Contents

Chapter 6

Perturbative methods

We have encountered a number of exactly solvable problems so far: The quantum harmonic oscillator, and the hydrogen atom. Unfortunately, the list of exactly solvable problems is actually shorted than one might wish. Usually, there is an algebraic structure in a problem that can be exploited that renders problems exactly solvable. Many problems in physics do not have this form. Then one has to have a general strategy of how to proceed. In this chapter, we will have a first look at approximate methods. In the center of this will be standard perturbation theory. We will look at this in some detail. The general mindset is that if one perturbs a known system a little bit, then one should expect that the solution still looks quite similar to the original one. And indeed, this intuition is true to a large extent (unless one is really in the situation of quantum many-body theory).

6.1 Time independent perturbation theory

Consider a Hamiltonian $H^{(0)}$ that has a "known" solution: The eigenvalues and eigenvectors are assumed to be known. Let us assume that such eigenvectors exist (and disregard mathematical fine print in case that one has a continuous spectrum). So we can write the eigenvalue decomposition as

$$H^{(0)} = \sum_j E_j^{(0)} |\psi_j^{(0)}\rangle\langle\psi_j^{(0)}|. \tag{6.1}$$

We are now interested to grasp properties of

$$H(\lambda) = H^{(0)} + \lambda V, \tag{6.2}$$

for some real λ and a peturbing Hamiltonian V, in case that λV is "small". More precisely, what we want to estimate is the eigenvalue decomposition of H,

$$H(\lambda) = \sum_j E_j(\lambda) |\psi_j(\lambda)\rangle\langle\psi_j(\lambda)|, \tag{6.3}$$

in terms of the eigenvalue decomposition of $H^{(0)}$.

6.1.1 Non-degenerate situation

We first consider the simpler case of a non-degenerate spectrum of $H^{(0)}$. We assume that for $\lambda \in [0, \Lambda]$, so "for a sufficiently small perturbation", $H^{(0)} + \lambda V$ has non-degenerate eigenvalues $E_j(\lambda)$ associated with eigenvectors $|\psi_j(\lambda)\rangle$, each of them with an analytic dependence on λ in the above interval. To find the eigenvectors and eigenvalues, we obviously need so solve

$$H(\lambda)|\psi_j(\lambda)\rangle = E_j(\lambda)|\psi_j(\lambda)\rangle. \tag{6.4}$$

If $\lambda \in [0, \Lambda]$, then we can make use of a power series expansion in λ and write

$$E_j(\lambda) = \sum_{n=0}^{\infty} \lambda^n E_j^{(n)}, \tag{6.5}$$

$$|\psi_j(\lambda)\rangle = \sum_{n=0}^{\infty} \lambda^n |\psi_j^{(n)}\rangle. \tag{6.6}$$

The first of the two series is referred to as *Rayleigh-Schroedinger series*. Inserting this series into the eigenvalue expression (6.4) gives

$$\left(H^{(0)} + \lambda V\right) \sum_{n=0}^{\infty} \lambda^n |\psi_j^{(n)}\rangle = \sum_{k=0}^{\infty} \lambda^k E_j^{(k)} \sum_{n=0}^{\infty} \lambda^n |\psi_j^{(n)}\rangle. \tag{6.7}$$

Reordering the series and shifting the indices, we get

$$\begin{aligned} & (H^{(0)} - E_j^{(0)})|\psi_j^{(0)}\rangle \\ + & \sum_{n=1}^{\infty} \lambda^n \left(H^{(0)}|\psi_j^{(n)}\rangle + V|\psi_j^{(n-1)}\rangle - \sum_{k=0}^{n} E_j^{(k)}|\psi_j^{(n-k)}\rangle \right) = 0. \end{aligned} \tag{6.8}$$

This can only be true for all λ if the expression vanishes for each pre-factor of each of the powers of λ, i.e.,

$$(H^{(0)} - E_j^{(0)})|\psi_j^{(0)}\rangle = 0, \tag{6.9}$$

$$(H^{(0)} - E_j^{(0)})|\psi_j^{(1)}\rangle = (E_j^{(1)} - V)|\psi_j^{(0)}\rangle, \tag{6.10}$$

$$(H^{(0)} - E_j^{(0)})|\psi_j^{(n)}\rangle = (E_j^{(1)} - V)|\psi_j^{(n-1)}\rangle + \sum_{k=2}^{n} E_j^{(k)}|\psi_j^{(n-k)}\rangle, \tag{6.11}$$

for $n = 2, 3, \ldots$. What is more, we still have the freedom to pick the normalization of the vectors $|\psi_j^{(n)}\rangle$ for $n = 1, 2, \ldots$. This is delicate, as this must be consistent with the normalization of the family of state vectors $|\psi_j(\lambda)\rangle$. It is most convenient to take

$$\langle \psi_j^{(0)} | \psi_j(\lambda) \rangle = 1 \tag{6.12}$$

which means that

$$\langle\psi_j^{(0)}|\psi_j^{(n)}\rangle = 0 \tag{6.13}$$

for $n = 1, 2, \ldots$. The contributions of all other orders are hence orthogonal to the one of zeroth order. Eqs. (6.9, 6.11, 6.13) gives a solution to the problem at hand.

The scalar product of Eq. (6.11) with $|\psi_j^{(0)}\rangle$ together with Eq. (6.13) gives the perturbed energy values: In fact, we get

$$0 = \langle\psi_j^{(0)}|(E_j^{(1)} - V)|\psi_j^{(n-1)}\rangle + \sum_{k=2}^{n} E_j^{(k)}\langle\psi_j^{(0)}|\psi_j^{(n-k)}\rangle. \tag{6.14}$$

Both in case of $n = 1$ and $n \geq 2$ – two cases that can be individually treated – one gets

$$E_j^{(1)} = \langle\psi_j^{(0)}|V|\psi_j^{(0)}\rangle \tag{6.15}$$

and more generally,

$$E_j^{(n)} = \langle\psi_j^{(0)}|V|\psi_j^{(n-1)}\rangle \tag{6.16}$$

for all j and all $n \geq 1$. But we can go further than that. We can also compute the scalar product of Eq. (6.11) with $|\psi_l^{(0)}\rangle$ with $l \neq j$: Using

$$H^{(0)}|\psi_l^{(0)}\rangle = E_l^{(0)}|\psi_l^{(0)}\rangle, \tag{6.17}$$

we get first

$$\langle\psi_l^{(0)}|\psi_j^{(1)}\rangle = \frac{1}{E_j^{(0)} - E_l^{(0)}}\langle\psi_l^{(0)}|V|\psi_j^{(0)}\rangle, \tag{6.18}$$

and, more generally, for $n \geq 2$,

$$\langle\psi_l^{(0)}|\psi_j^{(n)}\rangle = \frac{1}{E_j^{(0)} - E_l^{(0)}}\left(\langle\psi_l^{(0)}|V|\psi_j^{(n-1)}\rangle - \sum_{k=1}^{n-1} E_j^{(k)}\langle\psi_l^{(0)}|\psi_j^{(n-k)}\rangle\right). \tag{6.19}$$

Using the fact that the eigenvectors of $H^{(0)}$ are complete, from Eqs. (6.13, 6.19) we find that

$$\begin{aligned} |\psi_j^{(n)}\rangle &= \sum_{l=0}^{\infty} |\psi_l^{(0)}\rangle\langle\psi_l^{(0)}|\psi_j^{(n)}\rangle \\ &= \sum_{l\neq j} \frac{1}{E_j^{(0)} - E_l^{(0)}} \\ &\times \left(\langle\psi_l^{(0)}|V|\psi_j^{(n-1)}\rangle - \sum_{k=1}^{n-1} E_j^{(k)}\langle\psi_k^{(0)}|\psi_j^{(n-k)}\rangle\right)|\psi_l^{(0)}\rangle. \end{aligned} \tag{6.20}$$

This is the desired expression. We can simplify the notation by introducing the operator

$$S_j = \sum_{k \neq j} \frac{|\psi_k^{(0)}\rangle\langle\psi_k^{(0)}|}{E_j^{(0)} - E_k^{(0)}}. \tag{6.21}$$

This gives

$$|\psi_j^{(n)}\rangle = S_j \left(V|\psi_j^{(n-1)}\rangle - \sum_{k=1}^{n-1} |\psi_j^{(n-k)}\rangle \right). \tag{6.22}$$

Using this expression, we can determine all perturbed eigenvectors in an iterative fashion. Usually, only the first orders of this expansion are computed. We summarize these expressions in a box.

First terms in non-degenerate perturbation theory:

$$E_j^{(1)} = \langle\psi_j^{(0)}|V|\psi_j^{(0)}\rangle, \tag{6.23}$$

$$|\psi_j^{(1)}\rangle = S_j V|\psi_j^{(0)}\rangle = \sum_{l \neq j} \frac{\langle\psi_l^{(0)}|V|\psi_j^{(0)}\rangle}{E_j^{(0)} - E_l^{(0)}}|\psi_l^{(0)}\rangle, \tag{6.24}$$

$$E_l^{(2)} = \langle\psi_j^{(0)}|VS_jV|\psi_j^{(0)}\rangle = \sum_{l \neq j} \frac{|\langle\psi_j^{(0)}|V|\psi_l^{(0)}\rangle|^2}{E_j^{(0)} - E_l^{(0)}}. \tag{6.25}$$

In order to compute corrections of first order, therefore, we merely need the matrix elements

$$\{\langle\psi_l^{(0)}|V|\psi_j^{(0)}\rangle : l, j = 0, 1, \dots\}, \tag{6.26}$$

from these we can then compute the corrections in the eigenvalues as well as the eigenvectors. The energy shift is obtained from expectation values of the perturbing Hamiltonian in the unperturbed eigenvectors. The new eigenvectors are suitable linear combinations of the unperturbed eigenvectors. For higher orders, one can iteratively proceed as described above. Usually, in small orders, this procedure gives rise to a very good approximation of the perturbed problem.

A number of remarks are in order at this point.

- When considering the ground state energy, the correction $E_j^{(2)}$ is always non-positive.
- When judging the strength of a perturbation, rather not λ alone is relevant, but rather $\lambda\|V\|$, where $\|.\|$ is the operator norm of V, so its largest eigenvalue.
- In general, it is far from clear that the above series are convergent in any interval $[0, \Lambda]$ with $\lambda > 0$. One should not be too worried about this,

however. Often, the expressions for the energy eigenvalues can be analytically continued to the complex λ plane cut along the negative real axis, and one finds that the perturbed energy eigenvalues give rise to a strong asymptotic expansion around $\lambda = 0$. For practical purposes, this means that – although the series may not be strictly convergent – the "first perturbative terms still yield a very good approximation". Only when then again considering very large orders of the series, the approximation becomes worse again.

6.1.2 An example: The anharmonic oscillator

In order to exemplify the strategy outlined in the last subsection, we investigate the problem of an anharmonic oscillator with Hamiltonian

$$\begin{aligned} H(\lambda) &= H^{(0)} + \lambda V, && (6.27)\\ H^{(0)} &= \frac{1}{2m}P^2 + \frac{k}{2}X^2, && (6.28)\\ V &= X^4. && (6.29) \end{aligned}$$

Using the above expressions, we find for the perturbed ground state energy

$$E_0(\lambda) = 1 + \frac{3}{4}\lambda - \frac{21}{16}\lambda^2 + O(\lambda^3). \tag{6.30}$$

Note that the second order contribution is negative. Numerically, one can go evaluate the above expressions to any order – although this is not the point of perturbation theory: the computational effort will grow exponentially with the order. Say, pick $\lambda = 0.1$ and let us denote with $E_0(m)$ the full perturbed ground state energy up to order m, so

$$E_0(m) = \sum_{n=0}^{m} \lambda^j E_0^{(n)}. \tag{6.31}$$

Then we get the following values:

n	$E_0(m)$
1	1.075 000
2	1.061 875
3	1.067 078
4	1.064 062
5	1.066 300

As a comparison, the true value of the ground state energy, numerically computed for $\lambda = 0.1$, is 1.065285. The approximation is therefore very soon very good. For practical purposes, one often considers corrections to the first order that is non-vanishing, so often first or second order (which already provides a good approximation). It is worth mentioning that due to the fact that

strictly speaking, the series is often not convergent, the very high order terms can be very much off. E.g.,

$$E_0(20) = -9.919902, \tag{6.32}$$

so is totally wrong.

6.1.3 Degenerate situation

So far, we have considered the non-degenerate situation. In case of degeneracies, one needs some modifications. Let us assume that

$$H^{(0)}|\psi_j^{(0)}\rangle = E|\psi_j^{(0)}\rangle \tag{6.33}$$

for $j = 1, \ldots, s$, so $|\psi_1^{(0)}\rangle, \ldots, |\psi_s^{(0)}\rangle$ are degenerate. Since the perturbation series is an expansion of the form

$$|\psi_j^{(1)}\rangle = \sum_{l \neq j} \frac{\langle \psi_l^{(0)}|V|\psi_j^{(0)}\rangle}{E_j^{(0)} - E_l^{(0)}} |\psi_l^{(0)}\rangle, \tag{6.34}$$

it is clear that something must go wrong when the above formulae are naively applied. But the solution is also quite obvious: One has to use a basis such that the terms in the nominator vanish whenever the terms in the denominator vanish, in order not to have divergent terms. Let us introduce the Hermitian matrix $M = M^\dagger$ with components

$$M_{l,j} = \langle \psi_l^{(0)}|V|\psi_j^{(0)}\rangle. \tag{6.35}$$

Obviously, we can diagonalize this matrix M with a unitary, so find a U with

$$U^\dagger M U = D, \tag{6.36}$$

D being diagonal, and

$$UU^\dagger = U^\dagger U = \mathbb{1}. \tag{6.37}$$

This means that

$$\sum_{l,j} \langle \psi_l^{(0)}|U_{m,l}^* V U_{j,k}|\psi_j^{(0)}\rangle = 0 \text{ for } m \neq k. \tag{6.38}$$

In this way, one obtains unitaries U and therefore D and can then proceed as before in order to get the perturbed expressions.

6.2 Time-dependent perturbation theory

6.2.1 Interaction picture

So far, we have discussed the perturbation of eigenvectors and eigenvalues of Hamiltonians. For time-dependent problems, there is a specific kind of perturbation theory known as *time-dependent perturbation theory*. This links directly

to the discussion of time evolution that we have studied earlier. We consider time-dependent Hamiltonians of the form

$$H(t) = H^{(0)} + \lambda V(t). \tag{6.39}$$

Of course, time evolution is now governed by the Schroedinger equation. Here, we would like to follow again a perturbative approach, however, where V is again treated as a small perturbation of the problem dictated by $H^{(0)}$ alone, which in turn is assumed to be simply solvable. What is more, we assume that

$$V(t) = 0, \text{ for } t \leq t_0 \tag{6.40}$$

for some time t_0.

For $t \leq t_0$ the state vector is denoted by $|\psi^{(0)}(t)\rangle$, following

$$i\hbar \frac{\partial}{\partial t}|\psi^{(0)}(t)\rangle = H^{(0)}|\psi^{(0)}(t)\rangle. \tag{6.41}$$

For $t > t_0$, the state vector follows the Schroedinger equation

$$i\hbar \frac{\partial}{\partial t}|\psi(t)\rangle = (H^{(0)} + \lambda V(t))|\psi(t)\rangle, \tag{6.42}$$

with initial condition

$$|\psi(t)\rangle = |\psi^{(0)}(t)\rangle \tag{6.43}$$

for $t \leq t_0$. In order to proceed, it is helpful to separate off the part of the dynamics due to $H^{(0)}$: We anticipated such an approach already earlier when we discussed the Heisenberg and the Schroedinger picture. Here we introduce the *interaction picture* which in a sense interpolates between the two previous pictures. The state vectors in the interaction picture are given by

$$|\psi(t)\rangle_I = e^{iH^{(0)}t/\hbar}|\psi(t)\rangle. \tag{6.44}$$

Differentiating Eq. (6.44) in time and using the Schroedinger equation, one finds

$$i\hbar \frac{\partial}{\partial t}|\psi(t)\rangle_I = -H^{(0)}|\psi(t)\rangle_I + e^{iH^{(0)}t/\hbar}\left(H^{(0)} + \lambda V(t)\right)|\psi(t)\rangle. \tag{6.45}$$

In other words,

$$i\hbar \frac{\partial}{\partial t}|\psi(t)\rangle_I = \lambda V_I(t)|\psi(t)\rangle_I, \tag{6.46}$$

where we have introduced the following:

Interaction operator in the interaction picture: For $t \geq t_0$, define

$$V_I(t) = e^{iH^{(0)}t/\hbar}V(t)e^{-iH^{(0)}t/\hbar}. \tag{6.47}$$

Time integration now yields the integral equation

$$|\psi(t)\rangle_I = |\psi(t_0)\rangle_I + \frac{\lambda}{i\hbar}\int_{t_0}^{t} ds V_I(s)|\psi(s)\rangle_I. \tag{6.48}$$

Iterating this equation, one obtains

$$\begin{aligned} |\psi(t)\rangle_I &= |\psi(t_0)\rangle_I + \frac{\lambda}{i\hbar}\int_{t_0}^{t} ds V_I(s)|\psi(t_0)\rangle_I \\ &+ \frac{\lambda^2}{(i\hbar)^2}\int_{t_0}^{t} ds \int_{t_0}^{s} dr V_I(s)V_I(r)|\psi(t_0)\rangle_I + O(\lambda^3). \end{aligned} \tag{6.49}$$

This series is referred to as *von Neumann series*.

6.2.2 Applications: Transitions of first order

We now discuss a simple application of time-dependent perturbation theory: The theory of transitions of first order.

- Let us assume that initially, at time t_0, a system is in an eigenvector $|\psi_j^{(0)}\rangle$ of the unperturbed Hamiltonian $H^{(0)}$, so $|\psi(t_0)\rangle = |\psi_j^{(0)}\rangle$.
- At time t_0 we switch on $V(t)$.

We would like to know the probability of finding at time $t > t_0$ the system in

$$|\psi_k^{(0)}(t)\rangle = e^{-iH^{(0)}t/\hbar}|\psi_k^{(0)}\rangle = e^{-iE_k^{(0)}t/\hbar}|\psi_k^{(0)}\rangle \tag{6.50}$$

for some $k \neq j$. The probability amplitude for this transition (or rather, the probability of obtaining such an eigenvector upon measuring, but it should be clear what is meant) is

$$\langle\psi_k^{(0)}(t)|\psi(t)\rangle = \langle\psi_k^{(0)}|e^{iH^{(0)}t/\hbar}|\psi(t)\rangle = \langle\psi_k^{(0)}|\psi(t)\rangle_I. \tag{6.51}$$

Inserting $|\psi(t_0)\rangle = |\psi_j^{(0)}\rangle$ into Eq. (6.49), we find to first order

$$|\psi(t)\rangle_I = |\psi_j^{(0)}\rangle + \frac{\lambda}{i\hbar}\int_{t_0}^{t} ds V_I(s)|\psi_j^{(0)}\rangle + O(\lambda^2). \tag{6.52}$$

Hence we get for the transition amplitude

$$\begin{aligned} \langle\psi_k^{(0)}(t)|\psi(t)\rangle &= \delta_{j,k} + \frac{\lambda}{i\hbar}\int_{t_0}^{t} ds\langle\psi_k^{(0)}|V_I(s)|\psi_j^{(0)}\rangle + O(\lambda^2) \\ &= \delta_{j,k} + \frac{\lambda}{i\hbar}\int_{t_0}^{t} ds e^{i(E_k^{(0)}-E_j^{(0)})s/\hbar}\langle\psi_k^{(0)}|V(s)|\psi_j^{(0)}\rangle + O(\lambda^2). \end{aligned} \tag{6.53}$$

The probability of a transition $P_{j,k}(t)$ is the absolute value of this expression

$$\begin{aligned} P_{j,k}(t) &= |\langle\psi_k^{(0)}(t)|\psi(t)\rangle|^2 \\ &= \left|\frac{\lambda}{\hbar}\int_{t_0}^{t} ds e^{i(E_k^{(0)}-E_j^{(0)})s/\hbar}\langle\psi_k^{(0)}|V(s)|\psi_j^{(0)}\rangle\right|^2 + O(\lambda^4). \quad (6.54) \end{aligned}$$

This is a very useful expression in many contexts. Note that again, in order to evaluate this formula, we merely need to know the eigenvectors and eigenvalues of the unperturbed Hamiltonian $H^{(0)}$, which is assumed to be "simple". In the situation of a scattering in a continuum of modes, this formula is referred to as the "golden rule".

6.3 Some words on more general concepts of perturbation

Indeed, the remark that perturbation series are often not convergent should not be a reason for too much worry. For good reasons, low order perturbation theory is a work horse in quantum theory. Frankly, it is countlessly employed in studies in quantum physics. What is more, we have seen just the tip of the iceberg so far. There are many concepts of perturbation in quantum theory. In each instance, the intuition that "a small perturbation has a small effect" is at the basis of reasoning.

- A quite general idea of perturbation is *Kato's theory of perturbation*, which is based on a discussion of the resolvent of a Hamiltonian H, defined as the function
$$G(z) = \frac{1}{z\mathbb{1} - H}, \quad (6.55)$$
for $z \in \mathbb{C}$. A complete treatment of this is however beyond the scope of this course.

- In the many-body context, the concept of *relative boundedness* if often relevant, where a perturbation λV is small compared to $H^{(0)}$ in the sense that
$$(H^{(0)})^2 \geq \lambda^2 V^2. \quad (6.56)$$
This means – just as for the density operator that we considered earlier – that the eigenvalues of the operator $(H^{(0)})^2 - \lambda^2 V^2$ are non-negative.

- An intuition that is correct to a large extent is that "eigenvalues are better behaved than eigenvectors". In particular, degeneracies matter less. As an educational example, consider the matrices
$$A = \begin{bmatrix} 1+\varepsilon & 0 \\ 0 & 1-\varepsilon \end{bmatrix}, \quad (6.57)$$
$$B = \begin{bmatrix} 1 & \varepsilon \\ \varepsilon & 1 \end{bmatrix}. \quad (6.58)$$

The eigenvalues of A are $1+\varepsilon$ and $1-\varepsilon$. The same is true for B. The normalized eigenvectors of A are $[1,0]$ and $[0,1]$, while they are $[1,1]/\sqrt{2}$ and $[1,-1]/\sqrt{2}$ for B. As $\varepsilon \to 0$ the two matrices approach each other, but the eigenvectors do not.