Notes for Lecture 4

Time independent perturbation – examples, degenerate case

Here, we will study some examples of the perturbation theory. A new concept that we will introduce is the "degenerate perturbation."

We will see below that the expression "degenerate perturbation" is something of an oxymoron. Regardless, the physics of the degenerate perturbation theory is very important. Its physics is much more important than the physics of the non-degenerate perturbation. The good news is that, when properly understood, there is no real new formula to learn in the degenerate perturbation theory. Essentially all work has been done in the last LN, even for the degenerate perturbation case.

4.1 New symbols

The symbol $\doteq$ will be used to mean "is represented by." The symbol $\fallingdotseq$ will be used to mean "represents." These newly established symbols are noted in the updated LN 2.

For example, $\hat{p}_x \doteq -i\hbar\frac{\partial}{\partial x}$ (x representation), $\hat{p}_x \doteq \hbar k_x$ (k representation), and $\hat{S}_z \doteq \frac{\hbar}{2}\left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$, where in the latter example the S_z representation of a spin 1/2 system is considered. The ~~converse~~ equivalent statements are $-i\hbar\frac{\partial}{\partial x} \fallingdotseq \hat{p}_x$, $\hbar k_x \fallingdotseq \hat{p}_x$, and $\left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right) \fallingdotseq \hat{S}_z$.

4.2 Examples of non-degenerate perturbation

Let us consider couple of simple examples, to which we can apply the theory as we learned in the last lecture.

4.2.1 1D ∞ potential well example

Let us consider an infinite potential well in one dimension. $V(x) = 0$ for $0 < x < L$, while $V(x) = \infty$ otherwise. The energy eigenstate for this problem is of the form $A\sin(kx)$, where A is the normalization constant to determine. I will leave it up to you to determine A if you decide to do so. The wave vector k is quantized, since the wave function must vanish at $x = 0, L$ (why should it vanish?). Here, let us simply ask the following question.

A perturbation, $V_1(x) = V_0$ for $0 < x < L/2$ and zero otherwise, is applied, so that the total Hamiltonian is $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}) + V_1(\hat{x})$. What is the leading order correction to the energy eigenvalue?

To answer this, we do not need to know A. In fact, we do not know even that the wave function is sine form, if we just use the symmetry property, and the normalization condition.

As a starting point, we assume that

$$\hat{H}_0 \left|n^{(0)}\right\rangle = E_n^{(0)} \left|n^{(0)}\right\rangle \qquad\qquad \hat{H}_0 \equiv \frac{\hat{p}^2}{2m} + V(\hat{x}) \tag{4.1}$$

Arguably, $\xi \equiv x - \frac{L}{2}$ would be a better variable to bring out the symmetry clearly, since the potential V is has the parity symmetry, i.e. V remains unchanged when $\xi \leftrightarrow -\xi$. I will leave it up to you to re-express everything here in terms of ξ, rather than x, if that is what you prefer. In any case, as noted in page 20 of LN 2, in this case the parity – i.e. the evenness or the oddness of the state with respect to the center of the potential well – is conserved. As noted[1] in page 20 of LN 2, the energy eigenstate for $\hat{H}_0$ can be taken as the parity eigenstate, and so we assume that we did it. Thus, for any $\psi_n(x) \equiv \langle x \,|\, n\rangle$, we have

$$\psi_n(x)\big|_{x-\frac{L}{2} \to -\left(x-\frac{L}{2}\right)} = \pm\psi_n(x) \tag{4.2}$$

[1] We did not *prove* it there, as we were reviewing the subject. However, if you like to prove it from scratch, then it is not terribly difficult. Just use the fact $\left[\hat{H}_0, \hat{P}_\xi\right] = 0$ in the current case to show that an energy eigenstate *can* (not must; cf. Homework 1, problem 4) be chosen to be a parity eigenstate: $\hat{P}_\xi \left|n^{(0)}\right\rangle = \left|n^{(0)}\right\rangle$ or $-\left|n^{(0)}\right\rangle$.

Note that this means

$$\psi_n^*(x)\psi_n(x)|_{x-\frac{L}{2}\to-(x-\frac{L}{2})} \;=\; (\pm 1)^2\,\psi_n^*(x)\psi_n(x) \;=\; \psi_n^*(x)\psi_n(x) \tag{4.3}$$

I.e., the probability density $\rho_n(x) \equiv \psi_n^*(x)\psi_n(x)$ has the even parity. Now, let us calculate the first order energy correction, Eq. 3.16, with $\hat{H}_1 = V_1(\hat{x})$. Before we calculate this correction, note that it is possible to split any function into an even part and odd part. For the current function $V_1(x)$, we get

$$V_1(x) \;=\; V_{1,e}(x) + V_{1,o}(x) \tag{4.4}$$

$$V_{1,e}(x) \;=\; \frac{V_0}{2} \quad \text{if} \quad 0 < x < L \tag{4.5}$$

$$V_{1,o}(x) \;=\; \frac{V_0}{2} \quad \text{if} \quad 0 < x < \frac{L}{2}, \quad -\frac{V_0}{2} \quad \text{if} \quad \frac{L}{2} < x < L \tag{4.6}$$

where $V_1 = V_{1,e} = V_{1,o} = 0$ if $x < 0$ or $x > L$. $V_{1,e}(x)$ is an even function of ξ and $V_{1,o}(x)$ is an odd function of ξ.

The first order correction

$$E_n^{(1)} \;=\; \left\langle n^{(0)} \,|V_1(\hat{x})\,|\, n^{(0)} \right\rangle \tag{4.7}$$

$$\;=\; \int_0^L dx\; (\rho_n(x)V_{1,e}(x) + \rho_n(x)V_{1,o}(x)) \qquad \rho_n(x) = \psi_n^*(x)\psi_n(x) \tag{4.8}$$

$$\;=\; \int_0^L dx\,\rho_n(x)V_{1,e}(x) \qquad \text{parity w.r.t. } x = L/2 \tag{4.9}$$

$$\;=\; \frac{V_0}{2}\int_0^L dx\,\rho_n(x) \tag{4.10}$$

$$\;=\; \frac{V_0}{2} \qquad \text{since } \left\langle n^{(0)} \,\middle|\, n^{(0)} \right\rangle = 1 \tag{4.11}$$

This is the leading order correction. (If you have couple of minutes, and if necessary, please convince yourself that you could have obtained this solution, without writing anything down.) The second order correction will be much smaller than this. We will not calculate it here. However, if we were to calculate it, we would find that the above symmetry argument is not enough (A needs to be evaluated and used).

Let us ask the question "what is the perturbation parameter in this case?". Since $k_n = \pi n/L$ ($n = 1, 2, ...$: the wave length must be an integer fraction of $2L$, if the wave needs to vanish at 0 and L), we get

$$E_n^{(0)} \;=\; \frac{\hbar^2 k_n^2}{2m} \;=\; \frac{\hbar^2\pi^2 n^2}{2mL^2} \qquad n = 1, 2, 3, ... \tag{4.12}$$

and thus the energy spacing is given by

$$\Delta E_n^{(0)} \;=\; \frac{\hbar^2\pi^2(2n+1)}{2mL^2} \qquad n = 1, 2, 3, ... \tag{4.13}$$

So, the perturbation parameter is

$$\delta \sim \frac{E_n^{(1)}}{\Delta E_n^{(0)}} = \frac{1}{\pi^2(2n+1)} \frac{mV_0L^2}{\hbar^2} \tag{4.14}$$

Please double check that this quantity is indeed dimensionless (note that $\hbar$ has the dimension of angular momentum, or the *action*, $\int dt\, L$, where L is the Lagrangian).

4.2.2 Larmor precession example

Let us take a simple spin problem: an electron in a magnetic field. We shall assume that the electron does not have any orbital angular momentum. An example would be an electron in $1s$ orbital of an hydrogen atom. In the presence of a magnetic field, $\vec{B}_0$

$$\hat{H} = -\hat{\vec{\mu}} \cdot \hat{\vec{B}}_0 \tag{4.15}$$

where $\vec{\mu}$ is the magnetic moment due to spin only (we shall shortly discuss $\vec{\mu}$ more generally). We assume a time-independent uniform field. Then, in the real space ($\vec{x}$) representation, $\hat{\vec{B}}_0 \doteq B_0 \vec{e}_z$ ($\vec{e}_z$ is the unit vector along the z direction; we do not use the notation $\hat{z}$, since that can be misunderstood as a QM operator) where B_0 is a constant. Also, for an electron, the magnetic moment

$$\vec{\mu} = -\frac{g\mu_B \vec{S}}{\hbar} \qquad \text{electron, magnetic moment due to spin} \tag{4.16}$$

where

$$g = 2.002319... \qquad \text{gyromagnetic ratio for electron spin} \tag{4.17}$$

$$\mu_B = \frac{e\hbar}{2m_e} \qquad \text{Bohr magneton} = 5.788 \times 10^{-5}\ \text{eV/T} \tag{4.18}$$

Here, the Bohr magneton is given in the SI unit. In the cgs unit, it is given as $\mu_B = \frac{e\hbar}{2m_e c}$, and its value is 5.788×10^{-9} eV/gauss. Note that the proper SI unit is J/T (T for Tesla), and the proper cgs unit is erg/gauss (10,000 gauss = 1 Tesla). But in QM, eV is a much more useful unit, and so the units used above. Note that e is the absolute value of electron charge (1.6×10^{-19} C). However, a *much better way to know* e is through the fine structure constant $\alpha \equiv e^2/(4\pi\epsilon_0\hbar c)$ (SI unit) $= 1/137.0 = e^2/(\hbar c)$ (cgs), along with $\hbar c = 1973$ eVÅ. Lastly, m_e is the electron mass $m_e c^2 = 0.511$ MeV.

The gyromagnetic ratio for the orbital angular momentum is 1: i.e., the magnetic moment due to the *orbital* motion of an electron is given by

$$\vec{\mu}_l \;=\; -\frac{\mu_B \vec{L}}{\hbar} \qquad \text{electron, } \vec{L} = \text{orbital angular momentum} \tag{4.19}$$

Although this contribution is of no concern for the current problem, it is nevertheless helpful to derive it. Imagine a classical charge q (e.g. $-e$ for electron) rotating in a circle or radius r with angular frequency ω. Recall from E&M that the magnetic moment is defined as $I\vec{A}$, for a current loop, where $\vec{A}$ is the area vector, whose direction is determined from the direction of the current I by the right hand/screw rule. Thus, the magnitude of the magnetic moment is given by $\frac{|q|\omega}{2\pi}\,\pi r^2 = |q|mvr/(2m) = |q|L/(2m)$. For electron, $q = -e$, and so we get $|\vec{\mu}| = \mu_B L/\hbar$. Regarding the direction, we see that for a positive charge, the angular momentum and the magnetic moment are in the same direction, while they are in the opposite direction for a negative charge. Thus, the above formula for the electron is proved.

Now, back to the spin only problem. It took Dirac's relativistic QM ($g = 2$) and Feynman's QED to see why in the spin case g is not unity, but slightly greater than 2. Because the true value for g is very close to 2, it is acceptable to use Dirac's value $g = 2$ for explaining many experiments, when their finite energy/frequency resolution does not require knowing g beyond 2.00. Here, we will pretend that we do *not* know the value of g, and simply put

$$\mu_g \;\equiv\; \frac{g}{2}\,\mu_B \tag{4.20}$$

Then, the above Hamiltonian can be represented as (also, note that we change the notation: $\hat{H} \to \hat{H}_0$)

$$\hat{H}_0 \;\doteq\; \mu_g B_0 \sigma_z \;=\; \mu_g B_0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.21}$$

What is the typical magnitude of $\mu_B B_0$ ($\mu_g \approx \mu_B$)? Man-made magnetic field is ~ 100 T at the most, and typically much less. Thus, $\mu_B B_0$ is a few meV or much less. Recalling that the typical energy spacing of atomic energy levels is ~ 1 eV or higher, the energy due to an external magnetic field is very small. So, the above Hamiltonian could be considered a perturbation term (Zeeman term) for an atomic energy level. We will have a chance to do later on, and for the Zeeman term physics, including both the spin physics and the orbital physics. In the current example, however, we simply consider the above Hamiltonian as the unperturbed Hamiltonian, setting aside any orbital physics for the time being.

As the above Hamiltonian is diagonal already, the energy eigenvalues can be simply read off as $\pm\mu_g B_0$. **Under this Hamiltonian, a general spin state will precess**

with the Larmor frequency, $\omega_L = 2\mu_g B_0/\hbar$. The simplest way to see this is to use what we learned in Section 2.1.4 to see that the time evolution operator is $\hat{\mathcal{U}}(t) = \exp\left(-i\hat{H}t/\hbar\right) = \exp\left(-i2\mu_g B_0 \hat{S}_z t/\hbar^2\right)$, which is recognized as a spin rotation operator (Section 2.1.4), with the rotation angle $\Delta\theta = 2\mu_g B_0 t/\hbar$ from time 0 to t, which in turn means that the spin rotates at an angular frequency/velocity of $2\mu_g B_0/\hbar$, which we identify as the Larmor precession frequency ω_L.

It should be recognized that $\hbar\omega_L = 2\mu_g B$ = the energy difference ($\equiv \Delta E$ in this paragraph) between the ground state and the excited state of the above Hamiltonian $\hat{H}_0$. Why is this? First, note that the Larmor frequency represents the "classical natural frequency" of this system in the sense that $\left\langle \hat{\vec{S}} \right\rangle$ rotates at that frequency. Thus, if we are doing an IR absorption experiment the system will respond resonantly when the photon's frequency matches ω_L. Second, on the other hand, such a resonant absorption has a more lucid interpretation in QM. When the light frequency is such that $\hbar\omega$ matches the energy difference between the ground state and the excited state, then we have a resonant condition. This is why $\hbar\omega_L$ must be equal to $2\mu_g B$. Third, note that for two levels with different energy values, the expectation value of any observable $\hat{O}$ will have the time dependence at frequency given by $\Delta E/\hbar$. This is of course the general principle that we call "energy time uncertainty principle."

Now, suppose you were doing a spin precession experiment[2] like this, with a moderate field, $B_0 = 1$ Tesla for example. You realize that you did not take into account the earth magnetic field ($\vec{B}_E$), and would like to estimate its effect on your experiment. Since the earth magnetic field is very small (~ 0.5 gauss) compared with a Tesla field, we can apply the perturbation theory. Let us say that

$$\vec{B}_E = B_1\vec{e}_3 + B_2\vec{e}_1 \tag{4.22}$$

I.e., the earth magnetic field is assumed to lie in the xz plane, with the z axis being the direction of the applied field. Then, the **perturbing Hamiltonian** is given by

$$\hat{H}_1 \doteq \mu_g \begin{pmatrix} B_1 & B_2 \\ B_2 & -B_1 \end{pmatrix} \tag{4.23}$$

where $\sigma_x = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ has been used.

What is the first order correction for energy? What is the second order correction?

[2] As mentioned in the previous paragraph, this would be an experiment where an IR photon is resonantly absorbed by your prepared spin state. The prepared spin state must have components in both $|\uparrow\rangle$ and $|\downarrow\rangle$ for this experiment to work.

For the eigenstates of $\hat{H}_0$, i.e.,

$$|\uparrow\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad z\text{-spin up} \tag{4.24}$$

$$|\downarrow\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad z\text{-spin down} \tag{4.25}$$

we get

$$E_\uparrow^{(1)} = \left\langle \uparrow \left| \hat{H}_1 \right| \uparrow \right\rangle = \mu_g B_1 \tag{4.26}$$

$$E_\downarrow^{(1)} = \left\langle \downarrow \left| \hat{H}_1 \right| \downarrow \right\rangle = -\mu_g B_1 \tag{4.27}$$

How about the second order? Since the Hilbert space is two dimensional, the sum "$\sum_{m\neq n}$" involves only one term.

$$E_\uparrow^{(2)} = \frac{\left|\left\langle \downarrow \left| \hat{H}_1 \right| \uparrow \right\rangle\right|^2}{\mu_g B_0 - (-\mu_g B_0)} = \mu_g \frac{B_2^2}{2B_0} \tag{4.28}$$

$$E_\downarrow^{(2)} = \frac{\left|\left\langle \uparrow \left| \hat{H}_1 \right| \downarrow \right\rangle\right|^2}{-\mu_g B_0 - \mu_g B_0} = -\mu_g \frac{B_2^2}{2B_0} \tag{4.29}$$

Collecting these values, we get

$$\begin{aligned} E_\uparrow &= \mu_g B_0 + \mu_g B_1 + \mu_g \frac{B_2^2}{2B_0} + \ldots \\ &= \mu_g B_0 \left[1 + \frac{B_1}{B_0} + \frac{1}{2}\left(\frac{B_2}{B_0}\right)^2 + \ldots \right] \end{aligned} \tag{4.30}$$

$$E_\downarrow = -\mu_g B_0 \left[1 + \frac{B_1}{B_0} + \frac{1}{2}\left(\frac{B_2}{B_0}\right)^2 + \ldots \right] \tag{4.31}$$

The perturbed eigenstate is given as

$$|\uparrow\rangle_{\text{pert}} = |\uparrow\rangle + \frac{\left\langle \downarrow \left| \hat{H}_1 \right| \uparrow \right\rangle}{\mu_g B_0 - (-\mu_g B_0)} |\downarrow\rangle + \ldots = |\uparrow\rangle + \frac{B_2}{2B_0}|\downarrow\rangle + \ldots \tag{4.32}$$

$$|\downarrow\rangle_{\text{pert}} = |\downarrow\rangle + \frac{\left\langle \uparrow \left| \hat{H}_1 \right| \downarrow \right\rangle}{-\mu_g B_0 - \mu_g B_0} |\uparrow\rangle + \ldots = |\downarrow\rangle - \frac{B_2}{2B_0}|\uparrow\rangle + \ldots \tag{4.33}$$

These answers make it very clear that the perturbation parameters are B_1/B_0 and B_2/B_0. In order to reduce the effect of the Earth magnetic field on the experiment, it would be the best if $B_1 = 0$, i.e. if the applied field is perpendicular to the Earth magnetic field, since then the Larmor frequency changes only by second order. In any case, this problem is completely solvable for $\hat{H}_0 + \hat{H}_1$, and the exact solution can be shown to expand to the above solutions in the small $\hat{H}_1$ limit.

4.3 What is perturbation, really?

In the last LN, we saw that the perturbation parameter for the non-degenerate perturbation theory is the matrix element of $\hat{H}_1$ divided by the energy level spacing of unperturbed states. This actually makes a lot of sense, if you give it some thoughts. If you look at the main results of the last LN, you will see that as the result of the perturbation, the energy levels shift by $E_n^{(1)} \sim$ the matrix element of $\hat{H}_1$. And so, if the perturbation parameter is small, then *the "identities" of energy levels remain intact*, because the shifts in energies are negligible compared to the manner in which the energy levels are ordered before the perturbation is considered. This is the real sense of the perturbation! An energy level may be perturbed a little, but you can track it down to the original state. A large effect that can completely destroy the identity of an energy level cannot be called perturbation, in any reasonable sense.

In the "degenerate perturbation theory," this is exactly what happens – a large change that can completely destroy the identity of an energy level. So, this is why you might consider the expression "degenerate perturbation" an oxymoron.

Regardless of the name, the degenerate perturbation is very important, since its effect, if it coexists with non-degenerate perturbation effect, is much greater, as we shall see soon.

The correct way to understand what we mean by the "degenerate perturbation theory" is the following. Let us say that there are two degenerate energy levels. If we give them some identities, like A and B, then they are completely destroyed by any small perturbation. Come to think of it, degenerate levels do not have good identities in the first place! So, we can turn the logic around a little bit. We come up with a good way to give them identities. If we succeed in doing that, then, instead of destroying identities, we actually are creating identities. After that, all we need to do is to *re-apply* what we learned in the last LN!

4.4 Degenerate perturbation

For the non-degenerate perturbation theory of the last LN to hold, the energy levels must be distinct (Eq. 3.13).

In the degenerate perturbation theory, we assume the opposite[3].

$$E_n^{(0)} \;=\; E_m^{(0)} \qquad\qquad \text{for some } m\text{'s} \tag{4.34}$$

We define

$$D_n \equiv \{m \,|\, E_n^{(0)} \;=\; E_m^{(0)}\} \tag{4.35}$$

Note that the set D_n *includes* n itself, and the degenerate perturbation theory is necessary if the number of elements in D_n is 2 or greater.

It should be obvious why we cannot apply the old formulae, like Eq. 3.17 or 3.18, in the current case. For those m values belonging in D_n, we will have *divergent* terms!

What do we do in the degenerate case? It is quite simple actually. **We diagonalize the "sub-matrix" that represents the total Hamiltonian $\hat{H} = \hat{H}_0 + \hat{H}_1$ in the "sub-space," i.e. the small Hilbert space spanned only by $\{|m^{(0)}\rangle \mid m \in D_n\}$.** As before, $\hat{H}_0$ is the unperturbed Hamiltonian and $\hat{H}_1$ is the perturbation.

This is all there is, really! If the number of elements of D_n is small, like 2 or 3, which is a typical case we will deal with, then diagonalizing $\hat{H}$ in the sub-space, whose dimension is equal to the number of elements of D_n, is not a big deal.

You ask: OK, the sub-matrix is diagonalized, and what comes next? The answer is the following. If the sub-matrix diagonalization results in distinct $\hat{H}$ eigenvalues in the sub-space, then you are good to go, since you can **now use the non-degenerate perturbation theory of the last lecture with the new basis!** If the degeneracy in question is not lifted by $\hat{H}_1$, however, then your perturbation is doomed: it is a no go! There is nothing you can do within the perturbation theory. From now on, we will not consider this rare pathetic case[4] that does not permit a perturbation approach; we shall assume that the degeneracy of $\hat{H}_0$ is lifted by $\hat{H}_1$ within the sub-space.

Now, note that the new basis of the sub-space consists of linear combinations of states in the original basis set

$$B_n \equiv \{\left|m^{(0)}\right\rangle \mid m \in D_n\} \tag{4.36}$$

The new basis that diagonalizes the sub-matrix can be defined as

$$B'_n \equiv \{\left|m^{(0)\prime}\right\rangle \mid m \in D_n\} \tag{4.37}$$

[3]Here is a question that you can ponder. What if the equality symbol here is replaced by the symbol $\approx$? Would we then need to use the degenerate perturbation or the non-degenerate perturbation? The answer depends on how big the difference between $E_n^{(0)}$ and $E_m^{(0)}$ is compared to the matrix element of $\hat{H}_1$.

[4]It is often said that "Nature does not like the degeneracy." This is quite true. Unless protected by symmetry, the degeneracy is lifted by residual interactions, as a rule. So, what we assume here is very reasonable in reality.

where each $|m^{(0)\prime}\rangle$ is a linear combination of $|j^{(0)}\rangle$'s that belong in B_n. By assumption, the matrix representation of $\hat{H}_0$ using either basis set is $E_n^{(0)}$ times the identity matrix, and the matrix representation of $\hat{H}_1$ is diagonal with B'_n.

We form a new global basis set, by combining B'_n with the set of all $|m^{(0)}\rangle$'s with $m \notin D_n$, and proceed as in the last LN.

Zero-th order

$|n^{(0)\prime}\rangle$ and $E_n^{(0)}$ is the answer.

First order

The energy correction is $E_n^{(1)} = \left\langle n^{(0)\prime} \left| \hat{H}_1 \right| n^{(0)\prime} \right\rangle$. However, this has been already calculated in the above sub-matrix diagonalization step, since the diagonal elements of that sub-matrix are given by $E_n^{(0)} + \left\langle n^{(0)\prime} \left| \hat{H}_1 \right| n^{(0)\prime} \right\rangle = E_n^{(0)} + E_n^{(1)}$. **The first order correction in energy was calculated automatically in the sub-matrix diagonalization process, which gives energy eigenvalues $E_n^{(0)} + E_n^{(1)}$.**

For the eigenstate, we use Eq. 3.17, with the new global basis set. Note that $\hat{H}_1$ has zero non-diagonal elements between states belonging in B'_n, and so Eq. 3.17 becomes

$$\left|n^{(1)\prime}\right\rangle = \sum_{m \notin D_n} \frac{|m^{(0)}\rangle \left\langle m^{(0)} \left| \hat{H}_1 \right| n^{(0)\prime} \right\rangle}{E_n^{(0)} - E_m^{(0)}} \tag{4.38}$$

Second order

Using the above result for the state, and using Eq. 3.12, we get

$$E_n^{(2)} = \sum_{m \notin D_n} \frac{\left| \left\langle m^{(0)} \left| \hat{H}_1 \right| n^{(0)\prime} \right\rangle \right|^2}{E_n^{(0)} - E_m^{(0)}} \tag{4.39}$$

Degenerate perturbation

simply means **rebooting the problem so that it becomes a non-degenerate perturbation problem!**

The "rebooting" means the diagonalization of the sub-matrix of $\hat{H} = \hat{H}_0 + \hat{H}_1$ in the sub-space spanned by those states with state indices $m \in D_n$, where D_n is a set of indices for all degenerate states with $|n^{(0)}\rangle$ under $\hat{H}_0$. If the rebooting is successful, i.e. if the degeneracy is lifted by the diagonalization, then we can proceed with the non-degenerate perturbation method!

The "rebooting" means finding good zeroth order states $\{|n^{(0)\prime}\rangle\}$ to replace $\{|n^{(0)}\rangle\}$. The good zeroth order states are non-degenerate eigenstates of $\hat{H}$ in the $\hat{H}_0$-degenerate sub-space defined by D_n. If the rebooting is successful, then we are good for the zeroth order state *and* the first order correction in energy.

The first order correction in energy is precisely given by the difference in energy eigenvalue of the sub-matrix and $E_n^{(0)}$. For the first order correction in ket and the second order correction in energy, simply re-use the formulae for the non-degenerate case with the following modification: $m \neq n \to m \notin D_n$, and $|n^{(0)}\rangle \to |n^{(0)\prime}\rangle$.

4.5 Degenerate perturbation is more important

In general, for the same matrix elements of $\hat{H}_1$, the degenerate perturbation has a much stronger effect on the system than the non-degenerate perturbation does. So, this is why the degenerate perturbation theory is very important.

It is instructive to consider a simple case where the size of D_n is 2.

In the following two pages is a note to study with care. Also, this note should, I am hoping, rid you of any uncertainty about what it means to represent an operator with basis vectors.

Here in this note, the diagonal matrix elements of $\hat{H}$ are defined as

$$\varepsilon_j \equiv \left\langle j^{(0)} \left| \hat{H} \right| j^{(0)} \right\rangle \tag{4.40}$$

$$\equiv E_j^{(0)} + \left\langle j^{(0)} \left| \hat{H}_1 \right| j^{(0)} \right\rangle \tag{4.41}$$

So, the diagonal term has some correction due to the perturbation. This correction would be all there is for the first order correction in energy if the non-degenerate perturbation were applicable.

The reason why the degenerate perturbation is important is because off-diagonal matrix elements (e.g. "h_1" in the note) also have the first order effect if degenerate perturbation is applicable.

[This paragraph, which is a guide to the hand-written note in the following two pages, has been polished for improved clarity.] In the following two-page hand-written note, it is assumed that only two states, $|n^{(0)}\rangle$ and $\left|n+1^{(0)}\right\rangle$, are degenerate under $\hat{H}_0$. In the second page, it is *further assumed* that $\left\langle n^{(0)} \left| \hat{H}_1 \right| n^{(0)} \right\rangle = \left\langle n+1^{(0)} \left| \hat{H}_1 \right| n+1^{(0)} \right\rangle$ so that $\varepsilon_n = \varepsilon_{n+1}$. While this latter assumption is not essential to show that the degenerate perturbation has a greater effect than the non-degenerate perturbation, it is nevertheless a simplifying assumption for the discussion in that page. It is also an assumption that is frequently satisfied in degenerate perturbation problems. However, keep in mind that certain results are justified in a precise form only under this latter assumption – e.g. the result that *exactly half-and-half* mixture of the original unperturbed basis states contribute to each of the good unperturbed basis states. **However, an important point to take away is that, in a general degenerate perturbation case, the original states, $|n^{(0)}\rangle$ and $\left|n+1^{(0)}\right\rangle$, *mix very strongly* and thus lose their role of defining identities of the perturbed states, while the mixing is not exactly a half and half mixing in general; in contrast, in a non-degenerate perturbation, those original states remain the main parts of, and thus define the identities/characters of, the perturbed states.** In short, it is important to note that, even if $\left\langle n^{(0)} \left| \hat{H}_1 \right| n^{(0)} \right\rangle \neq \left\langle n+1^{(0)} \left| \hat{H}_1 \right| n+1^{(0)} \right\rangle$, the degenerate perturbation case involves a "strong level repulsion" (linear order in h_1) (as the difference between ε_n and ε_{n+1} is linear order in perturbation), in contrast to a non-degenerate case, which involves a "weak level repulsion" (second order in h_1) (as the difference between ε_n and ε_{n+1} is zero-th order in perturbation), where h_1 is the *off-diagonal* matrix element of $\hat{H}_1$.

- Represent $\hat{H}$ using basis $|0^{(0)}\rangle, |1^{(0)}\rangle, \ldots$

$$\hat{H} = \hat{H}_0 + \hat{H}_1 \qquad \hat{H}_0 |n^{(0)}\rangle = E_n^{(0)} |n^{(0)}\rangle$$

$$\hat{H} \doteq \begin{array}{c|cccccccc} & |0^{(0)}\rangle & |1^{(0)}\rangle & \cdots & |n-1^{(0)}\rangle & |n^{(0)}\rangle & |n+1^{(0)}\rangle & |n+2^{(0)}\rangle & \cdots \\ \hline \langle 0^{(0)}| & \mathcal{E}_0 & & & & & & & \\ \langle 1^{(0)}| & & \mathcal{E}_1 & & & & & & \\ \vdots & & & \ddots & & & & & \\ \langle n-1^{(0)}| & & & & \mathcal{E}_{n-1} & & & & \\ \langle n^{(0)}| & & & & & \mathcal{E}_n & h_1 & & \\ \langle n+1^{(0)}| & & & & & h_1 & \mathcal{E}_{n+1} & & \\ \langle n+2^{(0)}| & & & & & & & \mathcal{E}_{n+2} & \\ \vdots & & & & & & & & \ddots \end{array}$$

An ∞ dimensional matrix.

No matrix element is assumed to vanish.

- The listed bra vectors and ket vectors are there to remind you that the matrix elements inside the big [ ] correspond to

$$\langle i^{(0)} | \hat{H} | j^{(0)} \rangle \qquad i, j = 0, 1, 2, \ldots$$

row index ↑ ↑ column index

- Note that the diagonal elements $\mathcal{E}_i = E_i^{(0)} + \langle i^{(0)} | \hat{H}_1 | i^{(0)} \rangle$ $\mathcal{E}_i \equiv \langle i^{(0)} | \hat{H} | i^{(0)} \rangle$ already contain 1st order corrections.

- Assume $E_n^{(0)} = E_{n+1}^{(0)}$ and all other $E_i^{(0)}$'s are distinct from $E_n^{(0)}$. (Typical degenerate perturbation case)

If the "rebooting" is successful, then we have

$$\hat{H} \doteq \begin{array}{c|cccccccc} & |0^{(0)}\rangle & \cdots & |n-1^{(0)}\rangle & |n^{(0)\prime}\rangle & |n+1^{(0)\prime}\rangle & |n+2^{(0)}\rangle & \cdots \\ \hline \langle 0^{(0)}| & \mathcal{E}_0 & & & & & & \\ \vdots & & \ddots & & & & & \\ \langle n-1^{(0)}| & & & \mathcal{E}_{n-1} & & & & \\ \langle n^{(0)\prime}| & & & & E_n' & 0 & & \\ \langle n+1^{(0)\prime}| & & & & 0 & E_{n+1}' & & \\ \langle n+2^{(0)}| & & & & & & \mathcal{E}_{n+2} & \ddots \end{array}$$

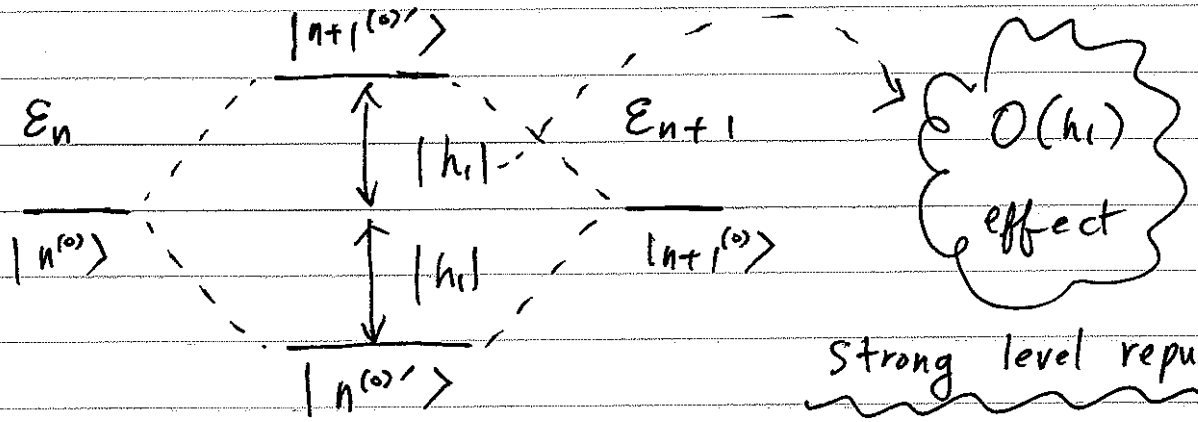
New basis

$|0^{(0)}\rangle \; |1^{(0)}\rangle \cdots |n-1^{(0)}\rangle$

$|n^{(0)\prime}\rangle \; |n+1^{(0)\prime}\rangle$

$|n+2^{(0)}\rangle \; \ldots$

Assume $\varepsilon_n = \varepsilon_{n+1}$ for simplicity.



Strong level repulsion

Both $|n^{(0)\prime}\rangle$ and $|n+1^{(0)\prime}\rangle$ are completely equal mixtures of $|n^{(0)}\rangle$ and $|n+1^{(0)}\rangle$.
(50% of $|n^{(0)}\rangle$ and 50% of $|n+1^{(0)}\rangle$)

$$|n^{(0)\prime}\rangle = \alpha\,|n^{(0)}\rangle + \beta\,|n+1^{(0)}\rangle \qquad |\alpha| = |\beta| = |\gamma| = |\delta| = \frac{1}{\sqrt{2}}$$

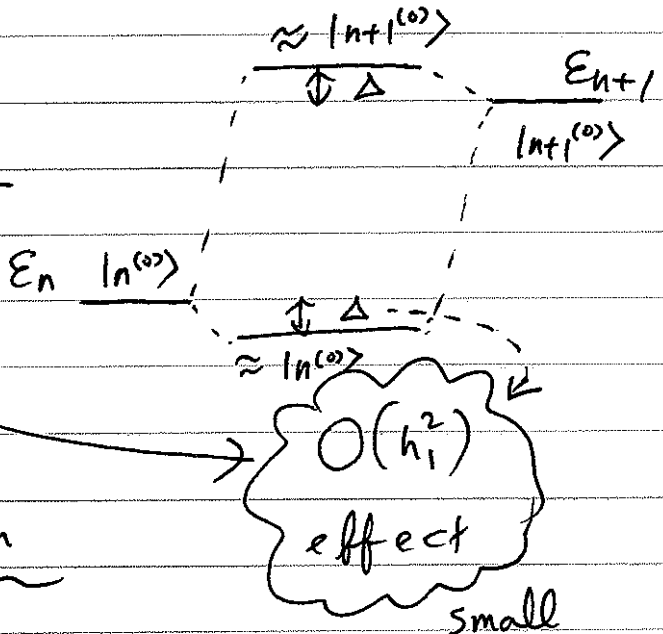
$$|n+1^{(0)\prime}\rangle = \gamma\,|n^{(0)}\rangle + \delta\,|n+1^{(0)}\rangle$$

Phases s.t. $\alpha^*\gamma + \beta^*\delta = 0$
(orthogonality)

So, the "identity" of $|n^{(0)}\rangle$ and $|n+1^{(0)}\rangle$ is completely lost in the new basis.

Non-degenerate case
(two state only problem)

$$\Delta \equiv \frac{|h_1|^2}{|\varepsilon_n - \varepsilon_{n+1}|}$$



Weak level repulsion

The "identity" of $|n^{(0)}\rangle$, $|n+1^{(0)}\rangle$ remain intact.