

§ 0.1 Perturbation Theory

Since most problems in quantum mechanics can't be solved directly, there are various methods in order to approximate the results. In perturbation theory, we have two main problems and two cases: time-independent and time-dependent perturbations with or without degeneration. As a first approach, we will consider nondegenerate and time independent perturbations.

§§ 0.1.1 Rayleigh-Schrödinger Perturbation Theory, Nondegenerate Case

Let $\hat{\mathcal{H}}$ be our non directly solvable Hamiltonian, divisible in a sum of a solvable Hamiltonian $\hat{\mathcal{H}}_0$ and a perturbation \hat{V} .

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{V} \quad (1)$$

For $\hat{\mathcal{H}}_0$ we already know the solution of the secular equation, and we label them as follows

$$\hat{\mathcal{H}}_0 |n_0\rangle = E_n^{(0)} |n_0\rangle \quad (2)$$

Where the 0 should be seen strictly as a label, otherwise told.

The full secular equation will then be the following

$$\hat{\mathcal{H}} |n\rangle = (\hat{\mathcal{H}}_0 + \hat{V}) |n\rangle = E_n |n\rangle \quad (3)$$

It is customary to have $\hat{V}' = \lambda \hat{V}$, where $0 \leq |\lambda| \leq 1$ is a parameter that can be manipulated in order to see the effect of the perturbation.

A nice example on how this works is given by a two state system, defined as follows

$$\begin{aligned} \hat{\mathcal{H}}_0 &= E_1^{(0)} |1_0\rangle \langle 1_0| + E_2^{(0)} |2_0\rangle \langle 2_0| \\ \hat{V} &= \lambda V_{12} |1_0\rangle \langle 2_0| + \lambda V_{21} |2_0\rangle \langle 1_0| \end{aligned} \quad (4)$$

In matrix representation, our Hamiltonian will then be the following

$$\mathcal{H}_{ij} = \begin{pmatrix} E_1^{(0)} & \lambda V_{12} \\ \lambda V_{21} & E_2^{(0)} \end{pmatrix} \quad (5)$$

Since it must be (obviously) an Hermitian operator, we have $V_{12} = V_{21}$. Calculating the eigenvalues is simple, and we get

$$E_{1,2} = \frac{E_1^{(0)} + E_2^{(0)}}{2} \pm \sqrt{\frac{(E_1^{(0)} - E_2^{(0)})^2}{4} + \lambda^2 V_{12}^2} \quad (6)$$

Since we are considering perturbations tied to a parameter, if $\lambda |V_{12}| \ll |E_1^{(0)} - E_2^{(0)}|$ we can approximate the square root with a power series, obtaining the following result

$$\begin{aligned} E_1 &= E_1^{(0)} + \frac{\lambda^2 |V_{12}|^2}{E_1^{(0)} - E_2^{(0)}} \\ E_2 &= E_2^{(0)} + \frac{\lambda^2 |V_{12}|^2}{E_2^{(0)} - E_1^{(0)}} \end{aligned} \quad (7)$$

Which are the perturbation-corrected eigenvalues.

Now it's simpler to grasp the *formal* development of the theory. We take our two secular equations (2) and (3), and rename the difference of eigenvalues found in the approximation (7) as follows $\Delta_n = E_n - E_n^{(0)}$.

Our new approximate Schrödinger equation can then be written as follows

$$\left(E_n^{(0)} - \hat{\mathcal{H}}_0\right) |n\rangle = \left(\lambda \hat{V} - \Delta_n\right) |n\rangle \quad (8)$$

And now here a little precaution: $E_n^{(0)} - \hat{\mathcal{H}}_0$ and $\lambda \hat{V} - \Delta_n$ are operators, and must be treated as such.

Back to our perturbation theory, we see right away that inverting the operator on the left isn't the way to go. It may act both on $|n_0\rangle, |n\rangle$, and therefore its inverse is ill-defined, but the right hand side comes to our rescue, and we can impose the following condition as an Ansatz

$$\langle n_0 | \left(\lambda \hat{V} - \Delta_n\right) |n\rangle = 0 \quad (9)$$

Now, we want to define properly the inverse of the operator on the left hand side, and we start by using a *complementary projection operator* $\hat{\phi}_n$ defined as follows

$$\hat{\phi}_n = 1 - |n_0\rangle \langle n_0| = \sum_{k \neq n} |k_0\rangle \langle k_0| \quad (10)$$

Now, in order to well-define the inverse operator we simply apply a projection beforehand

$$\frac{1}{E_n^{(0)} - \hat{\mathcal{H}}_0} \hat{\phi}_n = \sum_{k \neq n} \frac{1}{E_n^{(0)} - E_k^{(0)}} |k_0\rangle \langle k_0| \quad (11)$$

From the Ansatz (9) we have, evidently

$$\left(\lambda \hat{V} - \Delta_n\right) |n\rangle = \hat{\phi}_n \left(\lambda \hat{V} - \Delta_n\right) |n\rangle$$

So everything looks set up and fine, and it's tempting to find the perturbed eigenstates simply by inverting the first operator, but it simply doesn't work. Why? First of all, for $\lambda \rightarrow 0$ we *must* have $|n\rangle \rightarrow |n_0\rangle$, and Δ_0 , secondly because we need to add the solution to the homogeneous equation, hence, finally, we get the following result, naming this solution $c_n |n\rangle$

$$|n\rangle = c_n(\lambda) |n_0\rangle + \frac{1}{E_n^{(0)} - \hat{\mathcal{H}}_0} \hat{\phi}_n \left(\lambda \hat{V} - \Delta_n\right) |n\rangle \quad (12)$$

Where we have $c_n(\lambda) \rightarrow 1$ for $\lambda \rightarrow 0$, and $c_n(\lambda) = \langle n_0 | n \rangle$.

Simplifying the successive equations, we put $c_n(\lambda) = \langle n_0 | n \rangle = 1$ as our normalization condition, effectively removing a common multiplicative factor that appears. Then, easing the notation, we get

$$|n\rangle = |n_0\rangle + \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \left(\lambda \hat{V} - \Delta_n\right) |n\rangle \quad (13)$$

We also note that, from (9) that

$$\Delta_n = \lambda \langle n_0 | \hat{V} | n \rangle \quad (14)$$

Now everything is set. What we are searching depends only on the equations (13) and (14), and using the “smallness” of λ we approximate everything using power series, hence

$$\begin{aligned} |n\rangle &= \sum_{k=0}^{\infty} \lambda^k |n_k\rangle \\ \Delta_n &= \sum_{k=0}^{\infty} \lambda^k \Delta_n^{(k)} \end{aligned} \quad (15)$$

So, in order to evaluate the energy shift up to an order $\mathcal{O}(\lambda^N)$ it's sufficient to equate the coefficients of the powers of λ , putting simply the following condition $\Delta_n^{(N)} = \langle n_0 | \hat{V} | n_{N-1} \rangle$. It's evident how we need to know $|n_k\rangle$ only up to $\mathcal{O}(\lambda^{N-1})$. Adding all this in (15), we get

$$|n_0\rangle + \lambda |n_1\rangle + \dots = |n_0\rangle + \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \left(\lambda \hat{V} - \lambda \Delta_n^{(1)} - \dots \right) (|n_0\rangle + \lambda |n_1\rangle + \dots)$$

Therefore, for $\mathcal{O}(\lambda)$ we will get the following (remembering that $\hat{\phi}_n \Delta_n^{(1)} |n_0\rangle = 0$)

$$|n_1\rangle = \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \hat{V} |n_0\rangle \quad (16)$$

For $\mathcal{O}(\lambda^2)$ it gets trickier. Firstly we use the definition of $\Delta_n^{(2)}$, where

$$\Delta_n^{(2)} = \langle n_0 | \hat{V} \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \hat{V} | n_0 \rangle$$

Plugging it into the power series approximation up to order 2, we get therefore

$$\begin{aligned} |n_2\rangle &= \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \hat{V} \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} |n_0\rangle - \\ &\quad - \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \langle n_0 | \hat{V} | n_0 \rangle \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \hat{V} | n_0 \rangle \end{aligned} \quad (17)$$

Defining $\hat{\Phi} = \hat{\phi}/(E_n^{(0)} - \hat{\mathcal{H}}_0)$, we get the previous equations compacted

$$\begin{aligned} |n_1\rangle &= \hat{\Phi} \hat{V} |n_0\rangle \\ |n_2\rangle &= \hat{\Phi} \hat{V} \hat{\Phi} \hat{V} |n_0\rangle - \hat{\Phi} \langle \hat{V} \rangle_0 \hat{\Phi} \hat{V} |n_0\rangle \end{aligned} \quad (18)$$

It's evident that there is a trend in how next-order perturbations can be found, in this not-so-simple pattern.

Written explicitly, it's evident how this works

$$\begin{aligned} |n\rangle = & |n_0\rangle + \lambda \sum_{k \neq n} \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}} |k_0\rangle + \lambda^2 \sum_{k \neq n} \sum_{l \neq n} \frac{V_{kl} V_{ln}}{(E_n^{(0)} - E_k^{(0)})(E_n^{(0)} - E_l^{(0)})} |k_0\rangle - \\ & - \lambda^2 \sum_{k \neq n} \frac{V_{nn} V_{kn}}{(E_n^{(0)} - E_k^{(0)})^2} |k_0\rangle + \dots \end{aligned}$$

§§ 0.1.2 Rayleigh-Schrödinger Perturbation Theory, Degenerate Case

What we have defined so far, fails when the eigenstates we perturb are degenerate, since we supposed that there was only one well-defined eigenvalue $E_n^{(0)}$ for each eigenket.

Let's now suppose that we have a system, for which there is a g -fold degeneracy, hence there are g unperturbed eigenkets $|m_0\rangle$ for one single $E_D^{(0)}$ eigenvalue. Let's define the degenerate eigenspace as $D := \{|m_0\rangle \in \mathbb{H} \mid \hat{\mathcal{H}} |m_0\rangle = E_D^{(0)} |m_0\rangle, m = 1, \dots, g\}$. In general, the perturbation breaks the degeneracy, forming a new set of eigenkets $|l\rangle$ that do not coincide with the unperturbed set $|l_0\rangle$, although we can use the following projection

$$|l_0\rangle = \sum_{|m\rangle \in D} \langle m_0 | l_0 \rangle |m_0\rangle$$

Let's rewrite the Schrödinger equation for the new states $|l\rangle$, and define the projections $\hat{\pi}_0 = |m_0\rangle \langle m_0|$ and its coprojection $\hat{\pi}_1 = \hat{\mathbb{1}} - \hat{\pi}_0$. The Schrödinger equation then becomes

$$(E - \hat{\mathcal{H}}_0 - \lambda \hat{V}) |l\rangle = (E - E_D^{(0)} - \lambda \hat{V}) (\hat{\pi}_0 |l\rangle + \hat{\pi}_1 |l\rangle) = 0 \quad (19)$$

We separate the equation (19) multiplying on the left firstly by $\hat{\pi}_0$ and then by $\hat{\pi}_1$

$$\begin{aligned} (E - E_D^{(0)} - \lambda \hat{\pi}_0 \hat{V}) \hat{\pi}_0 |l\rangle - \lambda \hat{\pi}_0 \hat{V} \hat{\pi}_1 |l\rangle &= 0 \\ (E - \hat{\mathcal{H}}_0 - \lambda \hat{\pi}_1 \hat{V}) \hat{\pi}_1 |l\rangle - \lambda \hat{\pi}_1 \hat{V} \hat{\pi}_0 |l\rangle &= 0 \end{aligned} \quad (20)$$

From this equation we can then solve for $\hat{\pi}_1 |l\rangle$ and $\hat{\pi}_0 |l\rangle$

$$\begin{aligned} \hat{\pi}_1 &= \hat{\pi}_1 \frac{\lambda}{E - \hat{\mathcal{H}}_0 - \lambda \hat{\pi}_1 \hat{V} \hat{\pi}_1} \hat{\pi}_1 \hat{V} \hat{\pi}_0 |l\rangle \\ \left(E - E_D^{(0)} - \lambda \hat{\pi}_0 \hat{V} \hat{\pi}_0 - \lambda^2 \hat{\pi}_0 \hat{V} \hat{\pi}_1 \frac{1}{E - \hat{\mathcal{H}}_0 \lambda \hat{V}} \hat{\pi}_1 \hat{V} \hat{\pi}_0 \right) \hat{\pi}_0 |l\rangle &= 0 \end{aligned} \quad (21)$$

The general approximation to $\mathcal{O}(\lambda^n)$ will be given from the following general expression

$$\hat{\pi}_1 |l_1\rangle = \sum_{|k\rangle \notin D} \frac{V_{kl}}{E_D^{(0)} - E_k^{(0)}} |k_0\rangle$$

And, henceforth, in order to solve for $\mathcal{O}(\lambda)$, we get the following equation

$$(E - E_D^{(0)} - \lambda \hat{\pi}_0 \hat{V} \hat{\pi}_0) \hat{\pi}_0 |l_0\rangle = 0 \quad (22)$$

The energy shifts $\Delta^{(1)}$ will then be the diagonal elements of the perturbation $\langle l_0 | \hat{V} | l_0 \rangle$. We can immediately ask why a λ^2 appears in (21). This is given simply by the substitution we made in order to get the equation, but we already know that the energy shift at the first order is $E_i^{(1)} = E_D^{(0)} + \lambda v_i$, where v_i are the eigenvalues of the operator $\hat{\pi}_0 \hat{V} \hat{\pi}_0$. We assume that the degeneracy is completely resolved after the application of the perturbation, hence we get $E_i^{(1)} = \lambda(v_i - v_j) \neq 0$. Since there isn't anymore degeneration in this system, we solve using nondegenerate Rayleigh-Schrödinger perturbation theory, obtaining the corrections

$$\hat{\pi}_0 |l_i^1\rangle = \lambda \sum_{j \neq i} \frac{\hat{\pi}_0}{v_j - v_i} |l_j^0\rangle \langle l_j^0 | \hat{V} \hat{\pi}_1 \frac{1}{E_D^{(0)} - \hat{\mathcal{H}}_0} \hat{\pi}_1 \hat{V} |l_i^0\rangle \quad (23)$$

Since $\hat{\pi}_0 |l_j^0\rangle$ are eigenvectors of \hat{V} we get that the energy shift at the second order is simply

$$\Delta_l^{(2)} = \sum_{k \notin D} \frac{|V_{kl}|^2}{E_D^{(0)} - E_k^{(0)}} \quad (24)$$

§ 0.2 Variational Methods

Approximating through variational methods is done when searching for approximate ground states energies E_0 when exact values aren't available.

We start to “guess” the ground state by defining a trial ket $|\phi\rangle$. We then define the following

$$\langle \hat{\mathcal{H}} \rangle = \frac{\langle \phi | \hat{\mathcal{H}} | \phi \rangle}{\langle \phi | \phi \rangle} \quad (25)$$

Theorem 1. *There exists an upper bound to E_0 , hence*

$$\langle \hat{\mathcal{H}} \rangle \geq E_0$$

Proof. We can expand $|\phi\rangle$ as follows

$$|\phi\rangle = \sum_{k=0}^{\infty} |k\rangle \langle k | \phi \rangle$$

Where, $\hat{\mathcal{H}} |k\rangle = E_k |k\rangle$, hence it's an exact eigenket.

We can write $E_k = E_k - E_0 + E_0$ and evaluating $\langle \hat{\mathcal{H}} \rangle$ we have

$$\langle \hat{\mathcal{H}} \rangle = \frac{\sum_{k=0}^{\infty} |\langle k | \phi \rangle|^2 E_k}{\sum_{k=0}^{\infty} |\langle k | \phi \rangle|^2} = \frac{\sum_{k=0}^{\infty} |\langle k | \phi \rangle|^2 (E_k - E_0)}{\sum_{k=0}^{\infty} |\langle k | \phi \rangle|^2} + E_0 \geq E_0$$

Obviously, the equality is given iff $|\phi\rangle$ is the exact ground eigenket □

This method is really powerful, since for even a poor trial ket we have $\langle k|\emptyset\rangle \sim \mathcal{O}(\epsilon)$ and $\langle \hat{\mathcal{H}} \rangle - E_0 \sim \mathcal{O}(\epsilon^2)$.

Another way to say this is saying that if we variate $|\emptyset\rangle$, the Hamiltonian will be stationary with respect to $\delta|\emptyset\rangle$.

This method doesn't say what shape does the ket $|\emptyset\rangle$ has, hence we must guess them, using the system as a guide.

Practically, it's much more useful to define a parameter vector λ_i which will appear in the considered eigenket, and then find the minimum of $\langle \hat{\mathcal{H}} \rangle$, imposing the following equation

$$\frac{\partial \langle \hat{\mathcal{H}} \rangle}{\partial \lambda_i} = 0 \quad (26)$$

§ 0.3 Time Dependent Perturbation Theory

§§ 0.3.1 Dirac Interaction Picture

Let's begin considering a time dependent Hamiltonian that can be split in two parts

$$\hat{\mathcal{H}}(t) = \hat{\mathcal{H}}_0 + \hat{V}(t) \quad (27)$$

Where one piece is time independent and the other is time dependent. We suppose that $\hat{\mathcal{H}}_0$ is exactly solvable.

Let's suppose that at $t = 0$ the state ket is given by the following relation

$$|\alpha\rangle = \sum_n c_n(0) |n\rangle \quad (28)$$

Where $|n\rangle$ is the eigenvalue of $\hat{\mathcal{H}}_0$.

Our objective is to find some $c_n(t)$ such that

$$|\alpha(t)\rangle = \sum_n c_n(t) e^{-\frac{iE_n t}{\hbar}} |n\rangle \quad (29)$$

Now, in order to simplify our problem, we define the *Dirac picture*, or *Interaction picture*, where, having considered our Hamiltonian, we have

$$|\alpha(t)\rangle_I = e^{\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} |\alpha(t)\rangle_S \quad (30)$$

The observables in this picture will be defined as follows, in particular, for our perturbation \hat{V} , we have

$$\hat{V}_I = e^{\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \hat{V} e^{-\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \quad (31)$$

Remembering the relation between Schrödinger and Heisenberg picture, we have that

$$\begin{aligned} |\alpha\rangle_H &= e^{\frac{i\hat{\mathcal{H}} t}{\hbar}} |\alpha(t_0)\rangle_S \\ |\alpha(t_0)\rangle_S &= e^{-\frac{i\hat{\mathcal{H}}(t-t_0)}{\hbar}} |\alpha\rangle \\ \hat{A}_H &= e^{\frac{i\hat{\mathcal{H}} t}{\hbar}} \hat{A} e^{-\frac{i\hat{\mathcal{H}} t}{\hbar}} \end{aligned} \quad (32)$$

Hence, we have that the Dirac interaction picture will satisfy the following time dependent Schrödinger equation (remembering that $i\hbar\partial_t |\alpha(t)\rangle_I = i\hbar\partial_t (\exp(i\hat{\mathcal{H}}_0 t/\hbar) |\alpha(t)\rangle_S)$)

$$i\hbar \frac{\partial}{\partial t} |\alpha(t)\rangle_I = \hat{V}_I |\alpha(t)\rangle_I \quad (33)$$

Where the perturbation takes the place of the Hamiltonian in the time dependent equation.

It is also demonstrable that, for an observable \hat{A} , its interaction picture will satisfy the following differential equation

$$\frac{d\hat{A}_I}{dt} = \frac{1}{i\hbar} [\hat{A}_I, \hat{\mathcal{H}}_0] \quad (34)$$

It's obvious that this Dirac Interaction picture is the halfway between a Schrödinger and a Heisenberg representation picture.

Going back to (29), we have that in interaction picture it will simply become this

$$|\alpha(t)\rangle_I = \sum_n c_n(t) |n\rangle \quad (35)$$

We can now write a differential equation for $c_n(t)$.

$$i\hbar \frac{\partial}{\partial t} \langle n|\alpha(t)\rangle_I = \sum_m \langle n|\hat{V}_I|m\rangle \langle m|\alpha(t)\rangle_I \quad (36)$$

By definition, we have $c_n(t) = \langle n|\alpha(t)\rangle_I$, henceforth

$$i\hbar \frac{dc_n}{dt} = \sum_m V_{nm} e^{i\omega_{nm}t} c_m(t) \quad (37)$$

Where we have expanded the interaction-picture time dependence, and we have by definition of frequency

$$\omega_{nm} = \frac{E_n - E_m}{\hbar} = -\omega_{mn} \quad (38)$$

§§ 0.3.2 Dyson Series

Usually, exact solutions for $c_n(t)$ are not available, hence we must find a suitable approximation for our solution. One good way to start is to suppose that $c_n(t)$ can be expressed via the sum of different functions as follows

$$c_n(t) = \sum_{i=0}^{\infty} c_n^{(i)}$$

Where $c_n^{(i)}$ indicates the i -th transition amplitude.

This problem can be attacked, using that $c_n^{(0)} = \delta_{in}$ and then using it to define a differential equation for $c_n^{(1)}(t)$ and so on. Using operator theory, this problem can be

solved even in a better way.

The time evolution operator in the Dirac picture is defined as follows

$$|\alpha(t)\rangle = \hat{\mathcal{U}}_I(t) |\alpha\rangle$$

For which, we know already that it's solution to the following ODE

$$\begin{cases} i\hbar \frac{d\hat{\mathcal{U}}_I}{dt} = \hat{V}_I \hat{\mathcal{U}}_I \\ \hat{\mathcal{U}}_I(t_0) = \hat{\mathbb{1}} \end{cases}$$

This differential equation is equivalent to the following integral equation

$$\hat{\mathcal{U}}_I(t) = \hat{\mathbb{1}} - \frac{i}{\hbar} \int_{t_0}^t \hat{V}_I(t_1) \hat{\mathcal{U}}_I(t_1) dt_1$$

Iterating, we get

$$\hat{\mathcal{U}}_I(t) = \sum_{n=0}^{\infty} \prod_{k=1}^{\infty} \left(\frac{-i\hat{\mathbb{1}}}{\hbar} \right)^n \int_{t_0}^{t_{k-1}} V_I(t_k) dt_k \quad (39)$$

Which is equivalent to writing the following expression

$$\begin{aligned} \hat{\mathcal{U}}_I(t) &\approx \hat{\mathbb{1}} - \frac{i}{\hbar} \int_{t_0}^t V_I(t_1) dt_1 \approx \\ &\hat{\mathbb{1}} - \left(\frac{-i}{\hbar} \right)^2 \int_{t_0}^t V_I(t_1) V_I(t_2) dt_1 dt_2 + \dots \\ &\dots + \left(\frac{-i}{\hbar} \right)^n \int_{t_0}^t \int_{t_0}^{t_1} \dots \int_{t_0}^{t_{n-1}} V_I(t_1) \dots V_I(t_n) dt_1 \dots dt_n \end{aligned}$$

Through this approximation, it's virtually possible to compute $\hat{\mathcal{U}}_I(t)$ to any order, and therefore $c_n(t)$. This kind of computation is fundamental in fields like atomic physics, as we will see later.