$ and wavefunction $\psi_0 = A$.

First excited states, $n = \pm 1$. Eigenenergy $E_0 = \frac{\hbar^2}{2 m r^2}$ and wavefunctions $\psi_{\pm 1} = A e^{\pm i \varphi}$.

Second excited states, $n = \pm 2$. Eigenenergy $E_0 = \frac{2\hbar^2}{m r^2}$ and wavefunctions $\psi_{\pm 1} = A e^{\pm 2 i \varphi}$.

For $\Phi = \frac{2\pi\hbar}{q} m$, with $m \neq 0$

$$E_n = \frac{\hbar^2}{2 m r^2} \left( n - \frac{q\Phi}{2\pi\hbar} \right)^2 = \frac{\hbar^2 (n - m)^2}{2 m r^2}$$

(5.126)

The corresponding wavefunctions are

$$\psi_n = A e^{i n \varphi}$$

(5.127)

The ground state is $n = m$. Eigenenergy $E_0 = 0$ and wavefunction $\psi_0 = A$.

First excited states, $n = m \pm 1$. Eigenenergy $E_0 = \frac{\hbar^2}{2 m r^2}$ and wavefunctions $\psi_{\pm 1} = A e^{i (m \pm 1) \varphi}$. 


Second excited states, \( n = m \pm 2 \). Eigenenergy \( E_0 = \frac{3 \hbar^2}{2m} \) and wavefunctions \( \psi_{\pm 1} = A \ e^{i(m\pm 2)\varphi} \).

In general, magnetic flux \( \Phi \) and \( \Phi' = \Phi + \frac{2\pi\hbar}{q} m \) have the same eigenenergies, but their wavefunctions are NOT the same.

### 5.4. Example: The Aharonov-Bohm Effect

Also known as the AB effect.

Consider a particle with charge \( q \) moving near a solenoid. We assume that the particle can NOT go inside the solenoid. As a result, \( B = 0 \) for the particle.

\[
H = \frac{(-i\hbar \nabla - q\vec{A})^2}{2m} + V(\vec{r})
\]

The Schrodinger equation

\[
\frac{(-i\hbar \nabla - q\vec{A})^2}{2m} \psi + V(\vec{r}) \psi = E \psi
\]

Now, define

\[
\psi(\vec{r}) = e^{ig(\vec{r})} \psi'(\vec{r})
\]

where

\[
g(\vec{r}) = \frac{q}{\hbar} \int_O^\vec{r} d\vec{r}' \cdot \vec{A}
\]

\( O \) is an arbitrary reference point. The integral here is a line integral. The line starts from the reference point \( O \) and ends at \( \vec{r} \).

#### 5.4.1. Is \( g(\vec{r}) \) well defined?

Q: From \( O \) to \( \vec{r} \), we can define many different paths, do they all give the same \( g(\vec{r}) \)?

The answer to this question is very important. Because, if different paths give different \( g(\vec{r}) \), then the function \( g(\vec{r}) \) is ill-defined (the value of this function is NOT uniquely defined by \( \vec{r} \)). As a result, the relation that we assumed above become ill-defined:

\[
\psi(\vec{r}) = e^{ig(\vec{r})} \psi'(\vec{r})
\]

Luckily, this problem will not happen in this setup.

Assuming that we have two different paths from \( O \) to \( \vec{r} \), \( I \) and \( II \). Since these two paths have the same starting and ending point, they form a loop \( C = I - II \). The minus sign here means that for path \( II \). We will move in the opposite direction, from \( \vec{r} \) to \( O \).

Here, we assume that the loop \( C \) does NOT enclose the solenoid.

Using paths \( I \) and \( II \), we can define two different \( g \) function

\[
g_I(\vec{r}) = \frac{q}{\hbar} \int_I d\vec{r}' \cdot \vec{A}
\]

\[
g_{II}(\vec{r}) = \frac{q}{\hbar} \int_{II} d\vec{r}' \cdot \vec{A}
\]

We compute the difference between them and show that \( g_I(\vec{r}) - g_{II}(\vec{r}) = 0 \)

\[
g_I(\vec{r}) - g_{II}(\vec{r}) = \frac{q}{\hbar} \int_I d\vec{r}' \cdot \vec{A} - \frac{q}{\hbar} \int_{II} d\vec{r}' \cdot \vec{A} = \frac{q}{\hbar} \int_{C} d\vec{r}' \cdot \vec{A} = \frac{q}{\hbar} \int_{C} d\vec{r}' \cdot \vec{A} \]

\[
= \frac{q}{\hbar} \int_{\text{inside } C} ds \nabla \cdot \vec{A} = \frac{q}{\hbar} \int_{\text{inside } C} ds B = \frac{q}{\hbar} \Phi_{\text{inside } C}
\]

Because we have assumed that there is no B field inside \( C \), \( \Phi_{\text{inside } C} = 0 \). Thus, we showed that different paths give the same \( g \) function, i.e. this
function is independent of the choice of paths. It only depends on the starting and ending points.

5.4.2. Schrodinger Equation

\[
\left( -i \hbar \nabla - q \vec{A} \right)^2 \psi + V(\vec{r}) \psi = E \psi
\]  

(5.136)

We substitute \( \psi(\vec{r}) \) using

\[
\psi(\vec{r}) = e^{i \vec{q} \cdot \vec{r}} \psi'(\vec{r})
\]

(5.137)

Here, we know that

\[
\left( -i \hbar \nabla - q \vec{A} \right) \psi = \left( -i \hbar \nabla - q \vec{A} \right) \left[ e^{i \vec{q} \cdot \vec{r}} \psi'(\vec{r}) \right] = -i \hbar \nabla e^{i \vec{q} \cdot \vec{r}} \psi' - q \vec{A} e^{i \vec{q} \cdot \vec{r}} \psi' = -i \hbar e^{i \vec{q} \cdot \vec{r}} \nabla \psi' - q \vec{A} e^{i \vec{q} \cdot \vec{r}} \psi'
\]

(5.138)

In the first term,

\[
\nabla g(\vec{r}) = \frac{q}{\hbar} \int_0^\vec{r} d\vec{r}', \vec{A} = \frac{q}{\hbar} \vec{A}(\vec{r})
\]

(5.139)

Thus

\[
\left( -i \hbar \nabla - q \vec{A} \right) \psi = -i \hbar e^{i \vec{q} \cdot \vec{r}} \nabla \psi' - q \vec{A} e^{i \vec{q} \cdot \vec{r}} \psi' = -i \hbar e^{i \vec{q} \cdot \vec{r}} \nabla \psi' - q \vec{A} e^{i \vec{q} \cdot \vec{r}} \psi'
\]

(5.140)

Now, we can take the second derivative,

\[
\left( -i \hbar \nabla - q \vec{A} \right)^2 \psi = -i \hbar e^{i \vec{q} \cdot \vec{r}} \nabla \psi' - q \vec{A} e^{i \vec{q} \cdot \vec{r}} \psi' = -i \hbar e^{i \vec{q} \cdot \vec{r}} \nabla \psi' - q \vec{A} e^{i \vec{q} \cdot \vec{r}} \psi'
\]

(5.141)

Using the same procedure, we find that

\[
\left( -i \hbar \nabla - q \vec{A} \right)^2 \psi = -i \hbar e^{i \vec{q} \cdot \vec{r}} \nabla \psi' = -i \hbar e^{i \vec{q} \cdot \vec{r}} \nabla \psi' = -h^2 e^{i \vec{q} \cdot \vec{r}} \nabla^2 \psi'
\]

(5.142)

And thus,

\[
-\frac{\hbar^2}{2m} e^{i \vec{q} \cdot \vec{r}} \nabla^2 \psi' + V(\vec{r}) e^{i \vec{q} \cdot \vec{r}} \psi' = E e^{i \vec{q} \cdot \vec{r}} \psi'
\]

(5.143)

Thus,

\[
-\frac{\hbar^2}{2m} \nabla^2 \psi' + V(\vec{r}) \psi' = E \psi'
\]

(5.144)

This equation is the Schrodinger equation for the same potential, but with no magnetic field!

In other words, if we know how a particle move in the absence of the solenoid, then we immediately know the motion of the same particle in the presence of a solenoid.

(a) The eigenenergies are exactly the same

(b) The eigenwavefunctions are almost the same, except that the phase is now different.

\[
\psi(\vec{r}) = e^{i \vec{q} \cdot \vec{r}} \psi'(\vec{r})
\]

(5.145)

where

\[
g(\vec{r}) = \frac{q}{\hbar} \int_0^\vec{r} d\vec{r}', \vec{A}
\]

(5.146)

Here, \( \psi' \) is the wavefunction without the solenoid. \( \psi \) is the wavefunction when we add the solenoid.

Key conclusion: the only thing that we get after introducing the solenoid is an extra phase factor.
The Aharonove-Bohm Effect

The phase factor that we found has a direct experimental impact. Consider an interference experiment, we split particle a beam into two and then recombine them. We know that particles are waves, so we will see interference patterns, same as light or other waves. If the phase difference between the two paths is $2n\pi$, we see more particles (constructive interference). If the phase difference is $(2n+1)\pi$, we see less number of particles (destructive interference).

Now we introduce a solenoid between the two paths. The solenoid will introduce an extra phase factor for our particles. For particle taking path I, the phase factor is

$$g_I = \frac{q}{\hbar} \int_I d\vec{r} \cdot \vec{A}$$

For the other path (II), the phase factor is

$$g_II = \frac{q}{\hbar} \int_{II} d\vec{r} \cdot \vec{A}$$

Now, the phase difference between the two paths become

$$\Delta\phi = \Delta\phi' + g_I - g_II$$

where $\Delta\phi'$ is the phase difference before we introduce the solenoid.

As we have shown above,

$$g_I - g_II = \frac{q}{\hbar} \Phi$$

Here, $\Phi \neq 0$ is the total flux inside the solenoid.

$$\Delta\phi = \Delta\phi' + \frac{q}{\hbar} \Phi$$

The solenoid changes the phase difference between the two paths. i.e., it shifts the interference pattern.

The extreme case happens when $\frac{q}{\hbar} \Phi = (2n+1)\pi$,

$$\Delta\phi = \Delta\phi' + (2n+1)\pi$$

There, a constrictive interference become descriptive, and vice versa. In other words, by introducing the solenoid, dark lines in the interference pattern become bright lines, and vice versa.

NOTE 1: in an experiment, we always put the solenoid there, but changes its current. When $I = 0$, there is no B field, and thus it is Equivalent to having no solenoid. By increasing $I$, we can continuously increase $\Phi$, which shift the interference pattern continuously.

NOTE 2: The AB effect offers a way to measure $q$. By increasing $\Phi$ continuously, there will be a point where the interference pattern recovers the one at $\Phi = 0$. At that point, $\frac{q}{\hbar} \Phi = 2\pi$. By measuring this flux, we get $q = \frac{2\pi \hbar}{\Phi}$. In fractional quantum Hall effect, the particles has fractional charge (e.g., $e/3$ or $e/5$ ...). The direct experimental evidence comes from the AB effect measurement. In a superconductor, the charge carriers have charge 2e (a pair of electrons, which are known as Copper pairs). The smoking gun evidence there also comes from AB effect (there, the AB effect leads to a new phenomenon known as the Josephson effect, which is seeing experimentally and the experimental data shows charge 2e, instead of e).
5.5. T-dependent Hamiltonian (an simple example)

Consider a system, whose \( H \) changes at different time.

5.5.1. Review on spin-1/2

Here, we have a spin-1/2 particle (with magnetic dipole \( \vec{\mu} = g \mu_B \hat{S}/\hbar \)). Here \( g \) is a dimensionless constant, known as the \( g \) – factor. For an electron \( g = -2 \). \( \mu_B \) is the Bohr magneton \( \mu_B = e \hbar /2m \) and \( \hat{S} \) is the spin quantum number.

For spin-1/2, we can use up and down states as the basis |up\rangle and |down\rangle. Both these two states are eigenstates of \( L_z \) (i.e. the spin angular momentum is along the \( z \) direction). The former has angular momentum +\( \hbar/2 \), while the other has -\( \hbar/2 \). All quantum spin states can now be presented as a linear superposition of |up\rangle and |down\rangle

\[
| \psi \rangle = \alpha \text{up} \rangle + \beta \text{down} \rangle
\]

We can use a column vector \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) to represent the spin up state, and \( \begin{pmatrix} 0 \\ 1 \end{pmatrix} \) to represent spin down. Any quantum spin state can now be presented as a linear superposition of these two states

\[
| \psi \rangle = \alpha \text{up} \rangle + \beta \text{down} \rangle = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}
\]

Note: any spin state is now a 2-component vector (in general, for particles with higher spins \( S > 1/2 \), a quantum state is a \( 2S+1 \)-dimensional vector).

Since quantum states are 2-component vectors, quantum operators are \( 2 \times 2 \) matrices

\[
S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}
\]

\[
S_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}
\]

\[
S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\]

Here, we assume that the particle cannot move (ignore kinetic energy) and we apply an external magnetic field \( B \)

\[
H = -\vec{B} \cdot \vec{\mu} = -g \frac{\mu_B}{\hbar} \vec{B} \cdot \hat{S} = -g \frac{|e|}{2m} \vec{B} \cdot \hat{S}
\]

For an electron, \( g = -2 \)

\[
H = \frac{|e|}{m} \frac{\hbar}{m} B_0 \hat{S} = \frac{|e|}{m} \frac{\hbar}{m} \begin{pmatrix} B_x & B_y \\ B_y & -B_x \end{pmatrix}
\]

\[
H = \frac{|e|}{2m} \frac{\hbar}{m} B_0 \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix}
\]

In spherical coordinate, \( \vec{B} = (B_x, B_y, B_z) = B_0 (\sin \phi \cos \theta, \sin \phi \sin \theta, \cos \theta) \), so

\[
H = \frac{|e|}{2m} \frac{\hbar}{m} B_0 \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix}
\]

Define \( \omega_1 = |e| B_0 /m \)

\[
H = \frac{\hbar \omega_1}{2} \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix}
\]

5.5.2. T-dependent B field

Now consider a rotating B field.
In spherical coordinate, we have \( B = B_0 \), which is a constant and \( \theta = \alpha \), which is also a constant. \( \phi = \omega t \) changes with time \( t \).

\[
H(t) = \frac{\hbar \omega_1}{2} \begin{pmatrix} \cos \alpha & \sin \alpha e^{-i\omega t} \\ -\sin \alpha e^{i\omega t} & -\cos \alpha \end{pmatrix}
\]  

(5.162)

In x-y-z coordinate, we have a constant \( B_z = B_0 \cos \alpha \). However, the x- and y-components are changing with angular frequency \( \omega \), \( B_x = B_0 \sin \theta \cos \phi = B_0 \sin \alpha \cos \omega t \) and \( B_y = B_0 \sin \theta \sin \phi = B_0 \sin \alpha \sin \omega t \).

### 5.5.3. Eigenstates and eigenenergies at each time \( t \)

At each time \( t \), we have a different Hamiltonian

\[
H(t) = \frac{\hbar \omega_1}{2} \begin{pmatrix} \cos \alpha & \sin \alpha e^{-i\omega t} \\ -\sin \alpha e^{i\omega t} & -\cos \alpha \end{pmatrix}
\]  

(5.163)

We can define \( \omega_1 = -\frac{\hbar B_0}{m} \), so

\[
H(t) = \frac{\hbar \omega_1}{2} \begin{pmatrix} \cos \alpha & \sin \alpha e^{-i\omega t} \\ -\sin \alpha e^{i\omega t} & -\cos \alpha \end{pmatrix}
\]  

(5.164)

For each of these Hamiltonians, we can find its eigenenergies and eigenstates

\[
H(t) \chi = E \chi
\]  

(5.165)

For each time \( t \), we have a different \( H \) and thus we expect that for different \( t \), we get different eigenenergy and eigenstates.

\[
\frac{\hbar \omega_1}{2} \begin{pmatrix} \cos \alpha & \sin \alpha e^{-i\omega t} \\ -\sin \alpha e^{i\omega t} & -\cos \alpha \end{pmatrix} \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = E \begin{pmatrix} \beta \\ \gamma \end{pmatrix}
\]  

(5.166)

This is an eigenvalue problem for a 2×2 matrix, we will skip the details and jump to the final answer. It turns out that for any time \( t \), the eigenenergies remains the same \( E_+ = \frac{\epsilon \hbar}{2m} B_0 \) and \( E_- = -\frac{\epsilon \hbar}{2m} B_0 \). The former corresponds to spin anti-parallel to \( B \) at that moment and the latter is spin parallel to \( B \).

The eigenstate for \( E_+ \) is

\[
\chi_+ = \begin{pmatrix} \cos \frac{\alpha}{2} \\ e^{i\omega t} \sin \frac{\alpha}{2} \end{pmatrix}
\]  

(5.167)

and the \( E_- \) state is

\[
\chi_- = \begin{pmatrix} \sin \frac{\alpha}{2} \\ -e^{i\omega t} \cos \frac{\alpha}{2} \end{pmatrix}
\]  

(5.168)

### 5.5.4. If we state from one eigenstate at \( t = 0 \), does the system remains as an eigenstates at \( t > 0 \)?

If \( H \) is independent of \( t \), the answer is yes. If we state from an eigenstate of \( H \) (say \( \psi_\alpha \)) at \( t = 0 \), then at later time, the wavefunction only pick up a phase.

\[
\psi_\alpha(t) = \psi_\alpha e^{-i E_\alpha t}
\]  

(5.169)
\( \psi_n(t) \) is still the (same) eigenstate of \( H \)

But for \( t \)-dependent \( H \). The answer is \( \text{NO.} \)

### 5.5.5. Solution for the \( t \)-dependent Schrodinger’s equation

If we assume that at time \( t = 0 \), the electron starts out with spin along \( B(t = 0) \), (i.e., we start from one of the two eigenstates of \( H \) at \( t = 0 \)), we get initial condition

\[
\psi(0) = \chi_+(0) \tag{5.170}
\]

In the same time, we know that the state need to follow the \( t \)-dependent Schrodinger’s equation

\[
i \hbar \frac{d}{dt} \psi(t) = H(t) \psi(t) \tag{5.171}
\]

This differential equation and the initial condition mentioned above, uniquely determine \( \psi(t) \). If we find the solution, we will see that \( \psi(t) \neq \chi_+(t) \) for \( t > 0 \), although we start from \( \psi(0) = \chi_+(0) \).

The Schrodinger’s equation:

\[
i \hbar \frac{d}{dt} \beta(t) = \gamma(t) \tag{5.172}
\]

\[
\frac{i \hbar}{\hbar} \left( \frac{d\beta(t)}{dt} \right) = \frac{\hbar \omega_1}{2} \left( \cos \alpha \beta(t) + \sin \alpha e^{-i \omega t} \gamma(t) \right) \tag{5.173}
\]

Two differential equations:

\[
\left\{
\begin{array}{l}
i \hbar \frac{d\beta(t)}{dt} = \frac{\hbar \omega_1}{2} \left( \cos \alpha \beta(t) + \sin \alpha e^{-i \omega t} \gamma(t) \right) \\
i \hbar \frac{d\gamma(t)}{dt} = \frac{\hbar \omega_1}{2} \left( \sin \alpha e^{i \omega t} \beta(t) - \cos \alpha \gamma(t) \right)
\end{array}
\right. \tag{5.174}
\]

The initial condition

\[
\frac{\beta(t = 0)}{\gamma(t = 0)} = \chi_+(t = 0) = \begin{pmatrix} \cos \frac{\alpha}{2} \\ \sin \frac{\alpha}{2} \end{pmatrix} \tag{5.175}
\]

which means

\[
\begin{align*}
\beta(0) &= \cos \frac{\alpha}{2} \\
\gamma(0) &= \sin \frac{\alpha}{2}
\end{align*} \tag{5.176}
\]

With this initial condition and the Shrodinger’s equation, we get one and only one solution

\[
\psi(t) = \begin{pmatrix} \cos \left( \frac{\lambda t}{2} \right) - i \frac{\omega}{\lambda} \sin \left( \frac{\lambda t}{2} \right) \cos \left( \frac{\alpha}{2} \right) e^{-i \omega t/2} \\ \cos \left( \frac{\lambda t}{2} \right) - i \frac{\omega}{\lambda} \sin \left( \frac{\lambda t}{2} \right) \sin \left( \frac{\alpha}{2} \right) e^{i \omega t/2} \end{pmatrix} \tag{5.177}
\]

where

\[
\lambda = \sqrt{\omega^2 + \omega_1^2 - 2 \omega \omega_1 \cos \alpha} \tag{5.178}
\]

### 5.5.6. How did we find the solution? (not required)

\[
\left\{
\begin{array}{l}
i \hbar \frac{d\beta(t)}{dt} = \frac{\hbar \omega_1}{2} \left( \cos \alpha \beta(t) + \sin \alpha e^{-i \omega t} \gamma(t) \right) \\
i \hbar \frac{d\gamma(t)}{dt} = \frac{\hbar \omega_1}{2} \left( \sin \alpha e^{i \omega t} \beta(t) - \cos \alpha \gamma(t) \right)
\end{array}
\right. \tag{5.179}
\]

Define
\[ \beta(t) = f(t) e^{-i \omega t/2} \]  
\[ \gamma(t) = g(t) e^{i \omega t/2} \]  
\[ \frac{d}{dt} (f(t) e^{-i \omega t/2}) = -i \frac{\omega}{2} (\cos \alpha f(t) + \sin \alpha g(t)) e^{-i \omega t/2} \]  
\[ \frac{d}{dt} (g(t) e^{i \omega t/2}) = -i \frac{\omega}{2} (\cos \alpha f(t) + \sin \alpha g(t)) e^{i \omega t/2} \]  
So we have

\[ \frac{d}{dt} \big( e^{-i \omega t/2} \big) = -i \frac{\omega}{2} f(t) e^{-i \omega t/2} = \frac{\omega}{2} \sin \alpha \]  
\[ \frac{d}{dt} \big( e^{i \omega t/2} \big) = -i \frac{\omega}{2} g(t) e^{i \omega t/2} = -i \frac{\omega}{2} \sin \alpha \]  
\[ \frac{d}{dt} f(t) - i \frac{\omega}{2} \sin \alpha g(t) \]  
\[ \frac{d}{dt} g(t) + i \frac{\omega}{2} \sin \alpha f(t) \]  
For these differential equations, all terms are proportional to \( f \) or \( g \) (or their derivative) and all the coefficients are constants, independent of time \( t \). For this structure, the solution must be exponential functions of \( t \)

\[ \begin{cases} 
\sigma(t) = A e^{i \omega t} \\
g(t) = B e^{i \omega t} 
\end{cases} \]

\[ i \omega' A e^{i \omega t} = -i \frac{\omega}{2} \sin \alpha \]  
\[ i \omega' B e^{i \omega t} = -i \frac{\omega}{2} \sin \alpha \]  
\[ \omega' A = -i \frac{\omega}{2} \sin \alpha \]  
\[ \omega' B = -i \frac{\omega}{2} \sin \alpha \]  
Now we convert the differential equations to algebra equations. This equation is actually an eigenvalue problem for a 2x2 matrix

\[ \begin{pmatrix} 
\frac{\omega}{2} \sin \alpha & \omega \\
\omega & \frac{\omega}{2} \sin \alpha 
\end{pmatrix} \begin{pmatrix} 
A \\
B 
\end{pmatrix} = \omega' \begin{pmatrix} 
A \\
B 
\end{pmatrix} \]

where \( \omega' \) is the eigenvalue for the matrix \( M = \begin{pmatrix} 
\frac{\omega}{2} \sin \alpha & \omega \\
\omega & \frac{\omega}{2} \sin \alpha 
\end{pmatrix} \), while \( \begin{pmatrix} 
A \\
B 
\end{pmatrix} \) is the eigenvector

The eigenvalues for of a matrix is the solution of equation \( \det(\omega' I - M) = 0 \)

\[ \det \left( \omega' - \frac{\omega}{2} \sin \alpha \right) \]  
\[ \omega' = \pm \sqrt{\left( \frac{\omega}{2} \sin \alpha \right)^2 + \left( \frac{\omega}{2} \cos \alpha \right)^2} \]  
so the eigenvalues are

\[ \omega' = \pm \frac{\lambda}{2} \]

We can define

\[ \lambda = \sqrt{\omega'^2 + \omega_2^2 - 2 \omega \omega \cos \alpha} \]

\[ \omega' = \pm \frac{\lambda}{2} \]
For \( \omega' = \lambda / 2 \), the eigenvector satisfies
\[
\begin{pmatrix}
\frac{\sin \omega}{2} \\
\frac{\sin \omega}{2}
\end{pmatrix}
\begin{pmatrix}
\frac{\sin \alpha}{2} \\
-\frac{\sin \alpha}{2}
\end{pmatrix}
= \frac{\lambda}{2}
\begin{pmatrix}
A \\
B
\end{pmatrix}
\]
(5.195)

The solution is
\[
\begin{pmatrix}
A_s \\
B_s
\end{pmatrix}
= \begin{pmatrix}
\frac{\sin \alpha}{2} \\
\frac{\sin \alpha}{2}
\end{pmatrix}
= \begin{pmatrix}
\frac{\sin \alpha}{2} \\
\frac{\sin \alpha}{2}
\end{pmatrix}
(5.196)

For \( \omega' = -\lambda / 2 \), the eigenvector is
\[
\begin{pmatrix}
A_l \\
B_l
\end{pmatrix}
= \begin{pmatrix}
\frac{\sin \alpha}{2} \\
\frac{\sin \alpha}{2}
\end{pmatrix}
(5.197)

So any solution of the Schrödinger equation is
\[
\psi(t) = X e^{i \frac{\lambda}{2} t} \begin{pmatrix} A_s e^{-i \omega t/2} \\ B_s e^{i \omega t/2} \end{pmatrix} + Y e^{-i \frac{\lambda}{2} t} \begin{pmatrix} A_l e^{-i \omega t/2} \\ B_l e^{i \omega t/2} \end{pmatrix}
\]
(5.198)

Using the initial condition
\[
\psi(0) = X \begin{pmatrix} A_s \\ B_s \end{pmatrix} + Y \begin{pmatrix} A_l \\ B_l \end{pmatrix} = \begin{pmatrix} (X + Y) \sin \alpha & \sin \alpha \\ (X + Y) \cos \alpha & -\cos \alpha \end{pmatrix} = \begin{pmatrix} \cos \frac{\alpha}{2} \\ \sin \frac{\alpha}{2} \end{pmatrix}
\]
(5.199)

We know that
\[
X + Y = \frac{2 \cos \frac{\alpha}{2}}{\omega \sin \alpha} = \frac{2 \cos \frac{\alpha}{2}}{\omega \sin \alpha} = \frac{1}{2 \omega \sin \alpha}
\]
(5.200)

\[
X - Y = \frac{\omega - \omega_1 \cos \alpha}{\lambda} - \frac{2 \sin \alpha}{\lambda} = \frac{\omega - \omega_1 \cos \alpha}{\lambda \omega_1 \sin \frac{\alpha}{2}} - \frac{2 \omega_1 \sin \frac{\alpha}{2}}{\lambda \omega_1 \sin \frac{\alpha}{2}} = \frac{\omega - \omega_1 (\cos \alpha + 2 \sin \frac{\alpha}{2})}{\lambda \omega_1 \sin \frac{\alpha}{2}} = \frac{\omega - \omega_1}{\lambda \omega_1 \sin \frac{\alpha}{2}}
\]
(5.201)

So we find that
\[
X = \frac{\lambda + \omega - \omega_1}{2 \lambda \omega_1 \sin \frac{\alpha}{2}}
\]
(5.202)
\[
Y = \frac{\lambda - \omega + \omega_1}{2 \lambda \omega_1 \sin \frac{\alpha}{2}}
\]
(5.203)

So we have
\[
\psi(t) = \frac{\lambda + \omega - \omega_1}{2 \lambda \omega_1 \sin \frac{\alpha}{2}} e^{i \frac{\lambda}{2} t} \begin{pmatrix} \cos \frac{\alpha}{2} + i (\omega - \omega_1) \sin \frac{\alpha}{2} \omega \sin \alpha - \lambda \ e^{i \omega t/2} \\ \cos \frac{\alpha}{2} \ e^{-i \omega t/2} \end{pmatrix} + \frac{\lambda - \omega + \omega_1}{2 \lambda \omega_1 \sin \frac{\alpha}{2}} e^{-i \frac{\lambda}{2} t} \begin{pmatrix} \cos \frac{\alpha}{2} + i (\omega + \omega_1) \sin \frac{\alpha}{2} \omega \sin \alpha - \lambda \ e^{i \omega t/2} \\ \cos \frac{\alpha}{2} \ e^{-i \omega t/2} \end{pmatrix}
\]
(5.204)
\[
\begin{align*}
&= \left( \cos \frac{\lambda t}{2} + i \frac{\omega}{\lambda} \sin \frac{\lambda t}{2} \right) \cdot \cos \frac{\alpha}{2} \cdot e^{-i \omega t/2} \\
&= \left( \cos \frac{\lambda t}{2} - i \frac{\omega}{\lambda} \sin \frac{\lambda t}{2} \right) \cdot \cos \frac{\alpha}{2} \cdot e^{i \omega t/2}
\end{align*}
\]

Exactly the same as the solution that we shown above

\[
\psi(t) = \left( \cos \frac{\lambda t}{2} - i \frac{\omega}{\lambda} \sin \frac{\lambda t}{2} \right) \cos \frac{\alpha}{2} \cdot e^{i \omega t/2}
\]

(5.205)

### 5.5.7. Probability for the system to change to another state

For \( H(t) \) at time \( t \), the two eigenstates are

\[
\chi_+ = \left( \cos \frac{\alpha}{2} \right) \cdot e^{i \omega t \sin \frac{\alpha}{2}}
\]

(5.206)

and

\[
\chi_- = \left( \sin \frac{\alpha}{2} \right) \cdot e^{-i \omega t \cos \frac{\alpha}{2}}
\]

(5.207)

We start from state \( \chi_- \) at \( t = 0 \). The probability to find the system in \( \chi_- \) at a later time \( t \) is

\[
P_- = \left| \langle \chi_- | \psi(t) \rangle \right|^2 = \left| \left( \sin \frac{\alpha}{2} \right) \cdot e^{-i \omega t \cos \frac{\alpha}{2}} \right| ^2 \\
= \left| \left( \cos \frac{\lambda t}{2} - i \frac{\omega}{\lambda} \sin \frac{\lambda t}{2} \right) \cos \frac{\alpha}{2} \cdot e^{-i \omega t/2} \right|^2 \\
= \left| \cos \frac{\lambda t}{2} - i \frac{\omega}{\lambda} \sin \frac{\lambda t}{2} \right|^2 \left| \cos \frac{\alpha}{2} \right|^2 \left| e^{-i \omega t/2} \right|^2 \\
= 2 \frac{\omega}{\lambda} \sin \frac{\lambda t}{2} \sin \frac{\alpha}{2} \cos \frac{\alpha}{2} \\
= 4 \left( \frac{\omega}{\lambda} \right)^2 \sin^2 \frac{\lambda t}{2} \cos^2 \frac{\alpha}{2} \\
= \left( \frac{\omega}{\lambda} \sin \alpha \right)^2 \sin \frac{\alpha}{2} \cos \frac{\alpha}{2} \\
= \left( \omega \sin \alpha \right)^2 \sin \frac{\lambda t}{2}
\]

(5.208)

Although we start from \( \chi_- \) at \( t = 0 \), at time \( t \), there is a nonzero probability for the system to become \( \chi_+ \).

Although we start from spin parallel to \( B \) at \( t = 0 \), at time \( t \), there is a nonzero probability for the system to become anti-parallel to \( B \).

This phenomenon (system start from a quantum state and ends at a different quantum state, by applying a \( t \)-dependent \( H \)) is called a quantum tunneling.

The probability to find the system in \( \chi_+ \) is

\[
P_+ = \left| \langle \chi_+ | \psi(t) \rangle \right|^2
\]

(5.209)

This is the probably for the system to remain in the same state \( (\chi_-) \). It requires no calculation to show that \( P_+ = 1 - P_- \), because the total probability must be 1. This is indeed the case.
If we rotate the Hamiltonian very slowly (\(\omega \ll \omega_1\)), the probability for quantum tunneling vanishes

\[
P_\tau = 1 - \left\langle \chi_\alpha | \psi(t) \right\rangle^2 = \left| \left( \cos \frac{\lambda t}{2} - i \frac{\omega - \omega_1}{\lambda} \sin \left( \frac{\lambda t}{2} \right) \right) e^{i\omega t/2} \right|^2
\]

\[
= \left| \cos \left( \frac{\lambda t}{2} \right) - i \frac{\omega - \omega_1}{\lambda} \sin \left( \frac{\lambda t}{2} \right) \right|^2 e^{-\omega t/2} + \left| \cos \left( \frac{\lambda t}{2} \right) - i \frac{\omega_1 + \omega}{\lambda} \sin \left( \frac{\lambda t}{2} \right) \right|^2 e^{i\omega t/2}
\]

\[
= \cos^2 \left( \frac{\lambda t}{2} \right) + \left( \frac{\omega \cos \alpha - \omega_1}{\lambda} \right)^2 \sin^2 \left( \frac{\lambda t}{2} \right)
\]

\[
= 1 - \sin^2 \left( \frac{\lambda t}{2} \right) + \frac{\omega \cos \alpha - \omega_1}{\lambda} \sin \left( \frac{\lambda t}{2} \right)
\]

\[
= 1 - \frac{\omega^2 - (\omega \cos \alpha)^2}{\lambda^2} \sin \left( \frac{\lambda t}{2} \right)
\]

Because \(\lambda = \sqrt{\omega^2 + \omega_1^2} - 2 \omega \omega_1 \cos \alpha\)

\[
P_\tau = 1 - \frac{\omega^2 - (\omega \cos \alpha)^2}{\lambda^2} \sin \left( \frac{\lambda t}{2} \right) = 1 - \frac{\omega^2 \sin^2 \alpha}{\omega^2 + \omega_1^2 - 2 \omega \omega_1 \cos \alpha} \sin \left( \frac{\lambda t}{2} \right)
\]

(5.211)

### 5.6. Adiabatic theorem

If we rotate the \(B\) field very slowly (\(\omega \ll \omega_1\)), the probability for quantum tunneling vanishes

\[
P_\tau = \left( \frac{\omega^2}{\lambda} \sin \alpha \right)^2 \sin \left( \frac{\lambda t}{2} \right) = \frac{\omega^2 \sin^2 \alpha}{\omega^2 + \omega_1^2 - 2 \omega \omega_1 \cos \alpha} \sin \left( \frac{\lambda t}{2} \right)
\]

which vanishes when \(\omega \to 0\).

In other words, if we changes \(H\) really slowly, the system will remain in the same quantum state.

If we start from the ground state, the system will always in the ground state.

If we start from the nth excited state, it will remain in the nth excited states.

etc.

“Very slowly” in physics is called “adiabatically”.

Conclusion: for adiabatic procedures (i.e., changing \(H\) really slowly), the system will remain in the same quantum state.

Comments:

(1) The rigorous proof can be found in the textbook, which utilize time-dependent perturbation theory

(2) How slow is “slow”?

Assume that we start from state \(n\). The energy difference between this state and other states is \(\Delta E = | E_n - E_m |\). For the smallest \(\Delta E\), we define a time scale \(\tau = \hbar/\Delta E\). “Slow” here means that when we change the Hamiltonian, any change need to take much longer than then \(\tau\).

In the example above, we need to require the rate of change \(\omega\) to be much slower than \(\frac{1}{\tau} = \frac{\Delta E}{\hbar}\). For that example, \(\Delta E = \hbar \omega_1\), so it means that \(\omega \ll \omega_1\).
5.7. Berry phase

The adiabatic theorem tells us that if we are slow and if we start from an quantum eigenstate of \( H(t = 0) \), we will always stay on the same quantum state. But it doesn’t mean that the wavefunction will remain the same. There could be a phase factor.

If we have a time-dependent Hamiltonian \( H(t) \). At each time, we can compute its eigenenergies and eigenwavefunctions. Now we focus on a specific eigenstates (say the \( n \)th excited state), whose eigenenergy is \( E_n \) and the eigenwavefunction is \( | \chi_n(t) \rangle \).

The adiabatic theorem told us that if we change the Hamiltonian really slowly, then when we start from the state \( | \chi_n(t) \rangle \), we will always stay on the same quantum state. i.e., if we have initial condition

\[
| \psi(t = 0) \rangle = | \chi_n(t = 0) \rangle
\]

then at later time

\[
| \psi(t) \rangle = e^{i\theta(t)} \chi_n(t)
\]

5.7.1. Example (same example discussed above)

If we rotate the B field really slowly, the system will always stay on the same quantum state. Say, we start from spin S parallel to B at \( t = 0 \). If we rotate B field extremely slowly, then at a later time \( t > 0 \), the spin will still be parallel to B and it will never become anti parallel.

Q: At time \( t \), the B field comes back to the initial direction (same as at \( t = 0 \)). What is the wavefunction?

A: It must be the same wavefunction up to a phase factor

\[
\psi(t = 2\pi/\omega) = \psi(t = 0) e^{i\phi}
\]

Q: What is this \( \phi \)?

A: it contains two parts

\[
\phi = \phi_1 + \phi_2
\]

First part, \( \phi_1 \) is a dynamic phase factor. For a static problem (\( H \) independent of \( t \)), we know that if we start from an eigenstate of \( H \), the time evolution is

\[
\psi_n(t) = \psi_n e^{-iE_n t/\hbar}
\]

The phase factor here \(-E_n t/\hbar\) is our \( \phi_1 \).

\[
\phi_1 = -E_n t/\hbar
\]

Comment #1: \( \phi_1 \) always arises in a quantum system (no matter \( H \) depends on time \( t \) or not).

Comment #2: \( \phi_1 \) is not a constant number. It depends on \( t \).

The second part \( \phi_2 \) is what we called the Berry phase. It has a geometric original and thus is also known as the geometric phase.

For the example discussed above, we can show that

\[
\psi(t) = A(t) \chi_+(t) + B(t) \chi_-(t)
\]

where

\[
\psi(t) = \begin{bmatrix}
\cos \left( \frac{\lambda t}{2} \right) - i \frac{\sin \omega t}{\lambda} \sin \left( \frac{\lambda t}{2} \right) \\
\cos \left( \frac{\lambda t}{2} \right) - i \frac{\sin \omega t}{\lambda} \sin \left( \frac{\lambda t}{2} \right)
\end{bmatrix}
\]

\[
\chi_+ = \begin{bmatrix}
\cos \frac{\alpha}{2} \\
e^{i\omega t} \sin \frac{\alpha}{2}
\end{bmatrix}
\]

\[
\chi_- = \begin{bmatrix}
\sin \frac{\alpha}{2} \\
-e^{i\omega t} \cos \frac{\alpha}{2}
\end{bmatrix}
\]

and the coefficients
\[ A = \left[ \cos\left(\frac{\lambda t}{2}\right) + i \frac{\omega \cos \alpha - \omega_1}{\lambda} \sin\left(\frac{\lambda t}{2}\right) \right] e^{-i \omega t/2} \]  
\[ B = i \left[ \frac{\omega}{\lambda} \sin\left(\frac{\lambda t}{2}\right) \right] e^{-i \omega t/2} \sin \alpha \]  

When we are doing everything really slowly, \( \omega \rightarrow 0 \), \( B = 0 \), and thus

\[ \psi(t) = A(t) \chi_s(t) \]

For \( \omega \rightarrow 0 \),

\[ \lambda = \sqrt{\omega^2 + \omega_1^2 - 2 \omega \omega_1 \cos \alpha} = \omega_1 - \cos \alpha \omega \]

As a result,

\[ A \approx \cos\left(\frac{\lambda t}{2}\right) + i \frac{\omega \cos \alpha - \omega_1}{\lambda} \sin\left(\frac{\lambda t}{2}\right) e^{-i \omega t/2} \approx \cos\left(\frac{\lambda t}{2}\right) + i \frac{\omega_1}{2} e^{-i \omega t/2} \]

we have

\[ \psi(t) = e^{-i \omega t/2} \chi_s(t) = e^{-i \left(\frac{\omega_1}{2} - \frac{\cos \alpha}{2} \omega t\right)} \chi_s(t) = e^{i \left(\frac{\omega_1}{2} - \frac{\cos \alpha}{2} \omega t\right)} \chi_s(t) \]

The phase factor contains two parts. The first part \(-\frac{\omega_1}{2} t\) is the dynamic phase factor. Because the energy of the \( \chi_s \) state is \( E_s = +\hbar \omega_1/2 \), we expect the dynamic phase factor to be \(-E_s t/h = -\frac{\omega_1}{2} t\)

The second part is our Berry’s phase

\[ \varphi_2 = -\frac{1 - \cos \alpha}{2} \omega t \]

At \( t = 2\pi/\omega \) (i.e. when the B field goes back to the initial direction),

\[ \varphi_2 = -\frac{1 - \cos \alpha}{2} \frac{2\pi}{\omega} = \pi \cos \alpha - 1 \]

When the B field rotate back to the initial direction, the Berry’s phase is a constant, independent of \( \omega \) (of course we need to assume \( \omega \) is very small, but the value of \( \omega \) doesn’t matter).

Comment #1: the Berry phase only arise when \( H \) changes.

Comment #2: the Berry phase is independent of details, e.g. how fast we rotate the B field, as long as it is slow enough; or how strong the B field is; or what is the value of the g-factor.

5.7.2. Geometric meaning

We can define a unit vector

\[ \vec{n}_B = \frac{\vec{B}}{\|\vec{B}\|} \]

This unit vector is one point on the unit sphere. When we rotate the direction of \( \vec{B} \), this unit vector draws a line on the unit sphere. When \( \vec{B} \) finally comes back to the initial direction, the unit vector draws a close loop on the unit sphere.

This close loop enclose an area on the unit sphere, whose area is \( \Omega_B \). The Berry’s phase is nothing but

\[ \varphi_B = -s \Omega_B \]

where \( s \) is the spin quantum number. For electrons \( s = 1/2 \). For general quantum particles, \( s \) is integer (bosons) or half integer (fermions).

For the case we discussed above, when we rotate B around the z axis, the \( \vec{n}_B \) draws a circle on the unit sphere. The (spherical) area enclosed by this circle is
\[ \Omega_B = \int_0^\alpha 2 \pi \sin \theta \, d\theta = -2 \pi \cos \theta \bigg|_{\theta=\alpha} = 2 \pi (1 - \cos \alpha) \]  

(5.233)

Thus,

\[ \varphi_B = -s \Omega_B = \frac{1}{2} \times 2 \pi (1 - \cos \alpha) = \pi (\cos \alpha - 1) \]  

(5.234)

Comment #1: In the example, we assume that \( \vec{n}_B \) draws a perfect circle on the unit sphere. If the path is not a circle, all the conclusions remains the same.

Comment #2: We need to put a +/- sign in front of \( \Omega_B \). The sign is determined by the direction (orientation) of the path. This is because, if we reverse the path (reverse everything), the phase need to flip sign. The convention that we take here is, we use choose one point inside the circle as our reference point. When \( \vec{n}_B \) draw the circle, the ending point of \( \vec{n}_B \) moves around my reference point. This motion defines an angular momentum (relative to my reference point). If the angular momentum is positive (pointing from the center of the unit sphere to outside), \( \Omega_B > 0 \). Otherwise, it is negative.

5.7.3. Quantization of spins

One thing the Berry phase tells us is that spins must quantize to integer or half-integer values. To see that, we look at the circle discussed above.

If we sit at the north pole, we find that \( \vec{n}_B \) draws a circle around and the angular momentum is positive (pointing up). The area of the circle is \( \Omega_B = 2 \pi (1 - \cos \alpha) \). The + sign here means that the angular motion has angular momentum along the positive direction (up).

If we sit at the south pole, we will also find that \( \vec{n}_B \) draws a circle around, but the angular momentum is negative (pointing down, relative to the south pole).

NOTE: the angular momentum is always along +z direction. For north pole, +z is up (from the center of the sphere to outside). But for south pole, +z is down, pointing from outside the sphere towards the center.

Here, if we sit a the south pole, we find that the area enclosed by the circle is negative

\[ \Omega_B' = -\int_0^\alpha 2 \pi \sin \theta \, d\theta = 2 \pi \cos \theta \bigg|_{\theta=\alpha} = 2 \pi (-1 + \cos \alpha) = -2 \pi (1 + \cos \alpha) \]  

(5.235)

It is easy to notice that

\[ \Omega_B - \Omega_B' = | \Omega_B | + | \Omega_B' | = \text{area of the whole unit sphere} = 4 \pi \]  

(5.236)

Now, we have two different Berry phases

\[ \varphi_B = -s \Omega_B \]  

(5.237)

and

\[ \varphi_B' = -s \Omega_B' = -s (4 \pi + \Omega_B) = -4 \pi s - s \Omega_B \]  

(5.238)

This two phases differ by 4\( \pi s \). But we know that they cannot be different (there is only one unique Berry phase). The only possible option here is that they different by 2\( \pi n \)

\[ 4 \pi s = 2 \pi n \]  

(5.239)

So

\[ s = n / 2 \]  

(5.240)

If \( n \) is even, \( s \) is an integer.

If \( n \) is odd, \( s \) is a half integer.

5.7.4. A generic way to compute Berry phase \( \varphi_B \)

If we have a time-dependent Hamiltonian \( H(t) \). At each time, we can compute its eigenenergies and eigenwavefunctions. Now we focus on a specific eigenstates (say the nth excited state), whose eigenenergy is \( E_n \) and the eigenwavefunction is \( | \chi_n(t) \rangle \).

The adiabatic theorem told us that if we change the Hamiltonian really slowly, then when we start from the state \( | \chi_n(t) \rangle \), we will always stay
on the same quantum state, i.e., if we have initial condition
\[ |\psi(t = 0)\rangle = |\chi_0(t = 0)\rangle \] (5.241)
then at later time
\[ |\psi(t)\rangle = e^{i \phi(t)} |\chi_0(t)\rangle \] (5.242)
This phase \( \phi \) contains two parts
\[ \phi(t) = -\frac{E_0(t)}{\hbar} t + \phi_0(t) \] (5.243)
The first part is the dynamic phase factor. The second part is the Berry phase
\[ \phi_0(t) = \int_0^t d\tau \left( \chi_0(\tau) | i \frac{d}{d\tau} | \chi_0(\tau) \right) \] (5.244)
For a time-dependent Hamiltonian, we change the Hamiltonian by making some of the control parameters time-dependent (e.g. \( B_x \) and \( R_j \)). Let’s use \( R_1(\tau), R_2(\tau), \ldots, R_n(\tau) \) to represent these time-dependent parameters.
\[
\frac{d}{d\tau} | \chi(\tau) \rangle = \frac{\partial | \chi(\tau) \rangle}{\partial R_1} \frac{dR_1(\tau)}{d\tau} + \frac{\partial | \chi(\tau) \rangle}{\partial R_2} \frac{dR_2(\tau)}{d\tau} + \cdots + \frac{\partial | \chi(\tau) \rangle}{\partial R_n} \frac{dR_n(\tau)}{d\tau}
\] (5.245)
\[
\phi_0(t) = \int_0^t d\tau \frac{dR_1(\tau)}{d\tau} \langle \chi(\tau) | i \frac{\partial}{\partial R_1} | \chi(\tau) \rangle + \int_0^t d\tau \frac{dR_2(\tau)}{d\tau} \langle \chi(\tau) | i \frac{\partial}{\partial R_2} | \chi(\tau) \rangle + \cdots + \int_0^t d\tau \frac{dR_n(\tau)}{d\tau} \langle \chi(\tau) | i \frac{\partial}{\partial R_n} | \chi(\tau) \rangle
\] (5.246)
\[
= \int_\mathcal{E} dR \cdot \chi(\tau) | i \nabla_R | \chi(\tau) \rangle
\] (5.247)
Here, all possible values of control parameters form a \( n \)-dimensional space, with coordinates \( \vec{R} = (R_1, R_2, \ldots, R_n) \). Each point in this space gives a set of values for my control parameters. For a time-dependent \( H \), it corresponds to a curve in this space, \( \vec{R}(\tau) \). For each time \( \tau \), the value of my control parameters is one point in the space with coordinate \( \vec{R}(\tau) \). \( \nabla_R \) is the \( n \)-dimensional gradient in this space and \( \int_\mathcal{E} d\vec{R} \) is a line integral in this space along the curve of \( \vec{R}(\tau) \).
If the Hamiltonian at time \( t = T \) goes back to the initial Hamiltonian at \( t = 0 \), this integral becomes a loop integral before for \( t = T, R_j = R_i \)
\[ \phi_0(T) = \oint d\vec{R} \cdot (\chi(\tau) | i \nabla_R | \chi(\tau) \rangle) \] (5.248)
This is really similar to the relation between magnetic flux and the vector potential
\[ \Phi_B = \oint d\vec{A} \cdot \vec{B} \] (5.249)
Remember, \( \Phi_B \) also introduces a phase factor for wavefunctions (the AB effect), there is in fact a very strong analogy between Berry phase and magnetic flux. In some sense, the Berry phase is a “magnetic flux” in the parameter space.

**Comment #1:** The general proof can be found in the textbook.

**Comment #2:** only works for adiabatic procedures. As we will show below, if we are not slow, the wavefunction will NOT satisfy \( |\psi(t)\rangle = e^{i \phi(t)} |\chi_0(t)\rangle \). Thus, it makes no sense to take about phase factor.

### 5.7.5. Back to our example

In our example, there are two control parameters \( B_x \) and \( B_y \), i.e. we have a 2D parameter space. As time changes, we get a curve in this space
\[ (B_x, B_y) = (B_0 \sin \alpha \cos \omega t, B_0 \sin \alpha \sin \omega t) \] (5.249)
which is a circle around the origin with radius \( B_0 \sin \alpha \). This circle defines my loop integral \( \oint d\vec{R} \).

My wavefunction
Here, we used the fact that \( B_0 \sin \alpha = \sqrt{B_x^2 + B_y^2} \)

\[
\langle \chi_0 (\tau) | \hat{\mathbf{r}} \rangle = \langle \chi_0 (\tau) | \mathbf{r} \rangle \approx \left( \frac{\cos \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \right) \left( \frac{\cos \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \right) = \left( \frac{\cos \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \right) \left( \frac{\cos \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \right)
\]

Thus

\[
\langle \chi_0 (\tau) | \hat{\mathbf{r}} \rangle = \left( \frac{\cos \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \right) \left( \frac{\cos \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \right) = \left( \frac{\cos \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \right) \left( \frac{\cos \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \right)
\]

The second component

\[
\langle \chi_0 (\tau) | \hat{B}_x \rangle \langle \chi_0 (\tau) \rangle = \left( \frac{\sin \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \right) \left( \frac{\sin \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \right) = \left( \frac{\sin \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \right) \left( \frac{\sin \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \right)
\]

So

\[
\langle \chi_0 (\tau) | \hat{\mathbf{r}} \rangle = \frac{\sin^2 \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} (B_x, -B_y) = -\frac{\sin^2 \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \hat{\phi}
\]

Here, we convert to the polar coordinate \( r \) and \( \phi \). And

\[
\hat{\phi} = \left( \frac{\hat{r}}{\sqrt{B_x^2 + B_y^2}}, -\frac{\hat{B}_x}{\sqrt{B_x^2 + B_y^2}} \right)
\]

is the unit vector along the circular direction

\[
\phi(t) = \int_0^{2\pi} d\phi \langle \chi_0 (\tau) | \hat{\mathbf{r}} \rangle = \int_0^{2\pi} d\phi \left( \frac{\sin^2 \frac{\alpha}{2}}{\sqrt{B_x^2 + B_y^2}} \right) = -\frac{\sin^2 \frac{\alpha}{2}}{2} \int_0^{2\pi} d\phi = -2\pi \frac{\sin^2 \alpha}{2} = -\pi (1 - \cos \alpha) = \pi (\cos \alpha - 1)
\]

Same as what we found above from directly solving for \( \psi(t) \).
5.8. Resonance and transition rate

5.8.1. Non-adiabatic regime

If we change B fast enough (not slowly), then the system could jump/tunnel to a different quantum state

\[ P_\downarrow > 0 \]  \hspace{1cm} (5.258)

Here, we can ask the question, which \( \omega \) maximizes the probability for the system to change from one state to another?

Here, for simplicity, we will only focus on the case of small \( \alpha \).

\[ \vec{B} = (B_0 \sin \alpha \cos \omega t, B_0 \sin \alpha \sin \omega t, B_0 \cos \alpha) \]  \hspace{1cm} (5.259)

At small \( \alpha \),

\[ \vec{B} \approx (\alpha B_0 \cos \omega t, \alpha B_0 \sin \omega t, B_0) \]  \hspace{1cm} (5.260)

The x and y components are much smaller than the z component.

Notice that the x and y components depends on time, where the z component is independent of time, this small \( \alpha \) regime is the case that time independent part dominate and the time-dependent part is small and can be treat as a small perturbation.

\[ H = \frac{|e|}{m} \begin{bmatrix} B_x \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + B_y \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + B_z \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{bmatrix} \]

\[ = \frac{|e|}{m} \begin{bmatrix} B_x \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + B_y \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + B_z \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \end{bmatrix} \]

\[ \approx \frac{|e|}{m} \begin{bmatrix} B_0 \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \alpha \frac{|e|}{m} B_0 \frac{\hbar}{2} \begin{bmatrix} \cos \omega \left( \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \sin \omega \left( \begin{pmatrix} 0 & 0 \\ 0 & -i \end{pmatrix} \right) \right) \end{bmatrix} \end{bmatrix} \]

We can rewrite the Hamiltonian as

\[ H(t) = H_0 + \alpha H(t) \]

where

\[ H_0 = \frac{|e|}{m} \begin{bmatrix} B_0 \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{bmatrix} \]

is the unperturbed Hamiltonian and it is independent of time.

\[ H(t) = \frac{|e|}{m} \begin{bmatrix} B_0 \frac{\hbar}{2} \begin{pmatrix} \cos \omega \left( \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \sin \omega \left( \begin{pmatrix} 0 & 0 \\ 0 & -i \end{pmatrix} \right) \right) \end{bmatrix} \end{bmatrix} = \frac{\hbar \omega_1}{2} \begin{pmatrix} 0 & e^{i\omega \tau} \\ e^{-i \omega \tau} & 0 \end{pmatrix} \]

depends on time. Because \( \alpha \) is small. This part is a small time-dependent perturbation.

At small \( \alpha \), we have a time-independent system with Hamiltonian \( H_0 \). And we apply a small perturbation, which depends on time \( t \), \( \alpha H(t) \).

\( H_0 \) has two eigenstates, spin up and down, with eigenenergy \( \pm \frac{\hbar \omega_1}{2} \) and \( -\frac{\hbar \omega_1}{2} \). If we start from the spin up state at \( t = 0 \), at time \( t \) the probability for the system to be in the other state is

\[ P_\downarrow = \frac{\omega_1^2 \sin^2 \alpha}{\omega^2 + \omega_1^2 - 2 \omega \omega_1 \cos \alpha} \sin \left( \frac{\lambda t}{2} \right) \]

\[ \lambda = \sqrt{\omega^2 + \omega_1^2 - 2 \omega \omega_1 \cos \alpha} = \sqrt{\omega^2 + \omega_1^2 - 2 \omega \omega_1} = | \omega - \omega_1 | \]

\[ P_\downarrow \approx \frac{\omega_1^2}{\omega^2 + \omega_1^2 - 2 \omega \omega_1} \sin^2 \left( \frac{\omega - \omega_1 | t |}{2} \right) = \frac{\omega^2}{(\omega - \omega_1)^2} \sin^2 \left( \frac{|\omega - \omega_1 | t }{2} \right) = \alpha^2 \omega^2 \left( \frac{\sin \left( \frac{(\omega - \omega_1 \tau)}{2} \right)}{\omega - \omega_1} \right)^2 \]

\[ \text{For any fixed } t, \text{ the maximum of this function is at } \omega = \omega_1. \]

Remember, \( H_0 \) has two eigenstates. And we are considering the transition between them. The energy difference between there two states is \( \Delta E = E_+ - E_- = \hbar \omega_1 / 2 \). Thus, the maximum transition probability happens at
This conclusion turns out to be universal. If we have a time-independent system with multiple energy levels. We prepare the system in one eigenstate with energy $E_i$, and then we apply a time-dependent perturbation with frequency $\omega$. If we want the system to turn into a different quantum state with energy $E_f$, the best frequency that we shall apply is $\omega = |E_f - E_i|/\hbar$.

This is what we called a resonance.

A resonance in quantum mechanics is not something really surprising. In classical mechanics, we have similar phenomenon. If a system has some intrinsic oscillation frequency $\omega_0$ (say a harmonic oscillator), then the system will response very strongly to external perturbations, if the frequency of the perturbation is the same as $\omega_0$. In quantum mechanics, the intrinsic frequency $\omega_0$ is the energy difference between to eigenstates of $H$ divided by $\hbar$.

### 5.8.2. Transition rate

If we apply the perturbation for longer time, the probability for the transition increases. So to be fair, we shall define the transition rate

$$ R(t) = \frac{dP(t)}{dt} = \frac{d}{dt} \alpha^2 \omega^2 \left( \frac{\sin \left( \frac{\omega - \omega_1}{2} \right)}{\omega - \omega_1} \right)^2 = \alpha^2 \omega^2 \frac{2 \sin \left( \frac{\omega - \omega_1}{2} \right) \cos \left( \frac{\omega - \omega_1}{2} \right)}{(\omega - \omega_1)^2} \frac{(\omega - \omega_1)}{2} $$

$$ = \alpha^2 \omega^2 \frac{\sin((\omega - \omega_1)t)}{2 (\omega - \omega_1)} $$

In the large $t$ limit (i.e., if we wait for long enough time)

$$ \lim_{t \to \infty} R(t) = \frac{\alpha^2 \omega^2}{2} \lim_{t \to \infty} \frac{\sin((\omega - \omega_1)t)}{(\omega - \omega_1)} = \frac{\alpha^2 \omega^2}{2} \pi \delta(\omega - \omega_1) $$

In the long time limit, the transition rate is a constant (independent of time)

$$ R = \frac{\alpha^2 \omega^2}{2} \pi \delta(\omega - \omega_1) = \frac{\alpha^2 \omega_1^2}{2} \pi \delta(\omega - \omega_1) $$

Here, the delta function tells us that if we want the transition to happen, when we wait for long enough time, we will find that the frequency of the perturbation have to match $|E_f - E_i|/\hbar$.

**Comment A**: the infinite that comes with the $\delta$ function is not a problem, for two reasons

1. This is transition rate not probability. The probability is the integral of the transition rate $P = \int dt R(t)$. As long as the integral is finite (and probability smaller than 1), it is safe.

2. The transition rate above only consider perturbation with a unique frequency. In reality, we cannot really have a unique frequency. The energy must be distributed in some energy range. For example, if we shine light on a system, the light cannot really has one single frequency, the intensity must have some distribution

$$ I = \int d\omega \rho(\omega) $$

where $I$ is the total intensity of the light, and $\rho(\omega)$ is the intensity density at some frequency. There, the total transition rate is

$$ R = \int d\omega R(\omega) \rho(\omega) = \int d\omega \rho(\omega) \frac{\alpha^2 \omega^2}{2} \pi \delta(\omega - \omega_1) = \frac{\alpha^2 \omega_1^2}{2} \pi \rho(\omega_1) $$

After the frequency integral, $R$ is no longer infinity.

**Comment B**: the coefficient $\frac{\alpha^2 \omega_1^2}{2} \pi$ has one clear physical meaning

$$ \langle f | a H' | i \rangle = \left( \begin{array}{c} 0 \\ 1 \end{array} \right) \frac{\hbar \omega_1}{2} \left( \begin{array}{c} e^{-i \omega t} \\ 0 \end{array} \right) \left( \begin{array}{c} 1 \\ 0 \end{array} \right) = \frac{\hbar \omega_1}{2} e^{i \omega t} $$

Here $| i \rangle = (1, 0)$ is the initial state (spin up state at $t = 0$) and $| f \rangle = (0, 1)$ is the final state that they system will transit to at some later time $t > 0$. We find that

$$ \omega = \Delta E / \hbar $$
\[
\langle f | a H' | i \rangle^2 = \left( \frac{\hbar \omega_i}{2} \right)^2
\]

(5.274)

In other words,
\[
R_{i \rightarrow f} = 2 \pi \frac{\omega_i}{\hbar} \delta(\omega - \omega_i) = 2 \pi \left| \langle f | a H' | i \rangle \right|^2 \delta(\omega - \omega_i) = 2 \pi \left| \langle f | a H' | i \rangle \right|^2 \delta \left( \omega - \frac{|E_f - E_i|}{\hbar} \right)
\]

(5.275)

The transition rate from an initial state i to a final state is
\[
R_{i \rightarrow f} = \int d\omega \rho(\omega) R(\omega) = \int d\omega \rho(\omega) \frac{2 \pi}{\hbar} \left| \langle f | a H' | i \rangle \right|^2 \delta \left( \omega - \frac{|E_f - E_i|}{\hbar} \right)
\]

(5.276)

If we wait enough time, the transition will happen and can only happen when the frequency of the perturbation matches the energy difference between the final and initial states.

If our perturbation do not have a unique frequency (in any real experiment, the frequency cannot be really unique), we shall need to use a function \( \rho(\omega) \) to describe the intensity distribution. Here,
\[
R_{i \rightarrow f} = \int d\omega \rho(\omega) R(\omega) = \int d\omega \rho(\omega) \frac{2 \pi}{\hbar} \left| \langle f | a H' | i \rangle \right|^2 \rho \left( \frac{|E_f - E_i|}{\hbar} \right)
\]

(5.277)
In this chapter, we study time-dependent Hamiltonians with the following structure
\[ H = H_0 + \lambda H'(t) \] (6.1)
and we assume that \( H_0 \) is much larger than \( \lambda H'(t) \), i.e. \( \lambda \ll 1 \).

In the previous chapter, we examined a specific example (spin-1/2 in a rotation magnetic field), and find that time-dependent perturbation leads to transitions between different eigenstates. In this chapter, we will show that all the conclusion that we obtained early on hold for general time-dependent perturbations (regardless of the details).

### 6.1. Step one: solving the time-independent part \( H_0 \)

Since \( H'(t) \) is a small perturbation, in the first step, we start our investigation by ignoring this small \( H'(t) \) part. Here, the Hamiltonian only contains the time-independent part \( H_0 \). For a time-independent Hamiltonian, like \( H_0 \), we can get all the eigenstates and eigenenergies,
\[ H_0 \psi_n = E_n \psi_n \] (6.2)
\[ H_0 \psi_a = E_a \psi_a \] (6.3)
\[ H_0 \psi_b = E_b \psi_b \] (6.4)

Here, we assume that \( E_a \neq E_b \). As a result, we can prove immediately, \( \langle \psi_a | \psi_b \rangle = 0 \).

The reason that we can focus on two eigenstates (and ignore others) is because we are studying here transitions between different states. For a single transition, it only involves a pair of (two) quantum state, the initial state and the final state, and other eigenstates are not very important for this transition. Here, we assume that our state is prepared as \( | \psi_a \rangle \) at time \( t = 0 \), and we want to know what is the probability for this system to be found on state \( | \psi_b \rangle \) latter.

### 6.2. Step two: write down the wavefunction

Because \( | \psi_a \rangle \) forms a complete basis, any quantum state can be written as their superpositions:
\[ | \psi \rangle = \sum_a c_a | \psi_a \rangle \] (6.5)

In particular, what we want to know here the solution of the time-dependent Schrodinger’s equation
\[ \frac{\hat{\partial}}{\hat{\partial}t} | \psi(t) \rangle = H(t) | \psi(t) \rangle \] (6.6)

Because any state can be written as superpositions of \( | \psi_a \rangle \), the solution to the time-dependent Schrodinger’s equation, which is also a quantum state, cannot be an exception,
\[ |\psi(t)\rangle = \sum_n \tilde{c}_n(t) \left| \psi_n \right\rangle \]  

(6.7)

Because \( |\psi(t)\rangle \) depends on time \( t \), the coefficients \( c_n(t) \) are functions of \( t \). In many cases, it is more convenient to define a new set of function as the coefficient

\[ c_n(t) = \tilde{c}_n(t) e^{iE_n t / \hbar} \]  

(6.8)

Equivalently,

\[ \tilde{c}_n(t) = c_n(t) e^{-iE_n t / \hbar} \]  

(6.9)

As a result,

\[ |\psi(t)\rangle = \sum_n c_n(t) e^{-iE_n t / \hbar} \left| \psi_n \right\rangle \]  

(6.10)

The advantage of using \( c_n \), instead of \( \tilde{c}_n \) is because, if \( H'_{(t)} = 0 \), \( c_n \) will become constants (independent of \( t \)).

In the previous section, we have said that we will only focus on two quantum states, \( |\psi_a\rangle \) and \( |\psi_b\rangle \), so here, we ignore all other states here, i.e. we set \( c_n(t) = 0 \) unless \( n \) is \( a \) or \( b \)

\[ |\psi(t)\rangle = c_a(t) e^{-iE_a t / \hbar} \left| \psi_a \right\rangle + c_b(t) e^{-iE_b t / \hbar} \left| \psi_b \right\rangle \]  

(6.11)

The normalization condition tells us \( \langle \psi(t) | \psi(t) \rangle = 1 \), so

\[ \langle \psi(t) | \psi(t) \rangle = | c_a(t) |^2 + | c_b(t) |^2 = 1 \]  

(6.12)

The initial condition says that at \( t = 0 \), our system is prepared to be in the \( |\psi_a\rangle \) state, i.e., \( |\psi(t = 0)\rangle = |\psi_a\rangle \), so

\[ c_a(0) = 1 \quad \text{and} \quad c_b(0) = 0 \]  

(6.13)

### 6.3. Step three: rewrite time-dependent Schrödinger equation

\[ i\hbar \frac{\partial}{\partial t} \left| \psi(t) \right\rangle = H(t) \left| \psi(t) \right\rangle \]  

(6.14)

And we know that

\[ |\psi(t)\rangle = c_a(t) e^{-iE_a t / \hbar} \left| \psi_a \right\rangle + c_b(t) e^{-iE_b t / \hbar} \left| \psi_b \right\rangle \]  

(6.15)

Normalization condition

\[ \langle \psi(t) | \psi(t) \rangle = | c_a(t) |^2 + | c_b(t) |^2 = 1 \]  

(6.16)

Initial condition

\[ c_a(0) = 1 \quad \text{and} \quad c_b(0) = 0 \]  

(6.17)

How do we solve this Schrodinger equation? Let's start from the l.h.s.,

\[ i\hbar \frac{\partial}{\partial t} \left| \psi(t) \right\rangle = i\hbar \frac{\partial}{\partial t} \left[ c_a(t) e^{-iE_a t / \hbar} \left| \psi_a \right\rangle + c_b(t) e^{-iE_b t / \hbar} \left| \psi_b \right\rangle \right] \]

\[ = i\hbar \left[ \frac{d}{dt} \left( c_a(t) e^{-iE_a t / \hbar} \left| \psi_a \right\rangle \right) + \frac{d}{dt} \left( c_b(t) e^{-iE_b t / \hbar} \left| \psi_b \right\rangle \right) \right] \]

\[ = i\hbar \left[ \frac{d}{dt} \left( c_a(t) e^{-iE_a t / \hbar} \right) \left| \psi_a \right\rangle + \frac{d}{dt} \left( c_b(t) e^{-iE_b t / \hbar} \right) \left| \psi_b \right\rangle \right] \]

\[ = \left[ i\hbar \frac{d}{dt} c_a(t) e^{-iE_a t / \hbar} + E_a c_a(t) e^{-iE_a t / \hbar} \right] \left| \psi_a \right\rangle + \left[ i\hbar \frac{d}{dt} c_b(t) e^{-iE_b t / \hbar} + E_b c_b(t) e^{-iE_b t / \hbar} \right] \left| \psi_b \right\rangle \]

\[ = \left[ E_a c_a(t) e^{-iE_a t / \hbar} \left| \psi_a \right\rangle + E_b c_b(t) e^{-iE_b t / \hbar} \left| \psi_b \right\rangle \right] + \left[ i\hbar \frac{d}{dt} c_a(t) e^{-iE_a t / \hbar} \left| \psi_a \right\rangle + i\hbar \frac{d}{dt} c_b(t) e^{-iE_b t / \hbar} \left| \psi_b \right\rangle \right] \]

The r.h.s.
\[ H(t) \psi(t) = [H_0 + \lambda H'(t)] c_a(t) e^{-iE_a t/\hbar} \psi_a + c_b(t) e^{-iE_b t/\hbar} \psi_b \]
\[ = H_0 \left[ c_a(t) e^{-iE_a t/\hbar} \psi_a + c_b(t) e^{-iE_b t/\hbar} \psi_b \right] + \lambda H'(t) \left[ c_a(t) e^{-iE_a t/\hbar} \psi_a + c_b(t) e^{-iE_b t/\hbar} \psi_b \right] 
\[ = \left[ c_a(t) e^{-iE_a t/\hbar} H_0 \psi_a + c_b(t) e^{-iE_b t/\hbar} H_0 \psi_b \right] + \lambda \left[ c_a(t) e^{-iE_a t/\hbar} H'(t) \psi_a + c_b(t) e^{-iE_b t/\hbar} H'(t) \psi_b \right] 
\[ = \left[ c_a(t) e^{-iE_a t/\hbar} E_a \psi_a + c_b(t) e^{-iE_b t/\hbar} E_b \psi_b \right] + \lambda \left[ c_a(t) e^{-iE_a t/\hbar} H'(t) \psi_a + c_b(t) e^{-iE_b t/\hbar} H'(t) \psi_b \right] \]

So we have

\[ \left[ E_a c_a(t) e^{-iE_a t/\hbar} \psi_a + E_b c_b(t) e^{-iE_b t/\hbar} \psi_b \right] + \left[ i\hbar \frac{d c_a(t)}{dt} e^{-iE_a t/\hbar} \psi_a + i\hbar \frac{d c_b(t)}{dt} e^{-iE_b t/\hbar} \psi_b \right] = \left[ c_a(t) e^{-iE_a t/\hbar} E_a \psi_a + c_b(t) e^{-iE_b t/\hbar} E_b \psi_b \right] \]

The first [...] on the l.h.s. and the first [...] on the r.h.s. are identical and thus they cancel, so we get

\[ i\hbar \frac{d c_a(t)}{dt} e^{-iE_a t/\hbar} \psi_a + i\hbar \frac{d c_b(t)}{dt} e^{-iE_b t/\hbar} \psi_b = \lambda c_a(t) e^{-iE_a t/\hbar} H'(t) \psi_a + \lambda c_b(t) e^{-iE_b t/\hbar} H'(t) \psi_b \]

We can multiply both sides by \( \langle \psi_a | \),

\[ i\hbar \frac{d c_a(t)}{dt} e^{-iE_a t/\hbar} \langle \psi_a | \psi_a \rangle + i\hbar \frac{d c_b(t)}{dt} e^{-iE_b t/\hbar} \langle \psi_a | \psi_b \rangle = \lambda c_a(t) e^{-iE_a t/\hbar} \langle \psi_a | H'(t) \psi_a \rangle + \lambda c_b(t) e^{-iE_b t/\hbar} \langle \psi_a | H'(t) \psi_b \rangle \]

Because \( \langle \psi_a | \psi_a \rangle = 1 \) and \( \langle \psi_a | \psi_b \rangle = 0 \), we get

\[ i\hbar \frac{d c_a(t)}{dt} e^{-iE_a t/\hbar} = \lambda c_a(t) e^{-iE_a t/\hbar} \langle \psi_a | H'(t) \psi_a \rangle + \lambda c_b(t) e^{-iE_b t/\hbar} \langle \psi_a | H'(t) \psi_b \rangle \]

\[ i\hbar \frac{d c_b(t)}{dt} = \lambda c_a(t) \langle \psi_a | H'(t) \psi_a \rangle + \lambda c_b(t) e^{-i(E_b-E_a) t/\hbar} \langle \psi_a | H'(t) \psi_b \rangle \]

Similarly, we can multiply the equation above by \( \langle \psi_b | \), and we will get,

\[ i\hbar \frac{d c_b(t)}{dt} = \lambda c_b(t) e^{i(E_b-E_a) t/\hbar} \langle \psi_b | H'(t) \psi_a \rangle + \lambda c_b(t) \langle \psi_b | H'(t) \psi_b \rangle \]

At the end of the day, we get a set of two equations:

\[ \begin{align*}
\frac{d c_a(t)}{dt} &= c_a(t) \langle \psi_a | H'(t) \psi_a \rangle + c_b(t) e^{-i(E_b-E_a) t/\hbar} \langle \psi_a | H'(t) \psi_b \rangle \\
\frac{d c_b(t)}{dt} &= c_a(t) e^{i(E_b-E_a) t/\hbar} \langle \psi_b | H'(t) \psi_a \rangle + c_b(t) \langle \psi_b | H'(t) \psi_b \rangle
\end{align*} \]

Let's define

\[ H'_{ab}(t) = \langle \psi_a | H'(t) \psi_b \rangle \]

\[ H'_{ba}(t) = \langle \psi_b | H'(t) \psi_a \rangle \]

\[ H'_{aa}(t) = \langle \psi_a | H'(t) \psi_a \rangle \]

\[ H'_{bb}(t) = \langle \psi_b | H'(t) \psi_b \rangle \]

and define

\[ \omega_0 = \frac{E_b - E_a}{\hbar} \]

So the equations now looks like

\[ \begin{align*}
\frac{d c_a(t)}{dt} &= \lambda H'_{aa}(t) c_a(t) + \lambda e^{i\omega_0 t/\hbar} H'_{ab}(t) c_b(t) \\
\frac{d c_b(t)}{dt} &= \lambda e^{i\omega_0 t/\hbar} H'_{ba}(t) c_a(t) + \lambda H'_{bb}(t) c_b(t)
\end{align*} \]

NOTE: By definition, \( H'_{aa}, H'_{bb}, H'_{ab} \) and \( H'_{ba} \) are numbers. They are no longer quantum operators, i.e., These equations are ordinary differential equations. There is no quantum operator or quantum wavefunction in these equations anymore.

In many cases, \( H'_{aa} \) and \( H'_{bb} \) are zero, so we can drop those two terms:
What should we do, if $H'_{aa}$ and $H'_{bb}$ are not zero?

Let's start from these two equations:

\[
\begin{align*}
\frac{d}{dt} \left( \frac{\partial}{\partial \alpha} H(t) \right) & = \lambda H(t) + \lambda e^{-i \omega t} H(t) c_0(t) \\
\frac{d}{dt} \left( \frac{\partial}{\partial \lambda} H(t) \right) & = \lambda e^{-i \omega t} H(t) c_0(t) + \lambda H(t) c_0(t)
\end{align*}
\]  

(6.33)

Now, we define

\[
d_a(t) = c_a(t) e^{\frac{i}{\hbar} \int H_{aa}(\tau) d\tau}
\]

(6.35)

and

\[
d_b(t) = c_b(t) e^{\frac{i}{\hbar} \int H_{bb}(\tau) d\tau}
\]

(6.36)

We can substitute $c_a$ and $c_b$ by $d_a$ and $d_b$ in the equations

\[
\begin{align*}
\frac{d}{dt} \left[ e^{\frac{i}{\hbar} \int H_{aa}(\tau) d\tau} d_a(t) \right] & = H_{aa}(t) e^{\frac{i}{\hbar} \int H_{aa}(\tau) d\tau} d_a(t) + e^{-i \omega t} H_{ab}(t) e^{\frac{i}{\hbar} \int H_{aa}(\tau) d\tau} d_b(t) \\
\frac{d}{dt} \left[ e^{\frac{i}{\hbar} \int H_{bb}(\tau) d\tau} d_b(t) \right] & = e^{i \omega t} H_{bb}(t) e^{\frac{i}{\hbar} \int H_{bb}(\tau) d\tau} d_b(t) + H_{ba}(t) e^{-i \omega t} H_{bb}(t) e^{\frac{i}{\hbar} \int H_{bb}(\tau) d\tau} d_b(t)
\end{align*}
\]  

(6.37)

For the first equation, the derivative on the l.h.s. can be written as

\[
\frac{d}{dt} \left[ e^{\frac{i}{\hbar} \int H_{aa}(\tau) d\tau} d_a(t) \right] = \frac{d}{dt} e^{\frac{i}{\hbar} \int H_{aa}(\tau) d\tau} \frac{dd_a(t)}{dt} + \left( e^{\frac{i}{\hbar} \int H_{aa}(\tau) d\tau} \right) \frac{dd_a(t)}{dt}
\]

(6.38)

The first term cancels the first term on the r.h.s., so the first equation becomes

\[
 i e^{-i \omega t} H'_{ab}(t) e^{\frac{i}{\hbar} \int H_{aa}(\tau) d\tau} d_b(t)
\]

Thus

\[
i \hbar \frac{d}{dt} d_a(t) = e^{-i \omega t} H'_{ab}(t) e^{\frac{i}{\hbar} \int H_{aa}(\tau) d\tau} d_b(t)
\]

(6.40)

Similarly, the second equation becomes

\[
i \hbar \frac{d}{dt} d_b(t) = e^{i \omega t} H'_{ba}(t) e^{-\frac{i}{\hbar} \int H_{bb}(\tau) d\tau} d_a(t)
\]

(6.41)

For $d_a$ and $d_b$, the equations take the same form as Eq. (6.33) above. As a result, even if $H_{aa}$ and $H_{bb}$ are not zero, we will still have the same type of equations and we can solve them in the same way.

### 6.4. Step four: solve the equations using perturbation theory

\[
\begin{align*}
\frac{d}{dt} \left( \frac{\partial}{\partial \alpha} H(t) \right) & = e^{-i \omega t} \lambda H(t) c_0(t) \\
\frac{d}{dt} \left( \frac{\partial}{\partial \lambda} H(t) \right) & = e^{i \omega t} \lambda H(t) c_0(t)
\end{align*}
\]

(6.42)

Here, we can write $c_a$ and $c_b$ as power series of $\lambda$

\[
c_a(t) = c_a^{(0)}(t) + \lambda c_a^{(1)}(t) + \lambda^2 c_a^{(2)}(t) + \ldots
\]

(6.43)

\[
c_b(t) = c_b^{(0)}(t) + \lambda c_b^{(1)}(t) + \lambda^2 c_b^{(2)}(t) + \ldots
\]

(6.44)

So
\[
\begin{align*}
\{ & \quad \frac{\text{i}}{\hbar} \frac{d}{dt} \frac{\text{i}}{\hbar} \frac{d}{dt} \lambda + \frac{\text{i}}{\hbar} \frac{d}{dt} \lambda^2 + \ldots = 0 + e^{-i \omega t} H_{ab} c_b^{(0)}(t) \lambda + e^{-i \omega t} H_{ab} c_b^{(1)}(t) \lambda^2 + e^{-i \omega t} H_{ab} c_b^{(2)}(t) \lambda^3 + \ldots \\
& \quad \frac{\text{i}}{\hbar} \frac{d}{dt} \frac{\text{i}}{\hbar} \frac{d}{dt} \lambda + \frac{\text{i}}{\hbar} \frac{d}{dt} \lambda^2 + \ldots = 0 + e^{i \omega t} H_{ba} c_a^{(0)}(t) \lambda + e^{i \omega t} H_{ba} c_a^{(1)}(t) \lambda^2 + e^{i \omega t} H_{ba} c_a^{(2)}(t) \lambda^3 + \ldots \\
\end{align*}
\] (6.45)

For these two equations to hold, they must hold for every order of \(\lambda\).

**Zeroth order:**

At the zeroth order, we have

\[
\begin{align*}
\frac{\text{i}}{\hbar} \frac{d}{dt} c_b^{(0)} &= 0 \\
\frac{\text{i}}{\hbar} \frac{d}{dt} c_b^{(0)} &= 0
\end{align*}
\] (6.46)

In order words, we get

\[
\begin{align*}
c_b^{(0)} &= c_1 \\
c_b^{(0)} &= c_2
\end{align*}
\] (6.47)

where \(c_1\) and \(c_2\) are two constants, independent of time. To determine the value of these two constants, we use the initial condition. The initial condition tells us \(c_b(0) = 1\) and \(c_b(0) = 0\), so we know that \(c_1 = 1\) and \(c_2 = 0\)

\[
\begin{align*}
c_b^{(0)} &= 1 \\
c_b^{(0)} &= 0
\end{align*}
\] (6.48)

**First order:**

At the first order, we have

\[
\begin{align*}
\frac{\text{i}}{\hbar} \frac{d}{dt} c_b^{(0)} &= e^{-i \omega t} H_{ab} c_b^{(0)}(t) \\
\frac{\text{i}}{\hbar} \frac{d}{dt} c_b^{(0)} &= e^{i \omega t} H_{ba} c_a^{(0)}(t)
\end{align*}
\] (6.49)

For the r.h.s., we know that \(c_b^{(0)}(t) = 0\) and \(c_a^{(0)}(t) = 1\) (from the zeroth order perturbation theory). Therefore,

\[
\begin{align*}
\frac{\text{i}}{\hbar} \frac{d}{dt} c_b^{(0)} &= 0 \\
\frac{\text{i}}{\hbar} \frac{d}{dt} c_b^{(0)} &= \lambda e^{i \omega t} H_{ba}
\end{align*}
\] (6.50)

So we have

\[
\begin{align*}
c_b^{(1)}(t) &= c_1 \\
c_b^{(1)}(t) &= -\frac{\lambda}{\hbar} \int_0^t e^{i \omega t'} H_{ba} (t') dt' + c_2
\end{align*}
\] (6.51)

where \(c_1\) and \(c_2\) are two constants, independent of time. To determine the value of these two constants, we use the initial condition.

We know that

\[
\begin{align*}
c_b(t) &= c_b^{(0)} + \lambda c_a^{(1)} = 1 + c_1 \lambda \\
c_b(t) &= c_b^{(0)} + \lambda c_a^{(1)} = -\frac{\hbar}{\lambda} \int_0^t e^{i \omega t'} H_{ba} (t') dt' + c_2 \lambda
\end{align*}
\] (6.52)

At \(t = 0\), these solutions have

\[
\begin{align*}
c_b(0) &= 1 + c_1 \lambda \\
c_b(0) &= c_2 \lambda
\end{align*}
\] (6.53)

Because the initial condition says that \(c_b(0) = 1\) and \(c_b(0) = 0\), these two constants must be both zero,

Thus, we get

\[
\begin{align*}
c_b^{(1)}(t) &= 0 \\
c_b^{(1)}(t) &= -\frac{\lambda}{\hbar} \int_0^t e^{i \omega t'} H_{ba} (t') dt'
\end{align*}
\] (6.54)

**Second order:**

At the second order, we have
As shown in the previous section, in the zeroth order approximation, \( c_0(t) = 1 \) and \( c_0(t) = 0 \). Thus

\[
\psi(t) = c_0(t) e^{-iE_0 t/\hbar} \psi_0 + c_0(t) e^{-iE_0 t/\hbar} \psi_0 = e^{-iE_0 t/\hbar} \psi_0
\]

i.e. the system will always stay in the initial state and there is no transition.
The physics of transition arises only in the first order (or higher order) perturbations. In this section, we will consider only the leading order effect, i.e. the first order, and ignore all higher order effects. In the first order approximation,

\[
\begin{align*}
\langle \psi_a | H_{\text{int}}(t) | \psi_b \rangle &= c_a(t) \psi_a(t) + c_b(t) \psi_b(t) \\
\langle \psi_a | H_{\text{int}}(t) | \psi_b \rangle &= -\lambda \langle \psi_a | \psi_b \rangle \\
\langle \psi_a | H_{\text{int}}(t') | \psi_b \rangle &= \frac{-i}{\hbar} \int_0^{t'} \langle \psi_a | H_{\text{int}}(t'') | \psi_b \rangle \, dt''
\end{align*}
\]

(6.66)

Therefore, the probability for the system to turn into the state \( | \psi_b \rangle \) is

\[
P_{a \rightarrow b} = \sum_b \sum_b |c_b(t)|^2 = -\frac{\lambda^2}{\hbar^2} \int_0^\infty \langle \psi_a | H_{\text{int}}(t') | \psi_b \rangle \, dt' \int_0^\infty \langle \psi_a | H_{\text{int}}(t') | \psi_b \rangle \, dt' \]

(6.67)

Now, consider a sinusoidal perturbation,

\[
H'(\vec{r}, t) = V(\vec{r}) \cos(\omega t)
\]

(6.68)

Then,

\[
H'_{ba}(t) = \langle \psi_b | H'(t) | \psi_a \rangle = \langle \psi_b | V(\vec{r}) \cos(\omega t) | \psi_a \rangle = \langle \psi_b | V(\vec{r}) | \psi_a \rangle \cos(\omega t)
\]

(6.69)

Now, let’s define \( V_{ba} = \langle \psi_b | V(\vec{r}) | \psi_a \rangle \). Different from \( H'_{ba}(t) \), which is a function of \( t \), \( V_{ba} \) is just a complex number, which doesn’t depend on time.

\[
H'_{ba}(t) = V_{ba} \cos(\omega t)
\]

(6.70)

As a result,

\[
P_{a \rightarrow b} = \frac{\lambda^2}{\hbar^2} \int_0^\infty \langle \psi_a | e^{i\omega t} H'_{ba}(t') \rangle \, dt' = \frac{\lambda^2}{\hbar^2} \left( \int_0^\infty \langle \psi_a | e^{i\omega t} V_{ba} \rangle \, dt' \right)^2 = \frac{\lambda^2}{\hbar^2} \left( V_{ba} \right)^2 \int_0^\infty \langle \psi_a | e^{i\omega t} \rangle \, dt' \int_0^\infty \langle \psi_a | e^{i\omega t} \rangle \, dt'
\]

(6.71)

The integral here can be evaluated

\[
\int_0^\infty e^{i\omega t} \cos(\omega t) \, dt = \frac{\lambda^2}{\hbar^2} \left( \int_0^\infty e^{i\omega t} \cos(\omega t) \, dt \right)^2 = \frac{\lambda^2}{\hbar^2} \left( \int_0^\infty e^{i\omega t} \cos(\omega t) \, dt \right)^2
\]

(6.72)

For simplicity, we know limit us to \( \omega \) close to \( \omega_0 \), i.e. the frequency of the perturbation is close to \( \omega_0 \). In this case, \( \omega_0 + \omega >> \omega_0 - \omega \), and thus the first term is much smaller than the second term (because it has a larger denominator). As a result, we can drop the \( \frac{\lambda^2}{\hbar^2} \) term.

\[
\int_0^\infty e^{i\omega t} \cos(\omega t) \, dt = \frac{-i}{\omega_0 + \omega} e^{i(\omega_0 + \omega)t} \left( \frac{1}{1-e^{-2\omega}} - \frac{1}{1-e^{-2\omega}} - \frac{1}{e^{-2\omega} - 1} \right)
\]

(6.73)

Therefore,

\[
P_{a \rightarrow b} = \frac{\lambda^2}{\hbar^2} \int_0^\infty \langle \psi_a | e^{i\omega t} \rangle \, dt' \int_0^\infty \langle \psi_a | e^{i\omega t} \rangle \, dt' = \lambda^2 \left( \int_0^\infty \langle \psi_a | e^{i\omega t} \rangle \, dt' \right)^2
\]

(6.74)

In our textbook (and many other literature), the coefficient \( \lambda \) is absorbed into \( H' \). Instead of \( H = H_0 + \lambda H'(t) \), one uses \( H = H_0 + H'(t) \). As a result, the final result will not contain \( \lambda \).

\[
P_{a \rightarrow b} = \frac{\lambda^2}{\hbar^2} \left( \int_0^\infty \langle \psi_a | e^{i\omega t} \rangle \, dt' \right)^2
\]

(6.75)

because \( \lambda \) is now observed into \( V_{ba} \).
Comment #1: the probability oscillates as a periodic function of time $t$. The frequency of the oscillation is NOT the driven frequency, but $|\omega - \omega_0|$. This is easy to notice if we rewrite $P$ as a cos function of $t$.

$$P_{a \rightarrow b} = \frac{|V_{ba}|^2}{\hbar^2} \frac{\sin^2 \left( \frac{(\omega_0 - \omega)t}{2} \right)}{(\omega_0 - \omega)^2} = \frac{|V_{ba}|^2}{2 \hbar^2} \frac{1 - \cos[(\omega_0 - \omega)t]}{(\omega_0 - \omega)^2}$$  

(6.76)

Comment #2: The average probability (which is half of the height of the peaks shown in the figure above) is proportional to $\frac{1}{(\omega_0 - \omega)^2}$. If the driven frequency is closer to $|E_b - E_a|/\hbar$, we have larger probability to transit from a to b.

Comment #3: The probability needs to be smaller than 1. In the first order perturbation theory, the result is not obviously smaller than 1. Instead, if $\omega_0 - \omega$ is really close to zero, the high of the peaks diverge. This is because we are using the perturbation theory, and the perturbation theory assume that all corrections a much smaller than the zeroth order result, i.e., we assumed that $V_{ba}$ is very small, such that $\frac{|V_{ba}|^2}{\hbar^2 (\omega_0 - \omega)^2} < 1$. When $\omega_0$ and $\omega$ are too close to each other, this approximation fail.

Comment #4: If we have a system with only two energy levels $\psi_a$ and $\psi_b$, the problem can be solved within the rotating frame approximation without using the perturbation theory. There, the solution is

$$P_{a \rightarrow b} = \frac{|V_{ab}|^2}{\hbar^2} \frac{\sin^2 \left( \frac{(\omega_0 + \omega)t}{2} \right)}{\omega_r^2}$$

Because $\omega_r = \sqrt{\omega_0^2 + |V_{ab}|^2 / \hbar^2}$. One can easily verify that result solution is indeed smaller than 1

$$P_{a \rightarrow b} = \frac{|V_{ab}|^2}{\hbar^2} \frac{\sin^2 \left( \frac{(\omega_0 + \omega)t}{2} \right)}{\omega_r^2} \leq \frac{|V_{ab}|^2}{\hbar^2} \frac{1}{\omega_0^2} \leq 1$$  

(6.78)

### 6.6. Transition rate

The transition rate is defined as:

$$R_{a \rightarrow b}(t) = \frac{dP_{a \rightarrow b}}{dt} = \frac{|V_{ab}|^2}{2 \hbar^2} \frac{d}{dt} \frac{1 - \cos[(\omega_0 - \omega)t]}{(\omega_0 - \omega)^2} = \frac{|V_{ab}|^2}{2 \hbar^2} \frac{\sin[(\omega_0 - \omega)t]}{(\omega_0 - \omega)}$$  

(6.79)

In the long time limit, $t \rightarrow \infty$, this is a delta function because

$$\lim_{t \rightarrow \infty} \frac{\sin[(\omega_0 - \omega)t]}{(\omega_0 - \omega)} = \pi \delta(\omega - \omega_0)$$  

(6.80)
Thus,

\[ R_{a \rightarrow b}(t) = \frac{\pi}{2 \hbar^2} \left| V_{ab} \right|^2 \delta(\omega - \omega_0) \]  

(6.81)

This result is a little bit different from what we have learned in the previous chapter, where we found that

\[ R_{i \rightarrow f} = \frac{2 \pi}{\hbar^2} \left| \langle f | H' | i \rangle \right|^2 \delta(\omega - \omega_0) \left( E_f - E_i \right) \]  

(6.82)

Notice that there is a factor 4 difference. This factor 4 is because in the previous chapter, the time-dependent perturbation is an exponential function \( \propto e^{i \omega t} \). In this chapter, it is a cosine function \( \propto \cos \omega t \). When we compare cos and exp functions, there is a factor 1/2 difference

\[ \cos \omega t = \frac{e^{i \omega t} + e^{-i \omega t}}{2} \]  

(6.83)

This 1/2 factor, because there is a square in the functional form of \( R \), become 1/4, which is why we have this difference.

In different literature, the formula may differ by a factor of 4 for reasons mentioned above.

### 6.7. Emission and Absorption

Consider an atom (e.g. hydrogen or other elements), and we apply an E&M wave (i.e. shine light) on it.

\[ H = H_0 + H'(t) \]  

(6.84)

where \( H_0 \) is the Hamiltonian of the atom, and \( H'(t) \) is the coupling to the E&M field. The coupling to E&M fields has two parts:

\[ H = \frac{(-i \hbar \nabla - q A)^2}{2m} + V(r) + q \varphi \]  

(6.85)

the magnetic field contribution \( A \) and electric field contribution \( \varphi \). It turns out that \( \varphi \) usually has a much strong effect, thus we can ignore the \( A \) part here and only consider charge \( q \) times the electric potential \( \varphi \). Thus

\[ H' = q \varphi(t) \]  

(6.86)

For an E&M wave with electric field polarized along the \( z \) direction, the electric field is along \( z \) and we assume that the electric field is uniform (same value everywhere)

\[ \vec{E} = E_0 \cos(\omega t) \hat{z} \]  

(6.87)

For such a field, the electric potential is

\[ \varphi(x, y, z, t) = -\vec{E} \cdot \vec{r} = -E_0 z \cos(\omega t) \]  

(6.88)

and thus

\[ H' = q \varphi(t) = -q E_0 z \cos(\omega t) \]  

(6.89)

#### 6.7.1. \( H_0 \)

For the unperturbed Hamiltonian, an atom has many (infinite) energy levels.

\[ H_0 \psi_n = E_n \psi_n \]  

(6.90)

Using the language of wavefunctions, the eigenequation reads

\[ H_0 \psi_n(\vec{r}) = E_n \psi_n(\vec{r}) \]  

(6.91)

Here, again, we consider two of the energy levels,

\[ H_0 \psi_a(\vec{r}) = E_a \psi_a(\vec{r}) \]  

and
\[ H_0 \psi_1(\vec{r}) = E_0 \psi_1(\vec{r}) \] (6.93)

### 6.7.2. First order perturbation for a monochromatic light source

In the first order perturbation theory, we have

\[ P_{\alpha \rightarrow \beta} = \frac{1}{\hbar^2} \frac{V_{\alpha \beta}}{(\omega_0 - \omega)^2} \left( \frac{1 - \cos[(\omega_0 - \omega) t]}{2 (\omega_0 - \omega)^2} \right) \] (6.94)

Here,

\[ V_{\alpha \beta} = \langle \psi_\beta | -q A_0 | \psi_\alpha \rangle = -q E_0 \langle \psi_\beta | z | \psi_\alpha \rangle = -q E_0 \int d\vec{r} \psi_\beta^* (\vec{r}) z \psi_\alpha (\vec{r}) = -E_0 \mathcal{P} \] (6.95)

Here, we can define

\[ \mathcal{P} = q \langle \psi_\beta | z | \psi_\alpha \rangle = q \int d\vec{r} \psi_\beta^* (\vec{r}) z \psi_\alpha (\vec{r}) \] (6.96)

This is a number that one can easily compute. For example, for a hydrogen atom, we know all the eigenwavefunctions, including \( \psi_\alpha \) and \( \psi_\beta \), and thus one can evaluate the integral and that will produce a number.

After figuring out the value of \( \mathcal{P} \), we know that

\[ P_{\alpha \rightarrow \beta} = \frac{1}{\hbar^2} \frac{V_{\alpha \beta}}{(\omega_0 - \omega)^2} \left( \frac{1 - \cos[(\omega_0 - \omega) t]}{2 (\omega_0 - \omega)^2} \right) \] (6.97)

In E&M, we learned that the energy density of an E&M is

\[ u = \frac{\epsilon_0}{2} E_0^2 \] (6.98)

Thus

\[ P_{\alpha \rightarrow \beta} = \frac{2 u}{\epsilon_0 \hbar^2} |\mathcal{P}|^2 \frac{\sin^2 \left[ \frac{(\omega_0 - \omega) t}{2} \right]}{(\omega_0 - \omega)^2} \] (6.99)

### 6.7.3. Non-monochromatic light

In reality, the light source is never monochromatic. The energy is always distributed in some range of frequencies. There, the total energy density (from all frequencies) are

\[ u = \int \rho(\omega) d\omega \] (6.100)

For each frequency, we can use the same frequency to obtain the probability of transition (induced by light at frequency \( \omega \))

\[ \frac{2 \rho(\omega)}{\epsilon_0 \hbar^2} |\mathcal{P}|^2 \frac{\sin^2 \left[ \frac{(\omega_0 - \omega) t}{2} \right]}{(\omega_0 - \omega)^2} \] (6.101)

And we can integrate all frequency to get the total probability

\[ P_{\alpha \rightarrow \beta}(t) = \int d\omega \frac{2 \rho(\omega)}{\epsilon_0 \hbar^2} |\mathcal{P}|^2 \frac{\sin^2 \left[ \frac{(\omega_0 - \omega) t}{2} \right]}{(\omega_0 - \omega)^2} \] (6.102)

We can also compute the transition rate:

\[ R_{\alpha \rightarrow \beta}(t) = \frac{dP_{\alpha \rightarrow \beta}}{dt} = \frac{1}{\hbar^2} \frac{d}{dt} \int d\omega \rho(\omega) \frac{1 - \cos[(\omega_0 - \omega) t]}{(\omega_0 - \omega)^2} = \frac{1}{\hbar^2} \int d\omega \rho(\omega) \left( \frac{1 - \cos[(\omega_0 - \omega) t]}{(\omega_0 - \omega)^2} \right) \] (6.103)

In the long time limit, \( t \rightarrow \infty \), \( \sin[(\omega_0 - \omega) t]/(\omega_0 - \omega) \) becomes \( \pi \delta(\omega - \omega_0) \), so
\[
R_{a \rightarrow b} = \frac{P^2}{\epsilon_0 h^2} \int d\omega \rho(\omega) \frac{\pi}{\epsilon_0 h^2} \int d\omega \rho(\omega) \delta(\omega - \omega_0) = \frac{\pi}{\epsilon_0 h^2} \rho(\omega_0)
\] (6.104)

Note: if it is a monochromatic light source, we will have a delta function in the transition rate. However, in reality, there is no monochromatic light source, and thus at the end of the day, there is no delta function in \(R\), i.e. no infinity.

### 6.7.4. Summary

If we shine light on atoms (assuming that they are all the same type of atoms and they don’t interact with one another, but these atoms can be on different energy levels). For an atom, which was initial at state \(\psi_a\), it can jump to a different state \(\psi_b\) if the light contains frequency \(\omega = |E_a - E_b| / h\)

**Absorption:** If the energy of the initial state is lower than the final state, due to energy conservation, the atom must take some energy from the light, which is known as “absorption”. And we say that the atom absorbs a photon at frequency \(\omega\). The energy that was absorbed in this procedure is \(|E_a - E_b|\), which matches \(\omega h\). For lights, the experimental phenomenon is that the intensity for light at this frequency becomes smaller.

**Emission:** The reverse procedure implies light emission. If the energy of the initial state is higher than the final state, due to energy conservation, the atom will give some energy from the light, which is known as “emission”. And we say that the atom emits out a photon at frequency \(\omega\). The energy that was sent out in this procedure is \(|E_a - E_b|\), which matches \(\omega h\), i.e. the energy of a photon is \(h \omega\). For lights, the experimental phenomenon is that the intensity for light at this frequency becomes larger.

The time-perturbation theory tells us that the absorption rate and the emission rate are identical, because \(|\langle \psi_a | H' | \psi_b \rangle| = |\langle \psi_b | H' | \psi_a \rangle|\).

**Spontaneous Emission:** If an atom was put on an excited state \(E_a\), and we do NOT have any E&M waves in our system, is that possible to make the atom go to a lower energy state (say \(E_b < E_a\)) by emit an photon?

Based on the perturbation theory and quantum mechanics, the answer is negative. However, the real answer is positive. And the correct answer is beyond quantum mechanics.

### 6.8. Spontaneous Emission

**Q:** Why is there spontaneous emission, when quantum mechanics tells us that \(P_{a \rightarrow b}\) is zero in the absence of perturbation \(H^*\)?

**A:** A (not very accurate) answer based on the fact that “absolutely no E&M wave in a system” is a fake concept. In quantum mechanics, there is no motionless particles, because of the uncertainty principle. This is known as the zero point energy (e.g. in a harmonic oscillator, classical mechanics says that the lowest possible energy of the system is zero, the spring staying at its rest length and the mass is not moving. However, this is not the case. Instead, in QM, the lowest possible energy of the system is \(\hbar \omega / 2\), known as the zero point energy). Similarly, it is impossible to have no E&M waves. As a result, it is always possible for an atom to transit from a high energy state to a low energy state.

### 6.8.1. Einstein’s A and B coefficients

Consider a system with atoms and lights, and everything is at thermal equilibrium.

Consider two quantum stats in an atom here, \(\psi_a\) and \(\psi_b\). We assume that there are \(N_a\) atoms in state \(\psi_a\) and \(N_b\) atoms in \(\psi_b\). And here, we assume that the energy of the b state is higher than a. We know that

\[
\frac{dN_b}{dt} = -N_b R_{b \rightarrow a} + N_a R_{a \rightarrow b} - N_b A
\] (6.105)

Here, on the r.h.s., the first term describes the transition from b to a. The transition rate for one atom is \(R_{b \rightarrow a}\). Since we have \(N_b\) atom, it will leads to a total transition rate \(-N_b R_{b \rightarrow a}\). The negative sign here means that this transition reduces the number of atoms on state b (because they turns into state a). The second term is the inverse procedure, from a to b. This one increases number of atoms on state b, and thus has a positive sign. The last term is spontaneous emission, which reduces the number of atoms on state b (spontaneous emission makes the atom goes from an high energy state, \(\psi_b\) to a lower energy state \(\psi_a\), and thus it reduces \(N_b\)). We assume the rate is \(A\).

For \(R_{b \rightarrow a}\) and \(R_{a \rightarrow b}\), our perturbation theory tells us that they are identical and both of them are proportional to the intensity of light at frequency
\[ \omega_0 = |E_a - E_b|/\hbar, \]

\[ R_{a \rightarrow b} = R_{b \rightarrow a} = \rho(\omega_0) B \]

where \( B \) is some number. Thus

\[ \frac{d N_b}{d t} = -N_b B \rho(\omega_0) + N_a B \rho(\omega_0) - N_b A \]

(6.107)

The coefficients \( A \) and \( B \) here are known as the Einstein's A and B coefficients. According to the perturbation theory,

\[ B = \frac{1}{3} \frac{\pi}{\epsilon_0 h^2} \]

(6.108)

and \( A = 0 \). However, as will be shown below, \( A \) is NOT zero.

**NOTE:** in our calculation above, we assume the E&M have a specific polarization and a unique wavevector. In reality, light (at a fixed frequency) can have wavevectors oriented in different directions and the polarization can be in different directions too. After averaging out these different directions, a factor 1/3 is obtained. This is where the extra 1/3 factor comes from.

In thermal equilibrium, \( N_b \) should not change with time and thus \( N_b = 0 \)

\[ \frac{d N_b}{d t} = -N_b B \rho(\omega_0) + N_a B \rho(\omega_0) - N_b A = 0 \]

(6.109)

Thus

\[ \rho(\omega_0) = \frac{N_a}{N_a B - N_b B} = \frac{A}{B} \]

\[ \frac{N_a}{N_b} = \exp \left( \frac{E_a}{k_B T} \right) = \exp \left( \frac{E_b - E_a}{k_B T} \right) = \exp \left( \frac{\hbar \omega_0}{k_B T} \right) \]

(6.111)

So

\[ \rho(\omega_0) = \frac{A}{B} \exp \left( \frac{\hbar \omega_0}{k_B T} \right) - 1 \]

(6.112)

In the study about black body radiation, we learned that the energy density of E&M waves satisfies the Planck distribution

\[ \rho(\omega_0) = \frac{1}{\pi^2 c^3} \frac{\hbar \omega_0^3}{\exp \left( \frac{\hbar \omega_0}{k_B T} \right) - 1} \]

(6.113)

Thus, we find that

\[ A/B = \frac{\hbar \omega_0^3}{\pi^2 c^3} \]

(6.114)

So

\[ A = \frac{\hbar \omega_0^3}{\pi^2 c^3} B = \frac{\hbar \omega_0^3}{\pi^2 c^3} \frac{\pi |P|^2}{3 \epsilon_0 \hbar^2} = \frac{|P|^2 \omega_0^2}{3 \pi \epsilon_0 \hbar c^5} \]

(6.115)

This is the rate of spontaneous emission (from \( b \) to \( a \)). This calculation tells us that in reality, even if we don't perturb the system, a high energy state will still have finite probability to go to a lower energy state, by emitting a photon. In other words, if we prepare an atom in a high energy eigen-state, even if we don’t do anything, it will emit some light and go down to a lower energy state.

**This means that any excited state have a finite life time!**

In the absence of external perturbations, \( \rho(\omega) = 0 \)
\[
\frac{dN_b}{dt} = -N_b R_{b\rightarrow a} + N_a R_{a\rightarrow b} - N_b A = -N_b A
\]  

(6.116)

Thus,
\[
\frac{dN_b}{dt} = -N_b A
\]  

(6.117)

The solution of this differential equation is
\[
N_b(t) = N_b(0) e^{-A t}
\]  

(6.118)

In other words, the number of particle on this high energy state is reducing exponentially as we increase \( t \). We can define a time scale \( \tau = 1/A \)
\[
N_b(t) = N_b(0) e^{-t/\tau}
\]  

(6.119)

This \( \tau \) is known as the lifetime of this state.

Of course, in reality, a high energy state \( \psi_b \) can go to multiple lower energy states, \( \psi_{a1}, \psi_{a2}, \ldots \), if \( \mathcal{H}_0 \) multiple eigenstates whose energies are lower than \( E_b \). Each of this procedure will have a spontaneous emission rate \( A_i, i = 1, 2, 3 \ldots \)
\[
\frac{dN_b}{dt} = -N_b A_1 - N_b A_2 - N_b A_3 - \ldots
\]  

(6.120)

Therefore,
\[
\frac{dN_b}{dt} = -(A_1 + A_2 + A_3 + \ldots) N_b
\]  

(6.121)

We can define total decay rate \( A \) as
\[
A = A_1 + A_2 + A_3 + \ldots
\]  

(6.122)

Thus the lifetime is
\[
\tau = 1/A = \frac{1}{A_1 + A_2 + A_3 + \ldots}
\]  

(6.123)

### 6.8.2. The origin of spontaneous emission

Q1: Where does this spontaneous emission come from? (Why \( A \) is NOT zero as the perturbation theory tells us?)

Q2: Why are A and B coefficients directly related?

A: This is because one assumption that we used above is wrong: \( R_{b\rightarrow a} = R_{a\rightarrow b} \)

Due to the time-reversal symmetry (physics laws remains the same as we reverse the arrow of time), the transition rate for a procedure must be identical to the inverse procedure. This is why we concluded that
\[
R_{b\rightarrow a} = R_{a\rightarrow b}
\]  

(6.124)

Q: How can this be wrong?

A: \( b \rightarrow a \) and \( a \rightarrow b \) are NOT inverse procedure of each other

If we only look at the atom, and forget about the E&M waves, \( b \rightarrow a \) and \( a \rightarrow b \) are indeed inverse procedure of each other. However, if we taken into account the E&M waves, they are NOT inverse of each other.

The origin of this error comes from the fact that in quantum mechanics, we treat particles and waves differently (they should be treated the same way due to the particle-wave duality). We treated particles as quantum particles (i.e. we use quantum theory treat particles, and particles follow totally different equations in QM, in comparison with classical physics). However, for E&M waves, we treat them same as classical waves (they follow the same Maxwell equations. There is nothing new beyond classical E&M).

Key: in QM, we couple classical waves with quantum particles. This is NOT a complete quantum theory, in which everything needs to be treated as a quantum object (particles as well as waves).

The correct way to quantize E&M waves will be discussed in quantum field theory, although other courses, like advanced mechanics, may
touch some aspect of the problem. After we quantize the E&M waves, we will find that \( R_{b \rightarrow a} \) and \( R_{a \rightarrow b} \) are slightly different

\[
R_{b \rightarrow a} \neq R_{a \rightarrow b}
\]

**Q: Why are \( R_{b \rightarrow a} \) and \( R_{a \rightarrow b} \) different?**

For absorption, an atom takes a photon and goes from a low energy state \((a)\) to a high energy state \((b)\). In addition to the transition \( a \rightarrow b \) for the atom, the number of photon is reduced by 1. As a result, if we start from a system with \( N \) photons, the final state has \( N - 1 \) photons. We can write down the initial and final quantum states (including both the atom and photons) as:

\[
|\psi_a; N\rangle \rightarrow |\psi_b, N - 1\rangle
\]

Here, to emphasize the photons, we write the absorption rate as:

\[
R_{a \rightarrow b, N}
\]

The inverse of this absorption is

\[
|\psi_b; N - 1\rangle \rightarrow |\psi_a, N\rangle
\]

For emission, the quantum state of the atom changes from \( \psi_b \) to \( \psi_a \), and in the same time, the atom emits a photon and thus the photon number increases by 1. If we start with the same number of photons, which is \( N \), the final state will have \( N + 1 \) photons

\[
|\psi_b; N\rangle \rightarrow |\psi_a, N + 1\rangle
\]

Here, to emphasize the photons, we write the emission rate as:

\[
R_{b \rightarrow a, N}
\]

If we compare the absorption and emission procedures, they are obviously NOT inverse to each other. Their photon numbers differ by 1. As a result, the transition rate also differ a little bit, and this is the origin of spontaneous emission. In fact, the inverse of emission (with \( N \) photons in the initial state) is absorption with \( N + 1 \) photons in the initial state. So we should have

\[
R_{b \rightarrow a, N} = R_{a \rightarrow b, N+1} \neq R_{a \rightarrow b, N}
\]

For absorption, we can use the same results that we obtained above:

\[
R_{a \rightarrow b} = \frac{\pi}{3} \frac{|P|^2}{\epsilon_0 c^2} \rho(\omega_0)
\]

Because E&M waves are photons, the total energy carried by E&M waves are

\[
u = 2 \int \frac{d\vec{k}}{(2\pi)^3} \frac{\hbar c}{h} n(\vec{k})
\]

Here, \( n(\vec{k}) \) are number (density) of photons with wavevector \( \vec{k} \), \( c = \hbar \omega \) is the energy of a photon. Here we sum over all wavevectors to compute the total energy. The over all factor “2” is because there are two polarizations for light.

\[
u = 2 \int \frac{4\pi k^2 dk}{(2\pi)^3} \frac{\hbar c}{h} n(k) = \int \frac{\hbar c}{\pi^2} k^3 n(k)
\]

Notice that \( c = \omega \), so we can change our integral parameter from \( k \) to \( \omega \).

\[
u = \int \frac{\hbar c}{\pi^2} k^3 n(k) = \int d\omega \frac{\hbar \omega^3}{\pi^2 c^3} n(\omega)
\]

Compare with the definition for \( \rho(\omega) \)

\[
u = \int d\omega \rho(\omega)
\]

we find that

\[
\rho(\omega) = \frac{\hbar}{\pi^2 c^3} \omega^3 n(\omega)
\]

And thus
\[ R_{a\rightarrow b,N} = \frac{\pi |P|^2}{3} \frac{\hbar}{\varepsilon_0} \rho(\omega_0) = \frac{\pi |P|^2}{3} \frac{\hbar}{\varepsilon_0} \rho(\omega_0) \frac{\pi}{\varepsilon_0} = \frac{\pi}{\varepsilon_0} \frac{\hbar}{\varepsilon_0} \rho(\omega_0) \] (6.136)

However, for emission, (an atom gives out a photon and goes from a high energy state to a low energy state), as shown above, it should be the same as absorption with photon number increased by 1, so it is

\[ R_{b\rightarrow a,N+1} = \frac{\pi |P|^2}{3} \frac{\hbar}{\varepsilon_0} \rho(\omega_0) \frac{\pi}{\varepsilon_0} = \frac{\pi}{\varepsilon_0} \frac{\hbar}{\varepsilon_0} \rho(\omega_0) \frac{\pi}{\varepsilon_0} = \frac{\pi}{\varepsilon_0} \frac{\hbar}{\varepsilon_0} \rho(\omega_0) \] (6.137)

Thus we find

\[ R_{b\rightarrow a,N} = \frac{\pi |P|^2}{3} \frac{\hbar}{\varepsilon_0} \rho(\omega_0) \frac{\pi}{\varepsilon_0} = \frac{\pi}{\varepsilon_0} \frac{\hbar}{\varepsilon_0} \rho(\omega_0) \frac{\pi}{\varepsilon_0} = \frac{\pi}{\varepsilon_0} \frac{\hbar}{\varepsilon_0} \rho(\omega_0) \] (6.138)

The difference between \( R_{b\rightarrow a} \) and \( R_{b\rightarrow a} \) is precisely the \( A \) coefficient assumed by Einstein. In this fully quantum theory, we don't need to assume an \( A \) coefficient,

\[ \frac{dN_b}{dt} = -N_b R_{b\rightarrow a,N} + N_a R_{a\rightarrow b,N} \] (6.139)

Because the quantization of E&M waves tells us that \( R_{a\rightarrow b} = R_{a\rightarrow b} + A \), we get

\[ \frac{dN_b}{dt} = -N_b (R_{b\rightarrow a,N} + A) + N_a R_{a\rightarrow b,N} = -N_b R_{a\rightarrow b} + N_a R_{a\rightarrow b} - N_b A = -N_b B \rho + N_a B \rho - N_b A \] (6.140)

The spontaneous emission term, assumed by Einstein naturally arises

### 6.8.3. Selection rule

The transition rate is proportional to \(|P|^2\)

\[ R_{a\rightarrow b} = \frac{\pi}{\varepsilon_0} \frac{\hbar}{\varepsilon_0} \rho(\omega_0) \] (6.141)

If \(|P|^2\) is zero, then this transition rate becomes zero. For E&M waves with polarization along \( z \), we have shown that

\[ P = q \langle \psi_b | z | \psi_a \rangle = q \int d\vec{r} \psi_b^*(\vec{r}) z \psi_a(\vec{r}) \] (6.142)

If we are considering eigenwavefunctions of a hydrogen atom, (assuming the \( a \) state has quantum numbers \( n, l, m \) and the \( b \) state has quantum numbers \( n', l' \) and \( m' \))

\[ P = q \int d\vec{r} \psi_{n'l'm'}^*(\vec{r}) z \psi_{n'l'm}(\vec{r}) \] (6.143)

It turns out that one can prove that most of the time \( P = 0 \), i.e. the transition rate is zero. To get a non-zero \( P \), the following two condition must be satisfied

\[ l' = l \pm 1 \] (6.144)

and

\[ m' = m \pm 1 \] (6.145)

These relations are called selection rules. They tells us that which transition is possible (and which is not). If we satisfy the selection rule, a transition rate is finite. Otherwise, the transition rate is zero even if we make the frequency matches \( \frac{E_a - E_b}{\hbar} \).

**Comment:** These selection rules come from angular momentum conservation. We can prove them purely based on angular momentum and their commutators (see our textbook for the proof).

**Question:** Is this transition rate absolutely zero when the selection rules are not satisfied?

**Answer:** Not really. The rate is usually much smaller, but will not be exactly zero for two reasons

1. The conclusion above is based on the first order perturbation theory. If one consider higher order perturbations, even if \( P \), there is still some...
transition rate (from higher order terms). Because they are from higher order terms, the rate is small. But it is not zero.

(2) The conclusions above only considers E&M waves. There are other channels to make the transition happen. For example, the collision between two atoms. In a gas of atoms, occasionally, two atoms collide with each other. Here, the two atom could exchange some energy, as a result, one of the atom may turn from state $a$ to $b$. Here, because it is NOT driving by E&M waves, the above calculation doesn’t applied and thus the selection rule become irrelevant.
Second quantization: where quantization and particles come from?

7.1. Lagrangian mechanics and canonical quantization

Q: How do we quantize a general system?

7.1.1. Lagrangian

Lagrangian mechanics is a reformulation of classical mechanics.

\[ L = T - V \]  \hspace{1cm} (7.1)

Here, the quantity \( L \) is called the Lagrangian, where \( T \) is the kinetic energy and \( V \) is the potential energy.

For example, for a pendulum as shown in the figure, the kinetic energy is

\[ T = \frac{1}{2} m r^2 \left( \frac{d\theta}{dt} \right)^2 = \frac{1}{2} m r \dot{\theta}^2 \]  \hspace{1cm} (7.2)

where \( \dot{\theta} \) represents time derivative of \( \theta \)

\[ V = -m g r \cos \theta \]  \hspace{1cm} (7.3)

And thus

\[ L(\dot{\theta}, \theta) = \frac{1}{2} m r^2 \dot{\theta}^2 + m g r \cos \theta \]  \hspace{1cm} (7.4)

Here, the Lagrangian depends on \( \theta \) and the time derivative of \( \theta \). In Lagrangian mechanics, \( \theta \) is called a coordinate. In this example, we have only one single coordinate, but in general, a more complicated system can have multiple coordinates. And in general, \( L \) depends on the coordinates and the time derivatives of these coordinates, e.g.

\[ L(q_1, \dot{q}_1, q_2, \dot{q}_2; \ldots; q_n, \dot{q}_n) \]  \hspace{1cm} (7.5)

where \( q_i \) are coordinates.

7.1.2. Equation(s) of motion

The equation of motion in Lagrangian mechanics:

\[ \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} \]  \hspace{1cm} (7.6)

For the pendulum example discussed above,
\[
\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} = \frac{\partial L}{\partial \theta}
\]  
(7.7)

So
\[
\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} = \frac{d}{dt} \frac{\partial T}{\partial \dot{\theta}} = \frac{d}{dt} \left( \frac{1}{2} m r^2 \dot{\theta}^2 \right) = \frac{d}{dt} \left( m r^2 \dot{\theta} \right) = m r^2 \dot{\theta}
\]  
(7.8)

and
\[
\frac{\partial L}{\partial \theta} = - \frac{\partial V}{\partial \theta} = \frac{\partial m g r \cos \theta}{\partial \theta} = -m g r \sin \theta
\]  
(7.9)

As a result, the EOM is
\[
m r^2 \ddot{\theta} = -m g r \sin \theta
\]  
(7.10)

\[
m r \dot{\theta} = -m g \sin \theta
\]  
(7.11)

which is the same as the second law of Newton:
\[
a = \frac{F}{m}
\]  
(7.12)

### 7.1.3. Hamiltonian, canonical coordinates and canonical momenta

From the Lagrangian, we can define a canonical momentum for every coordinate:
\[
p_i = \frac{\partial L}{\partial \dot{q}_i}
\]  
(7.13)

If the system has \(n\) coordinates, we will have \(n\) canonical momenta. If the coordinates we used are real space coordinates, these momenta coincide with the “momentum” we usually used in Newton’s theory. If the coordinates that we used here are angles, then the momenta will be angular momenta \(L\).

Hamiltonian:
\[
H = \sum p_i q_i - L = T + V
\]  
(7.14)

\(H\) is a function of \(p_i\) and \(q_i\)

For the pendulum example,
\[
p = \frac{\partial L}{\partial \dot{\theta}} = \frac{\partial L}{\partial \dot{\theta}} = \frac{\partial T}{\partial \dot{\theta}} = \frac{\partial \left( \frac{1}{2} m r^2 \dot{\theta}^2 \right)}{\partial \dot{\theta}} = m r^2 \dot{\theta}
\]  
(7.15)

Notice that \(\dot{\theta}\) is just the angular velocity \(\omega\). Here, this \(p\) is in fact \(m r \omega\), i.e. the angular momentum.
\[
H = p \dot{\theta} - L = p \dot{\theta} - \frac{1}{2} m r^2 \dot{\theta}^2 - m g r \cos \theta
\]  
(7.16)

Because \(p = m r^2 \dot{\theta}\), we can replace all \(\dot{\theta}\) by \(p/m r^2\)
\[
H = \frac{p}{m r^2} - \frac{1}{2} m r^2 \left( \frac{p}{m r^2} \right)^2 - m g r \cos \theta = \frac{p^2}{2 m r^2} - \frac{p^2}{2 m r^2} - m g r \cos \theta = \frac{p^2}{2 m r^2} - m g r \cos \theta
\]  
(7.17)

This Hamiltonian is indeed a function of \(p\) and \(\theta\)

### 7.1.4. Canonical quantization

In a quantum system, all physical observable becomes quantum operators, including \(H\), \(q_i\) and \(p_i\). The relation between them remains the same as in classical mechanics. The only ingredient that we need to add here is the “uncertainty relation”:
\[
[q_i, p_j] = i \hbar \delta_{i,j}
\]  
(7.18)
\[ [q_i, q_j] = [p_i, p_j] = 0 \]  
(7.19)

i.e., we request a coordinate NOT to commute with its canonical momentum. The commutator is assumed to be a constant, $i \hbar$ for any canonical pair $p_i$ and $q_j$.

### 7.1.5. Summary

Step 1: starting from classical mechanics

Step 2: define coordinates

Step 3: find the canonical momenta for each coordinate

Step 4: write down the Hamiltonian

Step 5: require canonical commutation relation

**Example:** for the example we considered above, we have only one coordinate and one momentum, and the Hamiltonian looks like

\[ H = \frac{p^2}{2 m r^2} - m g r \cos q \]  
(7.20)

When we quantize this system, we turn everything into quantum operators

\[ \hat{H} = \frac{\hat{p}^2}{2 m r^2} - m g r \cos \hat{q} \]  
(7.21)

And we request $[\hat{q}, \hat{p}] = i \hbar$.

Q: What is a cos function of a quantum operator?

A: We use Taylor expansions

\[ \cos x = 1 - \frac{x^2}{2} + \frac{x^4}{24} + \ldots \]  
(7.22)

So

\[ \cos \hat{q} = 1 - \frac{\hat{q}^2}{2} + \frac{\hat{q}^4}{24} + \ldots \]  
(7.23)

And thus

\[ \hat{H} = \frac{\hat{p}^2}{2 m r^2} - m g r \cos \hat{q} = \frac{\hat{p}^2}{2 m r^2} - m g r + \frac{m g r}{2} \hat{q}^2 - \frac{m g r}{24} \hat{q}^4 + \ldots \]  
(7.24)

We know that a constant in energy is not important, so we can drop the constant part $-m g r$

\[ \hat{H} = \frac{\hat{p}^2}{2 m r^2} - m g r \cos \hat{q} = \frac{\hat{p}^2}{2 m r^2} + \frac{m g r}{2} \hat{q}^2 - \frac{m g r}{24} \hat{q}^4 + \ldots \]  
(7.25)

The first two terms here are the Hamiltonian of a harmonic oscillator. And the rest part of the Hamiltonian is known as the non-harmonic part of the potential.

### 7.2. Creation and annihilation operators

Q: How do we handle such a quantum system?

A: By defining creation and annihilation operators.

#### 7.2.1. Creation and annihilation operators.

Define:
\[ a_i^+ = \frac{\sqrt{x} \ q_i - \frac{i}{\sqrt{x}} \ p_i}{\sqrt{2 \ h}} \]  
(7.26)

\[ a_i = \frac{\sqrt{x} \ q_i + \frac{i}{\sqrt{x}} \ p_i}{\sqrt{2 \ h}} \]  
(7.27)

where \( x \) is some arbitrary (positive) real number. It is easy to prove that \( a_i \) and \( a_i^\dagger \) are conjugate of each other,

\[ (a_i^\dagger)^\dagger = a_i \]  
(7.28)

and it is easy to verify that \([a_i, a_i^\dagger] = 1\)

\[ a_i a_i^\dagger - a_i^\dagger a_i = [a_i, a_i^\dagger] = \left[ \frac{\sqrt{x} \ q_i - \frac{i}{\sqrt{x}} \ p_i}{\sqrt{2 \ h}}, \frac{\sqrt{x} \ q_i - \frac{i}{\sqrt{x}} \ p_i}{\sqrt{2 \ h}} \right] = \frac{x}{2 \ h} [q_i, q_i] + \frac{i}{2 \ h} [p_i, q_i] - \frac{i}{2 \ h} [q_i, p_i] + \frac{1}{2 \ h x} [p_i, p_i] = -\frac{i}{h} [q_i, p_i] = -\frac{i}{h} \frac{\hbar}{\hbar} = 0 \]  
(7.29)

More generally, we can prove that

\[ [a_i, a_j^\dagger] = \delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \]  
(7.30)

For the example that we considered above, there is only one coordinate and one momentum, \( q \) and \( p \), and thus, we only have one creation operator and one annihilation operator

\[ a^\dagger = \frac{\sqrt{x} \ q - \frac{i}{\sqrt{x}} \ p}{\sqrt{2 \ h}} \]  
(7.31)

\[ a^\dagger = \frac{\sqrt{x} \ q + \frac{i}{\sqrt{x}} \ p}{\sqrt{2 \ h}} \]  
(7.32)

### 7.2.2. Vacuum

For simplicity, from now on, we will only focus on the example that discussed above, but please keep in mind that all conclusions can be generalized easily to more complicated Hamiltonians.

Here we make two assumptions:

**Assumption #1:** there is one (and only one) quantum state such that

\[ a | 0 \rangle = 0 \]  
(7.33)

We call this state the vacuum state.

**Assumption #2:** The states \( | 0 \rangle, a^\dagger | 0 \rangle, a^\dagger a^\dagger | 0 \rangle, a^\dagger a^\dagger a^\dagger | 0 \rangle \ldots \) form a complete basis of the Hilbert space.

Comment: These assumptions turn out to be true for most cases that we study in quantum mechanics.

### 7.2.3. particle number operator

Define particle number operator

\[ n = a^\dagger a \]  
(7.34)

We can easily prove that this operator is a Hermitian operator, \( n^\dagger = n \)

\[ n^\dagger = (a^\dagger a)^\dagger = a^\dagger (a^\dagger)^\dagger = a^\dagger a = n \]  
(7.35)
So this operator has real eigenvalues (which could be a physics measurable quantity).

We can further prove that
\[ n \left( \left| 0 \right> \right) = n \left( \left| 0 \right> \right) = a^\dagger a \left| 0 \right> = a^\dagger (a^\dagger a) | 0 \rangle = a^\dagger x 0 = 0 \] (7.36)
i.e. the vacuum state \( | 0 \rangle \) is an eigenstate of \( n \) with eigenvalue 0.

\[ n \left( \left| 0 \right> \right) = n \left( \left| 0 \right> \right) = a^\dagger a \left| 0 \right> = a^\dagger (a^\dagger a - a^\dagger a + a^\dagger a) | 0 \rangle = a^\dagger (a^\dagger a + a^\dagger a - a^\dagger a) | 0 \rangle = a^\dagger (1 + a^\dagger a) | 0 \rangle = a^\dagger | 0 \rangle + a^\dagger a^\dagger a | 0 \rangle = a^\dagger | 0 \rangle \] (7.37)
The last term is 0, because \( a | 0 \rangle = 0 \) by definition.

In other words, we find that \( a^\dagger | 0 \rangle \) is also an eigenstate of \( n \) with eigenvalue 1.

We can do the same thing and prove that \( \left( a^\dagger \right)^m | 0 \rangle \) is an eigenstate for the number operator \( n \), and the eigenvalue is \( m \).

**Bottom line #1:** the eigenvalues of the operator \( n \) is a non-negative integer 0, 1, 2, … Notice that number of particles are also non-negative integers, so it makes sense to call this operator the particle number operator. Here, the operator \( n \) measures the number of some fictitious particles.

**Bottom line #2:** \( \left( \left( a^\dagger \right)^m | 0 \right) \rangle \) with \( m \) being 0, 1, 2, … for a complete basis of the Hilbert space (as assumed above). Each state in this basis is an eigenstate of \( n \).

### 7.2.4. Hamiltonian: the harmonic part

\[ H_0 = \frac{\hat{p}^2}{2m r^2} - m g r \cos \hat{q} = \frac{\hat{p}^2}{2m r^2} + \frac{mg r}{2} \hat{q}^2 - \frac{mg r}{24} \hat{q}^4 + \ldots \] (7.38)

Let's keep only the first two terms, and ignore all higher order terms of \( \hat{q} \)

\[ H_0 = \frac{\hat{p}^2}{2m r^2} + \frac{mg r}{2} \hat{q}^2 \] (7.39)

The part that we ignored are

\[ \delta H = -\frac{mg r}{24} \hat{q}^4 + \ldots \] (7.40)

and \( H = H_0 + \delta H \)

For the harmonic part \( H_0 \)

\[ H_0 = \frac{\hat{p}^2}{2m r^2} + \frac{mg r}{2} \hat{q}^2 \] (7.41)

Because

\[ a^\dagger = \frac{\sqrt{x} q - \frac{1}{\sqrt{x}} i p}{\sqrt{2 \hbar}} \] (7.42)
\[ a = \frac{\sqrt{x} q + \frac{1}{\sqrt{x}} i p}{\sqrt{2 \hbar}} \] (7.43)

we know that

\[ q = \frac{\sqrt{\hbar}}{\sqrt{2x}} (a^\dagger + \bar{a}) \] (7.44)
\[ p = i \sqrt{\frac{x \hbar}{2}} (a^\dagger - \bar{a}) \] (7.45)
Thus,
\[
H_0 = \frac{\hbar^2}{2 m r^2} + \frac{m g r}{2} \hat{q}^2 + \frac{m g r}{2} \left( \frac{\sqrt{\hbar}}{\sqrt{2 x}} (a^+ - a) \right)^2 = - \frac{x}{4 m r^2} (a^+ - a)^2 + \frac{m g r}{4 x} \left( a^+ a + a a^+ \right)^2
\]  
(7.46)

Here
\[
(a^+ - a)^2 = (a^+ - a)(a^+ - a) = a^+ a^+ - a^+ a - a a^+ + a a
\]  
(7.47)

and
\[
(a^+ + a)^2 = (a^+ + a)(a^+ + a) = a^+ a^+ + a^+ a + a a^+ + a a
\]  
(7.48)

So
\[
H_0 = \left( - \frac{x}{4 m r^2} + \frac{m g r}{4 x} \right) (a^+ a^+ + a a) + \left( \frac{x}{4 m r^2} + \frac{m g r}{4 x} \right) (a^+ a + a a^+)
\]  
(7.49)

Now, we choose the value of \( x \), such that the first term vanish
\[
- \frac{x}{4 m r^2} + \frac{m g r}{4 x} = 0
\]  
(7.50)

i.e.
\[
x^2 = m^2 g r^3
\]  
(7.51)

and thus
\[
x = \sqrt{m^2 g r^3} = m r \sqrt{g r}
\]  
(7.52)

And thus
\[
H_0 = \left( \frac{x}{4 m r^2} + \frac{m g r}{4 x} \right) (a^+ a^+ + a a) + \left( \frac{x}{4 m r^2} + \frac{m g r}{4 x} \right) (a^+ a + a a^+)
\]  
\[
= \hbar \sqrt{\frac{g}{r}} \left( \frac{a^+ a + a a^+}{2} \right)
\]  
(7.53)

Because \([a, a^+] = a a^+ - a^+ a = 1\), \(a a^+ = a^+ a + 1\) and thus
\[
H_0 = \hbar \sqrt{g/r} \left( \frac{a^+ a + a a^+}{2} \right) = \hbar \sqrt{g/r} \left( \frac{2 a^+ a + 1}{2} \right) = \hbar \sqrt{g/r} \left( a^+ a + \frac{1}{2} \right)
\]  
(7.54)

Notice that \(\sqrt{g/r} = \omega\) is actually the angular frequency of the harmonic oscillator, and \(a^2 a\) is in fact the particle number operator \(n\) we defined above
\[
H_0 = \hbar \omega \left( n + \frac{1}{2} \right)
\]  
(7.55)

We can understand this Hamiltonian as the following: (1) each of our fictitious particle has energy \(\hbar \omega\). (2) The system also have an zero point energy (energy of the vacuum) \(\frac{\hbar \omega}{2}\). So the total energy of the system is
\[
H_0 = \hbar \omega \left( a + \frac{1}{2} \right)
\]  
(7.56)

Because the energy is carried by these fictitious particles, the total energy is proportional to particle number (up to the zero point energy \(\hbar \omega/2\)). Because particle number are quantized integers, the energy is quantized.

This is the origin of quantization
7.2.5. Hamiltonian: the non-harmonic part

How about the terms that we ignored?

We can treat them as interactions between particles (i.e. scatterings).

The bottom line:

(1) for a harmonic oscillator, we can treat it as non-interacting particles (each has energy $\hbar \omega$)

(2) for non-harmonic terms, we can treat term as interactions between these particles

7.2.6. How about waves?

For a harmonic oscillator, we know that we can think about it using these fictitious particles. This fictitious particles cannot move (they are confined spatially at the oscillator). This is a bit different from a real particle, which can move in space. How can we have a real particle? Using waves, instead of a single oscillator.

Waves is nothing but couple harmonic oscillators (we put a harmonic oscillator at each position $x$, and couple them together). For each position $x$, we have one harmonic oscillator, and thus we can define a creation operator at this point $a^\dagger(x)$ and also an annihilation operator $a(x)$ at this point.

We can also ask what is the particle number for each oscillator $n(x) = a^\dagger(x) a(x)$

The coupling between harmonic oscillators allows the particle at $x$ to move to $y$ (i.e. motion of a particle).

Non-harmonic parts gives interactions between this particles.

This is the foundation of particle wave duality.