

# Perturbation theory (quantum mechanics)

In quantum mechanics, **perturbation theory** is a set of approximation schemes directly related to mathematical perturbation for describing a complicated quantum system in terms of a simpler one. The idea is to start with a simple system for which a mathematical solution is known, and add an additional "perturbing" Hamiltonian representing a weak disturbance to the system. If the disturbance is not too large, the various physical quantities associated with the perturbed system (e.g. its energy levels and eigenstates) can, from considerations of continuity, be expressed as 'corrections' to those of the simple system. These corrections, being 'small' compared to the size of the quantities themselves, can be calculated using approximate methods such as asymptotic series. The complicated system can therefore be studied based on knowledge of the simpler one.

## *Applications of perturbation theory*

Perturbation theory is an important tool for describing real quantum systems, as it turns out to be very difficult to find exact solutions to the Schrödinger equation for Hamiltonians of even moderate complexity. The Hamiltonians to which we know exact solutions, such as the hydrogen atom, the quantum harmonic oscillator and the particle in a box, are too idealized to adequately describe most systems. Using perturbation theory, we can use the known solutions of these simple Hamiltonians to generate solutions for a range of more complicated systems. For example, by adding a perturbative electric potential to the quantum mechanical model of the hydrogen atom, we can calculate the tiny shifts in the spectral lines of hydrogen caused by the presence of an electric field (the Stark effect). This is only approximate because the sum of a Coulomb potential with a linear potential is unstable although the tunneling time (decay rate) is very long. This shows up as a broadening of the energy spectrum lines, something which perturbation theory fails to reproduce entirely.

The expressions produced by perturbation theory are not exact, but they can lead to accurate results as long as the expansion parameter, say  $\alpha$ , is very small. Typically, the results are expressed in terms of finite power series in  $\alpha$  that seem to converge to the exact values when summed to higher order. After a certain order  $n \sim 1/\alpha$ , however, the results become increasingly worse since the series are usually divergent (being asymptotic series). There exist ways to convert them into convergent series, which can be evaluated for large-expansion parameters, most efficiently by Variational method.

In the theory of quantum electrodynamics (QED), in which the electron-photon interaction is treated perturbatively, the calculation of the electron's magnetic moment has been found to agree with experiment to eleven decimal places. In QED and other quantum field theories, special calculation techniques known as Feynman diagrams are used to systematically sum the power series terms.

Under some circumstances, perturbation theory is an invalid approach to take. This happens when the system we wish to describe cannot be described by a small perturbation imposed on some simple system. In quantum chromodynamics, for instance, the interaction of quarks with the gluon field cannot be treated perturbatively at low energies because the coupling constant (the expansion parameter) becomes too large. Perturbation theory also fails to describe states that are not generated adiabatically from the "free model", including bound states and various collective phenomena such as solitons. Imagine, for example, that we have a system of free (i.e. non-interacting) particles, to which an attractive interaction is introduced. Depending on the form of the interaction, this may create an entirely new set of eigenstates corresponding to groups of particles bound to one another. An example of this phenomenon may be found in conventional superconductivity, in which the phonon-mediated attraction between conduction electrons leads to the formation of correlated electron pairs known as Cooper pairs. When faced with such systems, one usually turns to other approximation schemes, such as the variational method and the WKB approximation. This is because there is no analogue of a bound particle in the unperturbed model and the energy of a soliton typically goes as the *inverse* of the expansion parameter. However, if we "integrate" over the solitonic phenomena, the nonperturbative corrections in this case will be tiny; of the order of  $e^{-1/g}$  or  $e^{-1/g^2}$  in the perturbation parameter  $g$ . Perturbation theory can only detect solutions "close" to the unperturbed solution, even if there are other solutions (which typically blow up as the expansion parameter goes to zero).

The problem of non-perturbative systems has been somewhat alleviated by the advent of modern computers. It has become practical to obtain numerical non-perturbative solutions for certain problems, using methods such as density functional theory. These advances have been of particular benefit to the field of quantum chemistry. Computers have also been used to carry out perturbation theory calculations to extraordinarily high levels of precision, which has proven important in particle physics for generating theoretical results that can be compared with experiment.

### ***Time-independent perturbation theory***

Time-independent perturbation theory is one of two categories of perturbation theory, the other being time-dependent perturbation (see next section). In time-independent perturbation theory the perturbation Hamiltonian is static (i.e., possesses no time dependence). Time-independent perturbation theory was presented by Erwin Schrödinger in a 1926 paper, shortly after he produced his theories in wave mechanics. In this paper Schrödinger referred to earlier work of Lord Rayleigh, who investigated harmonic vibrations of a string perturbed by small inhomogeneities. This is why this perturbation theory is often referred to as **Rayleigh-Schrödinger perturbation theory**.

## First order corrections

We begin with an unperturbed Hamiltonian  $H_0$ , which is also assumed to have no time dependence. It has known energy levels and eigenstates, arising from the time-independent Schrödinger equation:

$$H_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle \quad , \quad n = 1, 2, 3, \dots$$

For simplicity, we have assumed that the energies are discrete. The  $(0)$  superscripts denote that these quantities are associated with the unperturbed system. Note the use of Bra-ket notation.

We now introduce a perturbation to the Hamiltonian. Let  $V$  be a Hamiltonian representing a weak physical disturbance, such as a potential energy produced by an external field. (Thus,  $V$  is formally a Hermitian operator.) Let  $\lambda$  be a dimensionless parameter that can take on values ranging continuously from 0 (no perturbation) to 1 (the full perturbation). The perturbed Hamiltonian is

$$H = H_0 + \lambda V$$

The energy levels and eigenstates of the perturbed Hamiltonian are again given by the Schrödinger equation:

$$(H_0 + \lambda V)|n\rangle = E_n|n\rangle.$$

Our goal is to express  $E_n$  and  $|n\rangle$  in terms of the energy levels and eigenstates of the old Hamiltonian. If the perturbation is sufficiently weak, we can write them as power series in  $\lambda$ :

$$\begin{aligned} E_n &= E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \\ |n\rangle &= |n^{(0)}\rangle + \lambda |n^{(1)}\rangle + \lambda^2 |n^{(2)}\rangle + \dots \end{aligned}$$

where

$$E_n^{(k)} = \frac{1}{k!} \frac{d^k E_n}{d\lambda^k}$$

and

$$|n^{(k)}\rangle = \frac{1}{k!} \frac{d^k |n\rangle}{d\lambda^k}.$$

When  $\lambda = 0$ , these reduce to the unperturbed values, which are the first term in each series. Since the perturbation is weak, the energy levels and eigenstates should not deviate too much from their unperturbed values, and the terms should rapidly become smaller as we go to higher order.

Plugging the power series into the Schrödinger equation, we obtain

$$\begin{aligned} (H_0 + \lambda V) (|n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \dots) \\ = \left( E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \right) (|n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \dots) \end{aligned}$$

Expanding this equation and comparing coefficients of each power of  $\lambda$  results in an infinite series of simultaneous equations. The zeroth-order equation is simply the Schrödinger equation for the unperturbed system. The first-order equation is

$$H_0|n^{(1)}\rangle + V|n^{(0)}\rangle = E_n^{(0)}|n^{(1)}\rangle + E_n^{(1)}|n^{(0)}\rangle$$

Operating through by  $\langle n^{(0)}|$ . The first term on the left-hand side cancels with the first term on the right-hand side. (Recall, the unperturbed Hamiltonian is hermitian). This leads to the first-order energy shift:

$$E_n^{(1)} = \langle n^{(0)}|V|n^{(0)}\rangle$$

This is simply the expectation value of the perturbation Hamiltonian while the system is in the unperturbed state. This result can be interpreted in the following way: suppose the perturbation is applied, but we keep the system in the quantum state  $|n^{(0)}\rangle$ , which is a valid quantum state though no longer an energy eigenstate. The perturbation causes the average energy of this state to increase by  $\langle n^{(0)}|V|n^{(0)}\rangle$ . However, the true energy shift is slightly different, because the perturbed eigenstate is not exactly the same as  $|n^{(0)}\rangle$ . These further shifts are given by the second and higher order corrections to the energy.

Before we compute the corrections to the energy eigenstate, we need to address the issue of normalization. We may suppose  $\langle n^{(0)}|n^{(0)}\rangle=1$ , but perturbation theory assumes we also have  $\langle n|n\rangle=1$ . It follows that at first order in  $\lambda$ , we must have  $\langle n^{(0)}|n^{(1)}\rangle + \langle n^{(1)}|n^{(0)}\rangle=0$ . Since the overall phase is not determined in quantum mechanics, without loss of generality, we may assume  $\langle n^{(0)}|n^{(1)}\rangle$  is purely real. Therefore,  $\langle n^{(0)}|n^{(1)}\rangle = -\langle n^{(1)}|n^{(0)}\rangle$ , and we deduce

$$\langle n^{(0)}|n^{(1)}\rangle = 0.$$

To obtain the first-order correction to the energy eigenstate, we insert our expression for the first-order energy correction back into the result shown above of equating the first-order coefficients of  $\lambda$ . We then make use of the resolution of the identity,

$$\begin{aligned} V|n^{(0)}\rangle &= \left( \sum_{k \neq n} |k^{(0)}\rangle \langle k^{(0)}| \right) V|n^{(0)}\rangle + \left( |n^{(0)}\rangle \langle n^{(0)}| \right) V|n^{(0)}\rangle \\ &= \sum_{k \neq n} |k^{(0)}\rangle \langle k^{(0)}|V|n^{(0)}\rangle + E_n^{(1)}|n^{(0)}\rangle, \end{aligned}$$

where  $|k^{(0)}\rangle$  is in the orthogonal complement of  $|n^{(0)}\rangle$ . The result is

$$(E_n^{(0)} - H_0) |n^{(1)}\rangle = \sum_{k \neq n} |k^{(0)}\rangle \langle k^{(0)}|V|n^{(0)}\rangle$$

For the moment, suppose that the zeroth-order energy level is not degenerate, i.e. there is no eigenstate of  $H_0$  in the orthogonal complement of  $|n^{(0)}\rangle$  with the energy  $E_n^{(0)}$ . We multiply through by  $\langle k^{(0)}|$ , which gives

$$(E_n^{(0)} - E_k^{(0)}) \langle k^{(0)}|n^{(1)}\rangle = \langle k^{(0)}|V|n^{(0)}\rangle$$

and hence the component of the first-order correction along  $|k^{(0)}\rangle$  since by assumption  $E_n^{(0)} \neq E_k^{(0)}$ . In total we get

$$|n^{(1)}\rangle = \sum_{k \neq n} \frac{\langle k^{(0)}|V|n^{(0)}\rangle}{E_n^{(0)} - E_k^{(0)}} |k^{(0)}\rangle$$

The first-order change in the  $n$ -th energy eigenket has a contribution from each of the energy eigenstates  $k \neq n$ . Each term is proportional to the matrix element  $\langle k^{(0)}|V|n^{(0)}\rangle$ , which is a measure of how much the perturbation mixes eigenstate  $n$  with eigenstate  $k$ ; it is also inversely proportional to the energy difference between eigenstates  $k$  and  $n$ , which means that the perturbation deforms the eigenstate to a greater extent if there are more eigenstates at nearby energies. We see also that the expression is singular if any of these states have the same energy as state  $n$ , which is why we assumed that there is no degeneracy.

## Second-order and higher corrections

We can find the higher-order deviations by a similar procedure, though the calculations become quite tedious with our current formulation. Our normalization prescription gives that  $2\langle n^{(0)}|n^{(2)}\rangle + \langle n^{(1)}|n^{(1)}\rangle = 0$ . Up to second order, the expressions for the energies and (normalized) eigenstates are:

$$\begin{aligned} E_n &= E_n^{(0)} + \lambda \langle n^{(0)}|V|n^{(0)}\rangle + \lambda^2 \sum_{k \neq n} \frac{|\langle k^{(0)}|V|n^{(0)}\rangle|^2}{E_n^{(0)} - E_k^{(0)}} + O(\lambda^3) \\ |n\rangle &= |n^{(0)}\rangle + \lambda \sum_{k \neq n} |k^{(0)}\rangle \frac{\langle k^{(0)}|V|n^{(0)}\rangle}{E_n^{(0)} - E_k^{(0)}} + \lambda^2 \sum_{k \neq n} \sum_{\ell \neq n} |k^{(0)}\rangle \frac{\langle k^{(0)}|V|\ell^{(0)}\rangle \langle \ell^{(0)}|V|n^{(0)}\rangle}{(E_n^{(0)} - E_k^{(0)})(E_n^{(0)} - E_\ell^{(0)})} \\ &\quad - \lambda^2 \sum_{k \neq n} |k^{(0)}\rangle \frac{\langle n^{(0)}|V|n^{(0)}\rangle \langle k^{(0)}|V|n^{(0)}\rangle}{(E_n^{(0)} - E_k^{(0)})^2} - \frac{1}{2} \lambda^2 |n^{(0)}\rangle \sum_{k \neq n} \frac{\langle n^{(0)}|V|k^{(0)}\rangle \langle k^{(0)}|V|n^{(0)}\rangle}{(E_n^{(0)} - E_k^{(0)})^2} + O(\lambda^3). \end{aligned}$$

Extending the process further, the third-order energy correction can be shown to be

$$E_n^{(3)} = \sum_{k \neq n} \sum_{m \neq n} \frac{\langle n^{(0)} | V | m^{(0)} \rangle \langle m^{(0)} | V | k^{(0)} \rangle \langle k^{(0)} | V | n^{(0)} \rangle}{(E_m^{(0)} - E_n^{(0)}) (E_k^{(0)} - E_n^{(0)})} - \langle n^{(0)} | V | n^{(0)} \rangle \sum_{m \neq n} \frac{|\langle n^{(0)} | V | m^{(0)} \rangle|^2}{(E_m^{(0)} - E_n^{(0)})^2}.$$

## Effects of degeneracy

Suppose that two or more energy eigenstates are degenerate. The first-order energy shift is not well defined, since there is no unique way to choose a basis of eigenstates for the unperturbed system. The calculation of the change in the eigenstate is problematic as well, because the operator

$$E_n^{(0)} - H_0$$

does not have a well-defined inverse.

Let  $D$  denote the subspace spanned by these degenerate eigenstates. No matter how small the perturbation is, in the degenerate subspace  $D$  the energy differences between the eigenstates  $H_0$  are zero, so complete mixing of at least some of these states is assured. Thus the perturbation can not be considered small in the  $D$  subspace and in that subspace the new Hamiltonian must be diagonalized first. These correct perturbed eigenstates in  $D$  are now the basis for the perturbation expansion:

$$|n\rangle = \sum_{k \in D} \alpha_{nk} |k^{(0)}\rangle + \lambda |n^{(1)}\rangle$$

where only eigenstates outside of the  $D$  subspace are considered to be small. For the first-order perturbation we need to solve the perturbed Hamiltonian restricted to the degenerate subspace  $D$

$$V |k^{(0)}\rangle = \epsilon_k |k^{(0)}\rangle \quad \forall |k^{(0)}\rangle \in D.$$

simultaneously for all the degenerate eigenstates, where  $\epsilon_k$  are first-order corrections to the degenerate energy levels. This is equivalent to diagonalizing the matrix

$$\langle k^{(0)} | V | l^{(0)} \rangle = V_{kl} \quad \forall |k^{(0)}\rangle, |l^{(0)}\rangle \in D.$$

This procedure is approximate, since we neglected states outside the  $D$  subspace. The splitting of degenerate energies  $\epsilon_k$  is generally observed. Although the splitting may be small compared to the range of energies found in the system, it is crucial in understanding certain details, such as spectral lines in Electron Spin Resonance experiments.

Higher-order corrections due to other eigenstates can be found in the same way as for the non-degenerate case

$$(E_n^{(0)} - H_0) |n^{(1)}\rangle = \sum_{k \notin D} (\langle k^{(0)} | V | n^{(0)} \rangle) |k^{(0)}\rangle.$$

The operator on the left hand side is not singular when applied to eigenstates outside  $D$ , so we can write

$$|n^{(1)}\rangle = \sum_{k \notin D} \frac{\langle k^{(0)} | V | n^{(0)} \rangle}{E_n^{(0)} - E_k^{(0)}} |k^{(0)}\rangle,$$

but the effect on the degenerate states is minuscule, proportional to the square of the first-order correction  $\epsilon_k$ .

Near-degenerate states should also be treated in the above manner, since the original Hamiltonian won't be larger than the perturbation in the near-degenerate subspace. An application is found in the nearly-free electron model, where near-degeneracy treated properly gives rise to an energy gap even for small perturbations. Other eigenstates will only shift the absolute energy of all near-degenerate states simultaneously.

### Generalization to Multi-parameter Case

The generalization of the time-independent perturbation theory to the multi-parameter case can be formulated more systematically using the language of differential geometry, which basically defines the derivatives of the quantum states and calculate the perturbative corrections by taking derivatives iteratively at the unperturbed point.

### Hamiltonian and Force Operator

From the differential geometric point of view, a parameterized Hamiltonian is considered as a function defined on the parameter manifold that maps each particular set of parameters  $x^\mu = (x^1, x^2, \dots)$  to an Hermitian operator  $H(x^\mu)$  that acts on the Hilbert space. The parameters here can be external field, interaction strength, or driving parameters in the quantum phase transition. Let  $E_n(x^\mu)$  and  $|n(x^\mu)\rangle$  be the  $n^{\text{th}}$  eigenenergy and eigenstate of  $H(x^\mu)$  respectively. In the language of differential geometry, the states  $|n(x^\mu)\rangle$  form a vector bundle over the parameter manifold, on which derivatives of these states can be defined. The perturbation theory is to answer the following question: given  $E_n(x_0^\mu)$  and  $|n(x_0^\mu)\rangle$  at a unperturbed reference point  $x_0^\mu$ , how to estimate the  $E_n(x^\mu)$  and  $|n(x^\mu)\rangle$  at  $x^\mu$  close to that reference point.

Without loss of generality, the coordinate system can be shifted, such that the reference point  $x_0^\mu = 0$  is set to be the origin. The following linearly parameterized Hamiltonian is frequently used

$$H(x^\mu) = H(0) + x^\mu F_\mu.$$

If the parameters  $x^\mu$  are considered as generalized coordinates, then  $F_\mu$  should be identified as the generalized force operators related to those coordinates. Different indices  $\mu$ 's label the different forces along different directions in the parameter manifold. For example, if  $x^\mu$  denotes the external magnetic field in the  $\mu$ -direction, then  $F_\mu$  should be the magnetization in the same direction.

### **Perturbation Theory as Power Series Expansion**

The validity of the perturbation theory lies on the adiabatic assumption, which assumes the eigenenergies and eigenstates of the Hamiltonian are smooth functions of parameters such that their values in the vicinity region can be calculated in power series (like Taylor expansion) of the parameters:

$$\begin{aligned} E_n(x^\mu) &= E_n + x^\mu \partial_\mu E_n + \frac{1}{2!} x^\mu x^\nu \partial_\mu \partial_\nu E_n + \dots, \\ |n(x^\mu)\rangle &= |n\rangle + x^\mu |\partial_\mu n\rangle + \frac{1}{2!} x^\mu x^\nu |\partial_\mu \partial_\nu n\rangle + \dots \end{aligned}$$

Here  $\partial_\mu$  denotes the derivative with respect to  $x^\mu$ . When applying to the state  $|\partial_\mu n\rangle$ , it should be understood as the Lie derivative if the vector bundle is equipped with non-vanishing connection. All the terms on the right-hand-side of the series are evaluated at  $x^\mu = 0$ , e.g.  $E_n \equiv E_n(0)$  and  $|n\rangle \equiv |n(0)\rangle$ . This convention will be adopted though out the this subsection, that all functions without the parameter dependence explicitly stated are assumed to be evaluated at the origin. The power series may converge slowly or even not converging when the energy levels are close to each other. The adiabatic assumption breaks down when there is energy level degeneracy, and hence the perturbation theory is not applicable in that case.

### **Hellmann-Feynman Theorems**

The above power series expansion can be readily evaluated if there is a systematic approach to calculate the derivatives to any order. Using the chain rule, the derivatives can be broken down to the single derivative on either the energy or the state. The Hellmann-Feynman theorems are used to calculate these single derivatives. The first Hellmann-Feynman theorem gives the derivative of the energy,

$$\partial_\mu E_n = \langle n | \partial_\mu H | n \rangle.$$

The second Hellmann-Feynman theorem gives the derivative of the state (resolved by the complete basis with  $m \neq n$ ),



$$\langle m|\partial_\mu n\rangle = \frac{\langle m|\partial_\mu H|n\rangle}{E_n - E_m},$$

$$\langle \partial_\mu m|n\rangle = \frac{\langle m|\partial_\mu H|n\rangle}{E_m - E_n}.$$

For the linearly parameterized Hamiltonian,  $\partial_\mu H$  simply stands for the generalized force operator  $F_\mu$ .

The theorems can be simply derived by applying the differential operator  $\partial_\mu$  to both sides of the Schrödinger equation  $H|n\rangle = E_n|n\rangle$ , which reads

$$\partial_\mu H|n\rangle + H|\partial_\mu n\rangle = \partial_\mu E_n|n\rangle + E_n|\partial_\mu n\rangle.$$

Then overlap with the state  $\langle m|$  from left and make use of the Schrödinger equation again  $\langle m|H = \langle m|E_m$ ,

$$\langle m|\partial_\mu H|n\rangle + E_m\langle m|\partial_\mu n\rangle = \partial_\mu E_n\langle m|n\rangle + E_n\langle m|\partial_\mu n\rangle.$$

Given that the eigenstates of the Hamiltonian always form a set of orthonormal basis  $\langle m|n\rangle = \delta_{mn}$ , both the cases of  $m = n$  and  $m \neq n$  can be discussed separately. The first case will lead to the first theorem and the second case to the second theorem, which can be shown immediately by rearranging the terms. With the differential rules given by the Hellmann-Feynman theorems, the perturbative correction to the energies and states can be calculated systematically.

### Correction of Energy and State

To the second order, the energy correction reads

$$E_n(x^\mu) = \langle n|H|n\rangle + \langle n|\partial_\mu H|n\rangle x^\mu + \sum_{m \neq n} \frac{\langle n|\partial_\nu H|m\rangle \langle m|\partial_\mu H|n\rangle}{E_n - E_m} x^\mu x^\nu + \dots$$

The first order derivative  $\partial_\mu E_n$  is given by the first Hellmann-Feynman theorem directly. To obtain the second order derivative  $\partial_\mu \partial_\nu E_n$ , simply applying the differential operator  $\partial_\mu$  to the result of the first order derivative  $\langle n|\partial_\nu H|n\rangle$ , which reads

$$\partial_\mu \partial_\nu E_n = \langle \partial_\mu n|\partial_\nu H|n\rangle + \langle n|\partial_\mu \partial_\nu H|n\rangle + \langle n|\partial_\nu H|\partial_\mu n\rangle.$$

Note that for linearly parameterized Hamiltonian, there is no second derivative  $\partial_\mu \partial_\nu H = 0$  on the operator level. Resolve the derivative of state by inserting the complete set of basis,

$$\partial_\mu \partial_\nu E_n = \sum_m (\langle \partial_\mu n | m \rangle \langle m | \partial_\nu H | n \rangle + \langle n | \partial_\nu H | m \rangle \langle m | \partial_\mu n \rangle),$$

then all parts can be calculated using the Hellmann-Feynman theorems. In terms of Lie derivatives,  $\langle \partial_\mu n | n \rangle = \langle n | \partial_\mu n \rangle = 0$  according to the definition of the connection for the vector bundle. Therefore the case  $m = n$  can be excluded from the summation, which avoids the singularity of the energy denominator. The same procedure can be carried on for higher order derivatives, from which higher order corrections are obtained.

The same computational scheme is applicable for the correction of states. The result to the second order is as follows

$$\begin{aligned} |n(x^\mu)\rangle = & |n\rangle + \sum_{m \neq n} \frac{\langle m | \partial_\mu H | n \rangle}{E_n - E_m} |m\rangle x^\mu \\ & + \left( \sum_{m \neq n} \sum_{l \neq n} \frac{\langle m | \partial_\mu H | l \rangle \langle l | \partial_\nu H | n \rangle}{(E_n - E_m)(E_n - E_l)} |m\rangle - \sum_{m \neq n} \frac{\langle m | \partial_\mu H | n \rangle \langle n | \partial_\nu H | n \rangle}{(E_n - E_m)^2} |m\rangle \right. \\ & \left. - \frac{1}{2} \sum_{m \neq n} \frac{\langle n | \partial_\mu H | m \rangle \langle m | \partial_\nu H | n \rangle}{(E_n - E_m)^2} |n\rangle \right) x^\mu x^\nu + \dots \end{aligned}$$

Both energy derivatives and state derivatives will be involved in deduction. Whenever a state derivative is encountered, resolve it by inserting the complete set of basis, then the Hellmann-Feynman theorem is applicable. Because differentiation can be calculated systematically, the series expansion approach to the perturbative corrections can be coded on computers with symbolic processing software like Mathematica.

## Effective Hamiltonian

Let  $H(0)$  be the Hamiltonian completely restricted either in the low-energy subspace  $\mathcal{H}_L$  or in the high-energy subspace  $\mathcal{H}_H$ , such that there is no matrix element in  $H(0)$  connecting the low- and the high-energy subspaces, i.e.  $\langle m | H(0) | l \rangle = 0$  if  $m \in \mathcal{H}_L, l \in \mathcal{H}_H$ . Let  $F_\mu = \partial_\mu H$  be the coupling terms connecting the subspaces. Then when the high energy degrees of freedoms are integrated out, the effective Hamiltonian in the low energy subspace reads

$$H_{mn}^{\text{eff}}(x^\mu) = \langle m | H | n \rangle + \langle m | \partial_\mu H | n \rangle x^\mu + \frac{1}{2!} \sum_{l \in \mathcal{H}_H} \left( \frac{\langle m | \partial_\mu H | l \rangle \langle l | \partial_\nu H | n \rangle}{E_m - E_l} + \frac{\langle m | \partial_\nu H | l \rangle \langle l | \partial_\mu H | n \rangle}{E_n - E_l} \right) x^\mu x^\nu + \dots$$

Here  $m, n \in \mathcal{H}_L$  are restricted in the low energy subspace. The above result can be derived by power series expansion of  $\langle m | H(x^\mu) | n \rangle$ .

## ***Time-dependent perturbation theory***

### **Method of variation of constants**

Time-dependent perturbation theory, developed by Paul Dirac, studies the effect of a time-dependent perturbation  $V(t)$  applied to a time-independent Hamiltonian  $H_0$ . Since the perturbed Hamiltonian is time-dependent, so are its energy levels and eigenstates. Therefore, the goals of time-dependent perturbation theory are slightly different from time-independent perturbation theory. We are interested in the following quantities:

- The time-dependent expectation value of some observable  $A$ , for a given initial state.
- The time-dependent amplitudes of those quantum states that are energy eigenkets (eigenvectors) in the unperturbed system.

The first quantity is important because it gives rise to the classical result of an  $A$  measurement performed on a macroscopic number of copies of the perturbed system. For example, we could take  $A$  to be the displacement in the  $x$ -direction of the electron in a hydrogen atom, in which case the expected value, when multiplied by an appropriate coefficient, gives the time-dependent dielectric polarization of a hydrogen gas. With an appropriate choice of perturbation (i.e. an oscillating electric potential), this allows us to calculate the AC permittivity of the gas.

The second quantity looks at the time-dependent probability of occupation for each eigenstate. This is particularly useful in laser physics, where one is interested in the populations of different atomic states in a gas when a time-dependent electric field is applied. These probabilities are also useful for calculating the "quantum broadening" of spectral lines (see line broadening).

We will briefly examine the ideas behind Dirac's formulation of time-dependent perturbation theory. Choose an energy basis  $|n\rangle$  for the unperturbed system. (We will drop the  $(0)$  superscripts for the eigenstates, because it is not meaningful to speak of energy levels and eigenstates for the perturbed system.)

If the unperturbed system is in eigenstate  $|j\rangle$  at time  $t = 0$ , its state at subsequent times varies only by a phase (we are following the Schrödinger picture, where state vectors evolve in time and operators are constant):

$$|j(t)\rangle = e^{-iE_j t/\hbar} |j\rangle$$

We now introduce a time-dependent perturbing Hamiltonian  $V(t)$ . The Hamiltonian of the perturbed system is

$$H = H_0 + V(t)$$

Let  $|\psi(t)\rangle$  denote the quantum state of the perturbed system at time  $t$ . It obeys the time-dependent Schrödinger equation,

$$H|\psi(t)\rangle = i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle$$

The quantum state at each instant can be expressed as a linear combination of the eigenbasis  $|n\rangle$ . We can write the linear combination as

$$|\psi(t)\rangle = \sum_n c_n(t) e^{-iE_n t/\hbar} |n\rangle$$

where the  $c_n(t)$  are undetermined complex functions of  $t$  which we will refer to as **amplitudes** (strictly speaking, they are the amplitudes in the Dirac picture). We have explicitly extracted the exponential phase factors  $\exp(-iE_n t/\hbar)$  on the right hand side. This is only a matter of convention, and may be done without loss of generality. The reason we go to this trouble is that when the system starts in the state  $|j\rangle$  and no perturbation is present, the amplitudes have the convenient property that, for all  $t$ ,  $c_j(t) = 1$  and  $c_n(t) = 0$  if  $n \neq j$ .

The absolute square of the amplitude  $c_n(t)$  is the probability that the system is in state  $n$  at time  $t$ , since

$$|c_n(t)|^2 = |\langle n|\psi(t)\rangle|^2$$

Plugging into the Schrödinger equation and using the fact that  $\partial/\partial t$  acts by a chain rule, we obtain

$$\sum_n \left( i\hbar \frac{\partial c_n}{\partial t} - c_n(t) V(t) \right) e^{-iE_n t/\hbar} |n\rangle = 0$$

By resolving the identity in front of  $V$ , this can be reduced to a set of partial differential equations for the amplitudes:

$$\frac{\partial c_n}{\partial t} = \frac{-i}{\hbar} \sum_k \langle n|V(t)|k\rangle c_k(t) e^{-i(E_k - E_n)t/\hbar}$$

The matrix elements of  $V$  play a similar role as in time-independent perturbation theory, being proportional to the rate at which amplitudes are shifted between states. Note, however, that the direction of the shift is modified by the exponential phase factor. Over times much longer than the energy difference  $E_k - E_n$ , the phase winds many times. If the time-dependence of  $V$  is sufficiently slow, this may cause the state amplitudes to oscillate. Such oscillations are useful for managing radiative transitions in a laser.

Up to this point, we have made no approximations, so this set of differential equations is exact. By supplying appropriate initial values  $c_n(0)$ , we could in principle find an exact (i.e. non-perturbative) solution. This is easily done when there are only two energy levels ( $n = 1, 2$ ), and the solution is useful for modelling systems like the ammonia molecule. However, exact solutions are difficult to find when there are many energy levels, and one instead looks for perturbative solutions, which may be obtained by putting the equations in an integral form:

$$c_n(t) = c_n(0) + \frac{-i}{\hbar} \sum_k \int_0^t dt' \langle n|V(t')|k\rangle c_k(t') e^{-i(E_k - E_n)t'/\hbar}$$

By repeatedly substituting this expression for  $c_n$  back into right hand side, we get an iterative solution

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

where, for example, the first-order term is

$$c_n^{(1)}(t) = \frac{-i}{\hbar} \sum_k \int_0^t dt' \langle n|V(t')|k\rangle c_k(0) e^{-i(E_k - E_n)t'/\hbar}$$

Many further results may be obtained, such as Fermi's golden rule, which relates the rate of transitions between quantum states to the density of states at particular energies, and the Dyson series, obtained by applying the iterative method to the time evolution operator, which is one of the starting points for the method of Feynman diagrams.

### Method of Dyson series

Time dependent perturbations can be treated with the technique of Dyson series. Taking Schrödinger equation

$$H(t)|\psi(t)\rangle = i\hbar \frac{\partial |\psi(t)\rangle}{\partial t}$$

this has the formal solution

$$|\psi(t)\rangle = T \exp \left[ -\frac{i}{\hbar} \int_{t_0}^t dt' H(t') \right] |\psi(t_0)\rangle$$

Being  $T$  the time ordering operator such that

$$TA(t_1)A(t_2) = A(t_1)A(t_2)$$

if  $t_1 > t_2$  and

$$TA(t_1)A(t_2) = A(t_2)A(t_1)$$

if  $t_2 > t_1$  so that the exponential will represent the following Dyson series

$$|\psi(t)\rangle = \left[ 1 - \frac{i}{\hbar} \int_{t_0}^t dt_1 H(t_1) - \frac{1}{\hbar^2} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 H(t_1)H(t_2) + \dots \right] |\psi(t_0)\rangle$$

Now, let us take the following perturbation problem

$$[H_0 + \lambda V(t)]|\psi(t)\rangle = i\hbar \frac{\partial |\psi(t)\rangle}{\partial t}$$

assuming that the parameter  $\lambda$  is small and that we are able to solve the problem  $H_0|n\rangle = E_n|n\rangle$ . We do the following unitary transformation going to interaction picture or Dirac picture

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}H_0(t-t_0)}|\psi_I(t)\rangle$$

and so the Schrödinger equation becomes

$$\lambda e^{\frac{i}{\hbar}H_0(t-t_0)}V(t)e^{-\frac{i}{\hbar}H_0(t-t_0)}|\psi_I(t)\rangle = i\hbar \frac{\partial |\psi_I(t)\rangle}{\partial t}$$

that can be solved through the above Dyson series as

$$|\psi_I(t)\rangle = \left[ 1 - \frac{i\lambda}{\hbar} \int_{t_0}^t dt_1 e^{\frac{i}{\hbar}H_0(t_1-t_0)}V(t_1)e^{-\frac{i}{\hbar}H_0(t_1-t_0)} - \frac{\lambda^2}{\hbar^2} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 e^{\frac{i}{\hbar}H_0(t_1-t_0)}V(t_1)e^{-\frac{i}{\hbar}H_0(t_1-t_0)}e^{\frac{i}{\hbar}H_0(t_2-t_0)}V(t_2)e^{-\frac{i}{\hbar}H_0(t_2-t_0)} + \dots \right] |\psi(t_0)\rangle$$

being this a perturbation series with small  $\lambda$ . Using the solution of the unperturbed problem  $H_0|n\rangle = E_n|n\rangle$  and  $\sum_n |n\rangle\langle n| = 1$  (for the sake of simplicity we assume a pure discrete spectrum), we will have till first order

$$|\psi_I(t)\rangle = \left[ 1 - \frac{i\lambda}{\hbar} \sum_m \sum_n \int_{t_0}^t dt_1 \langle m|V(t_1)|n\rangle e^{-\frac{i}{\hbar}(E_n-E_m)(t_1-t_0)}|m\rangle\langle n| + \dots \right] |\psi(t_0)\rangle$$

So, the system, initially in the unperturbed state  $|\alpha\rangle = |\psi(t_0)\rangle$ , due to the perturbation can go into the state  $|\beta\rangle$ . The corresponding probability amplitude will be

$$A_{\alpha\beta} = -\frac{i\lambda}{\hbar} \int_{t_0}^t dt_1 \langle \beta|V(t_1)|\alpha\rangle e^{-\frac{i}{\hbar}(E_\alpha-E_\beta)(t_1-t_0)}$$

and the corresponding transition probability will be given by Fermi's golden rule.

Time independent perturbation theory can be derived from the time dependent perturbation theory. For this purpose, let us write the unitary evolution operator, obtained from the above Dyson series, as

$$U(t) = 1 - \frac{i\lambda}{\hbar} \int_{t_0}^t dt_1 e^{\frac{i}{\hbar} H_0(t_1-t_0)} V(t_1) e^{-\frac{i}{\hbar} H_0(t_1-t_0)} \\ - \frac{\lambda^2}{\hbar^2} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 e^{\frac{i}{\hbar} H_0(t_1-t_0)} V(t_1) e^{-\frac{i}{\hbar} H_0(t_1-t_0)} e^{\frac{i}{\hbar} H_0(t_2-t_0)} V(t_2) e^{-\frac{i}{\hbar} H_0(t_2-t_0)} + \dots$$

and we take the perturbation  $V$  time independent. Using the identity

$$\sum_n |n\rangle \langle n| = 1$$

with  $H_0|n\rangle = E_n|n\rangle$  for a pure discrete spectrum, we can write

$$U(t) = 1 - \frac{i\lambda}{\hbar} \int_{t_0}^t dt_1 \sum_m \sum_n \langle m|V|n\rangle e^{-\frac{i}{\hbar}(E_n-E_m)(t_1-t_0)} |m\rangle \langle n| \\ - \frac{\lambda^2}{\hbar^2} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \sum_m \sum_n \sum_q e^{-\frac{i}{\hbar}(E_n-E_m)(t_1-t_0)} \langle m|V|n\rangle \langle n|V|q\rangle e^{-\frac{i}{\hbar}(E_q-E_n)(t_2-t_0)} |m\rangle \langle q| + \dots$$

We see that, at second order, we have to sum on all the intermediate states. We assume  $t_0 = 0$  and the asymptotic limit of larger times. This means that, at each contribution of the perturbation series, we have to add a multiplicative factor  $e^{-\epsilon t}$  in the integrands so that, the limit  $t \rightarrow \infty$  will give back the final state of the system by eliminating all oscillating terms but keeping the secular ones.  $\epsilon$  must be postulated as being arbitrarily small. In this way we can compute the integrals and, separating the diagonal terms from the others, we have

$$U(t) = 1 - \frac{i\lambda}{\hbar} \sum_n \langle n|V|n\rangle t - \frac{i\lambda^2}{\hbar} \sum_{m \neq n} \frac{\langle n|V|m\rangle \langle m|V|n\rangle}{E_n - E_m} t - \frac{1}{2} \frac{\lambda^2}{\hbar^2} \sum_{m,n} \langle n|V|m\rangle \langle m|V|n\rangle t^2 + \dots \\ + \lambda \sum_{m \neq n} \frac{\langle m|V|n\rangle}{E_n - E_m} |m\rangle \langle n| \\ + \lambda^2 \sum_{m \neq n} \sum_{q \neq n} \sum_n \frac{\langle m|V|n\rangle \langle n|V|q\rangle}{(E_n - E_m)(E_q - E_n)} |m\rangle \langle q| + \dots$$

where the time secular series yields the eigenvalues of the perturbed problem and the remaining part gives the corrections to the eigenfunctions. The unitary evolution operator is applied to whatever eigenstate of the unperturbed problem and, in this case, we will get a secular series that holds at small times.

### **Strong perturbation theory**

In a similar way as for small perturbations, it is possible to develop a strong perturbation theory. Let us consider as usual the Schrödinger equation

$$H(t)|\psi(t)\rangle = i\hbar \frac{\partial|\psi(t)\rangle}{\partial t}$$

and we consider the question if a dual Dyson series exists that applies in the limit of a perturbation increasingly large. This question can be answered in an affirmative way and the series is the well-known adiabatic series. This approach is quite general and can be shown in the following way. Let us consider the perturbation problem

$$[H_0 + \lambda V(t)]|\psi(t)\rangle = i\hbar \frac{\partial|\psi(t)\rangle}{\partial t}$$

being  $\lambda \rightarrow \infty$ . Our aim is to find a solution in the form

$$|\psi\rangle = |\psi_0\rangle + \frac{1}{\lambda}|\psi_1\rangle + \frac{1}{\lambda^2}|\psi_2\rangle + \dots$$

but a direct substitution into the above equation fails to produce useful results. This situation can be adjusted making a rescaling of the time variable as  $\tau = \lambda t$  producing the following meaningful equations

$$\begin{aligned} V(t)|\psi_0\rangle &= i\hbar \frac{\partial|\psi_0\rangle}{\partial \tau} \\ V(t)|\psi_1\rangle + H_0|\psi_0\rangle &= i\hbar \frac{\partial|\psi_1\rangle}{\partial \tau} \\ &\vdots \end{aligned}$$

that can be solved once we know the solution of the leading order equation. But we know that in this case we can use the adiabatic approximation. When  $V(t)$  does not depend on time one gets the Wigner-Kirkwood series that is often used in statistical mechanics. Indeed, in this case we introduce the unitary transformation

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}\lambda V(t-t_0)}|\psi_F(t)\rangle$$

that defines a **free picture** as we are trying to eliminate the interaction term. Now, in dual way with respect to the small perturbations, we have to solve the Schrödinger equation

$$e^{\frac{i}{\hbar}\lambda V(t-t_0)} H_0 e^{-\frac{i}{\hbar}\lambda V(t-t_0)} |\psi_F(t)\rangle = i\hbar \frac{\partial|\psi_F(t)\rangle}{\partial t}$$

and we see that the expansion parameter  $\lambda$  appears only into the exponential and so, the corresponding Dyson series, a **dual Dyson series**, is meaningful at large  $\lambda$ s and is



$$|\psi_F(t)\rangle = \left[ 1 - \frac{i}{\hbar} \int_{t_0}^t dt_1 e^{\frac{i}{\hbar} \lambda V(t_1-t_0)} H_0 e^{-\frac{i}{\hbar} \lambda V(t_1-t_0)} - \frac{1}{\hbar^2} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 e^{\frac{i}{\hbar} \lambda V(t_1-t_0)} H_0 e^{-\frac{i}{\hbar} \lambda V(t_1-t_0)} e^{\frac{i}{\hbar} \lambda V(t_2-t_0)} H_0 e^{-\frac{i}{\hbar} \lambda V(t_2-t_0)} + \dots \right] |\psi(t_0)\rangle.$$

After the rescaling in time  $\tau = \lambda t$  we can see that this is indeed a series in  $1/\lambda$  justifying in this way the name of **dual Dyson series**. The reason is that we have obtained this series simply interchanging  $H_0$  and  $V$  and we can go from one to another applying this exchange. This is called **duality principle** in perturbation theory. The choice  $H_0 = p^2/2m$  yields, as already said, a Wigner-Kirkwood series that is a gradient expansion. The Wigner-Kirkwood series is a semiclassical series with eigenvalues given exactly as for WKB approximation.

### Examples

#### Example of first order perturbation theory - Ground State Energy of the Quartic Oscillator

Let us consider the quantum harmonic oscillator with the quartic potential perturbation and the Hamiltonian

$$H = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \frac{m\omega^2 x^2}{2} + \lambda x^4$$

The ground state of the harmonic oscillator is

$$\psi_0 = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\alpha x^2/2}$$

( $\alpha = m\omega/\hbar$ ) and the energy of unperturbed ground state is

$$E_0^{(0)} = \frac{1}{2} \hbar \omega.$$

Using the first order correction formula we get

$$E_0^{(1)} = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \int e^{-\alpha x^2/2} x^4 e^{-\alpha x^2/2} dx = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \frac{\partial^2}{\partial \alpha^2} \int e^{-\alpha x^2}$$

or

$$E_0^{(1)} = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \frac{\partial^2}{\partial \alpha^2} \left(\frac{\pi}{\alpha}\right)^{\frac{1}{2}} = \frac{3}{4} \frac{1}{\alpha^2} = \frac{3}{4} \frac{\hbar^2}{m^2 \omega^2}$$

## Example of first and second order perturbation theory - Quantum Pendulum

Consider the quantum mathematical pendulum with the Hamiltonian

$$H = -\frac{\hbar^2}{2ma^2} \frac{\partial^2}{\partial \phi^2} - \lambda \cos \phi$$

with the potential energy  $-\lambda \cos \phi$  taken as the perturbation i.e.

$$V = -\cos \phi$$

The unperturbed normalized quantum wave functions are those of the rigid rotor and are given by

$$\psi_n(\phi) = e^{in\phi} / \sqrt{2\pi}$$

and the energies

$$E_n^{(0)} = \frac{\hbar^2 n^2}{2ma^2}$$

The first order energy correction to the rotor due to the potential energy is

$$E_n^{(1)} = -\frac{1}{2\pi} \int e^{-in\phi} \cos \phi e^{in\phi} = -\frac{1}{2\pi} \int \cos \phi = 0$$

Using the formula for the second order correction one gets

$$E_n^{(2)} = \frac{ma^2}{2\pi^2 \hbar^2} \sum_k \left| \int e^{-ik\phi} \cos \phi e^{in\phi} \right|^2 / (n^2 - k^2)$$

or

$$E_n^{(2)} = \frac{ma^2}{2\hbar^2} \sum_k |(\delta_{n,1-k} + \delta_{n,-1-k})|^2 / (n^2 - k^2)$$

or

$$E_n^{(2)} = \frac{ma^2}{2\hbar^2} \left( \frac{1}{2n-1} + \frac{1}{-2n-1} \right) = \frac{ma^2}{\hbar^2} \frac{1}{4n^2-1}$$