Chapter 17. Time-Independent Perturbation Theory of Non-Degenerate States

Section 17.1. Who Needs Perturbation Theory?

§ 1 *Introduction.* There are many problems in which the Hamiltonian has the form

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{1}$$

In this lecture we examine how one gets approximate solutions for the eigenstates and the eigenvalues of $\hat{H}$ if the following conditions are satisfied:

1. We have solved the eigenvalue problem

$$\hat{H}_0 \ |E_n\rangle = E_n \ |E_n\rangle, \tag{2}$$

   which means that we can calculate any matrix elements involving $|E_n\rangle$ for any operator defined in the same space

2. We also know that in some sense, to be specified later, $\hat{V}$ is substantially smaller than $\hat{H}_0$.

3. $\hat{V}$ is time independent. If all degrees of freedom are treated quantum mechanically this requirement is automatically fulfilled when we work in the Schrödinger representation.

When these requirements are satisfied it is possible to calculate an approximate solution of the eigenvalue problem

$$\hat{H} \ |\epsilon_n\rangle = \epsilon_n \ |\epsilon_n\rangle \tag{3}$$

which is essentially an expansion of ϵ_n and $|\epsilon_n\rangle$ in powers of $\hat{V}$. The procedure for doing this is called perturbation theory. The terms first order or second order perturbation theory refer to the power of $\hat{V}$ in this expansion.

Here are a few examples of problems that can be treated by perturbation theory:

1. The change in the spectrum of a molecule or in its dipole moment when it is exposed to an external, static electric or magnetic field.

2. Corrections to the harmonic approximation when the difference between the harmonic potential energy differs slightly from the real anharmonic one.

3. Corrections to the assumption that vibrations and rotation are independent.

4. Corrections to the energy eigenvalues due to the effect of spin-orbit coupling or other spin related effects (fine and hyperfine structure in spectra)

The procedure is sufficiently abstract to be applicable to any problem in physics or engineering that satisfies the conditions specified above. One could for example use it to find corrections to the vibrations of string or a membrane because the material's properties differ from those of a perfectly elastic one.

There are however problems to which perturbation theory is not applicable even if $\hat{V}$ is very small. The theory is such that there is a one-to-one connection between the eigenvalues of $\hat{H}_0$ and those of $\hat{H}$. Because of this, the method fails for problems in which the nature of the spectrum of $\hat{H}$ differs from that of the spectrum of $\hat{H}_0$. For example, the superconducting states of a metal cannot be derived by perturbation theory even though

the electron-phonon interaction causing it is very small; the superconducting state is a new state that is not obtained by shifting the energy eigenstates of $\hat{H}_0$. Another example is the case when the spectrum of $\hat{H}_0$ is discrete (denumerable) and that of $\hat{H}$ is continuous.[1]

To derive perturbation theory in a general way, we need to do some preparatory work. Here is a vague description of what we need to do.

1. We introduce projection operators which allow us to divide the space of states into two subspaces orthogonal to each other, which we call the P-subspace and the Q-subspace. The P-subspace contains the states of interest (whose properties we want to calculate or measure) the Q-subspace contains all other the states.

2. The projection operators allow us to partition an eigenvalue equation into two equations, one focusing on the P-subspace and the other on the Q-subspace.

3. Perturbation theory is then developed by assuming that the part of the Hamiltonian that couples the P-subspace to the Q-subspace is small. In other words the existence of the Q-subspace changes the properties of the states in the P-subspace by a small amount.

[1] K. O. Friedrichs, *Perturbation of Spectra in Hilbert Space*, American Mathematical Society, Providence RI, 1965; F. Rellich, *Perturbation Theory of Eigenvalue Problems*, Gordon and Breach Science Publishers, New York, 1969; E. M. Harrell II, Perturbation theory and atomic resonances since Schrödinger's time, pp. 227-248 in: P. Deift, F. Gesztesy. P. Perry, and W. Schlag, eds., *Spectral Theory and Mathematical Physics: A Festschrift in Honor of Barry Simon's 60th Birthday*, Proceedings of Symposia in Pure Mathematics 76.1, American Mathematical Society, Providence RI, 2007.

4. To develop this formalism we make use of Green's functions (or the resolvent of an operator) which is a concept of great usefulness in all physics.

There are other ways of deriving perturbation theory that might seem simpler, such as assuming that the eigenstates and the eigenvalues of $\hat{H}$ are power series of a small, scalar, coupling parameter. The method used here is more general and it introduces you to concepts that are used throughout physical chemistry, physics or engineering. The price to pay is spending some time to prepare the ground.

Section 17.2. Projection Operators and the Partitioning of the Eigenvalue Problem

§ 2 *Projection operators.* The premise of perturbation theory for $\hat{H}$ is that we have solved the eigenvalue problem Eq. 2 for $\hat{H}_0$ and we know the eigenstates $|E_n\rangle$. It is therefore convenient to use the eigenstates $|E_n\rangle$, $n = 1, 2, \ldots$, as a basis set. We will use repeatedly the fact that these eigenstates are orthonormal

$$\langle E_i \,|\, E_j \rangle = \delta_{ij} \tag{4}$$

and complete

$$\sum_{j=1}^{\infty} |E_j\rangle\langle E_j| = \hat{I} \tag{5}$$

where $\hat{I}$ is the unit operator.

The addition of $\hat{V}$ to $\hat{H}_0$ will change the states $|E_n\rangle$ into the "perturbed states" $|\epsilon_n\rangle$ and the eigenvalues E_n into the perturbed eigenvalues ϵ_n. The

question we ask is how to calculate these changes if we assume that $\hat{V}$ is small.

We are rarely interested in how every one of the states $|E_n\rangle$ changes. For example, you might want to study the emission spectrum of the hydrogen atom and you want to know what happens to it in the presence of an electric field. In this example $\hat{H}_0$ is the Hamiltonian of the atom when the field is absent and its eigenfunctions $|n, \ell, m\rangle$ will be used as a basis set. Here $n = 1, 2, \ldots$ labels the energy; $\ell = 0, 1, \ldots, n-1$, the angular momentum squared; and $m = -\ell, -\ell+1, \ldots, \ell-1, \ell$, the projection of the angular momentum on the OZ axis. You may have your own reasons to focus on what happens to the emission frequency from the states $|2, 0, 0\rangle$, $|2, 1, -2\rangle$, $|2, 1, 0\rangle$, $|2, 1, 1\rangle$ to the lower energy state $|1, 0, 0\rangle$ when you turn the electric field on. In this case you want a theory that will focus on these states. However, one cannot ignore the fact that $\hat{H}_0$ has other states because the perturbation $\hat{V}$ caused by the electric field couples the states you are interested in to the other states. However, if $\hat{V}$ is small we can treat the effect of the other states approximately.

I emphasize that although we focus on a few states, there is no limitation on which states they are and there is no limitation on how many there are. Focusing on only a few states at a time is beneficial because this reduces the dimensionality of the problem. Focusing on one state at a time is best and this is possible if the difference between the eigenvalues E_n is large. However, if the states of interest are degenerate I need to focus on all of them. In the general theory this is an insignificant detail. However, when the theory is used it makes a big difference whether the state of interest is degenerate or

nearly degenerate with other states. This is why we have a separate lecture for the case when we are interested in degenerate or nearly degenerate states.

We can now get back to the general theory. Assume that the fate of certain eigenstates $\{|E_n\rangle\}_{n\in P}$ of $\hat{H}_0$ are of special interest to me (here P is a set of integers indexing the states of interest). I define the *projection operator*:

$$\hat{P} \equiv \sum_{j\in P} |E_j\rangle\langle E_j| \tag{6}$$

in which the sum is over all indices contained in the set P. When I act with this operator on an arbitrary ket $|\phi\rangle$ I obtain

$$\hat{P}|\phi\rangle = \sum_{j\in P} |E_j\rangle\langle E_j|\phi\rangle \tag{7}$$

While $|\phi\rangle$ is defined in the whole space $\hat{P}|\phi\rangle$ is a vector in a *subspace* generated by the restricted basis set $\{|E_n\rangle\}_{n\in P}$. In what follows I use the term P-subspace for all vectors of the form given by Eq. 7.

I also define the operator $\hat{Q}$

$$\hat{Q} \equiv \sum_{j\in Q} |E_j\rangle\langle E_j| \tag{8}$$

where the set Q contains all indices not included in P and does not contain any index that is in P (the sets P and Q are disjoint and together contain all integers). The operator $\hat{Q}$ acts on a ket $|\phi\rangle$ defined in the full space and produces a ket that is defined in the Q-subspace:

$$\hat{Q}|\phi\rangle = \sum_{n\in Q} |E_n\rangle\langle E_n|\phi\rangle \tag{9}$$

The full space generated by $\{|\psi_i\rangle\}_{i=1}^{\infty}$ consists of two *orthogonal subspaces*, one containing all states of the form $\sum_{j\in P} |\psi_j\rangle\langle\psi_j\,|\,\phi\rangle = \hat{P}|\phi\rangle$ and the other

containing all states of the form $\sum_{j\in Q}|\psi_j\rangle\langle\psi_j\,|\,\phi\rangle = \hat{Q}|\phi\rangle$. Each ket in the full space is the sum of a ket in the P-subspace and a ket in the Q-subspace.

In the example of the hydrogen atom in the electric field, which we discussed above, the projection operator will be

$$\hat{P} = |1,0,0\rangle\langle 1,0,0| + \sum_{m=-1}^{1}|2,1,m\rangle\langle 2,1,m| \tag{10}$$

The projection operator includes all the states you are interested in. Therefore you want to treat their interaction with electric field accurately but it might be tolerable to treat approximately how the other states (not included in $\hat{P}$) affect the ones that interest you.

We can now list a number of properties of $\hat{P}$ and $\hat{Q}$ that follow from the fact that the basis set $\{|E_n\rangle\}_{n=1}^{\infty}$ is complete and orthonormal and consists of the eigenstates of $\hat{H}_0$.

$$\hat{P}^2 = \hat{P} \tag{11}$$

$$\hat{Q}^2 = \hat{Q} \tag{12}$$

$$\hat{P} + \hat{Q} = \hat{I} \tag{13}$$

$$\hat{P}\hat{Q} = \hat{Q}\hat{P} = 0 \tag{14}$$

$$\hat{P}\hat{H}\hat{Q} = \hat{P}\hat{V}\hat{Q} \tag{15}$$

$$\hat{Q}\hat{H}\hat{P} = \hat{Q}\hat{V}\hat{P} \tag{16}$$

$$\hat{Q}\hat{H}_0 = \hat{H}_0\hat{Q} \tag{17}$$

$$\hat{P}\hat{H}_0 = \hat{H}_0\hat{P} \tag{18}$$

In addition, $\hat{P}$ and $\hat{Q}$ are Hermitian operators. We will use these properties repeatedly in what follows.

Exercise 1 Show that Eqs. 11–18 follow from the relation $\langle \psi_i \,|\, \psi_j \rangle = \delta_{ij}$ and from the fact that $\hat{H}_0|E_n\rangle = E_n|E_n\rangle$.

Exercise 2 A diatomic molecule collides harshly with an atom and its vibrational state becomes $|\phi\rangle$. Define a projection operator that picks up those components of $|\phi\rangle$ whose energy is higher than $\frac{5}{2}\hbar\omega$, where ω is the vibrational frequency.

---

§ 3 *Partitioning the Schrödinger equation.* I will use the projection operators $\hat{P}$ and $\hat{Q}$ to split the Schrödinger equation

$$\hat{H}|\varepsilon\rangle = \varepsilon|\varepsilon\rangle \tag{19}$$

into *two coupled* equations, one for $\hat{P}|\varepsilon\rangle$ and the other for $\hat{Q}|\varepsilon\rangle$. Insert $\hat{P} + \hat{Q} = \hat{I}$ into Eq. 19, to obtain

$$\hat{H}(\hat{P} + \hat{Q})|\varepsilon\rangle = \varepsilon(\hat{P} + \hat{Q})|\varepsilon\rangle \tag{20}$$

Act with $\hat{P}$ on this equation (recall that $\hat{P}^2 = \hat{P}$ and $\hat{P}\hat{Q} = 0$) and you obtain

$$\hat{P}\hat{H}\hat{P}|\varepsilon\rangle + \hat{P}\hat{H}\hat{Q}|\varepsilon\rangle = \varepsilon\hat{P}|\varepsilon\rangle \tag{21}$$

Acting with $\hat{Q}$ on Eq. 20 gives

$$\hat{Q}\hat{H}\hat{Q}|\varepsilon\rangle + \hat{Q}\hat{H}\hat{P}|\varepsilon\rangle = \varepsilon\hat{Q}|\varepsilon\rangle \tag{22}$$

Equations 21 and 22 are equivalent to the Schrödinger equation, Eq. 19.

Exercise 3 Derive Eq. 19 from Eqs. 21 and 22.

This kind of formal rewriting of an equation gives nothing new, since no information is added to what we know from Eq. 19. Occasionally, however, such manipulation affords us a different look at the problem and suggests new approximations. In our case it allows us to treat the effect of $\hat{V}$ on the states in the P-space accurately and the approximate the way in which the Q-states affect the P-states.

§ 4 *What is the meaning of $\hat{P}\hat{H}\hat{P}$?* It is easy to understand the meaning of $\hat{P}\hat{H}\hat{P}$ if we use the definition $\hat{P} = \sum_{j\in P} |\psi_j\rangle\langle\psi_j|$ given by Eq. 6. This leads to

$$\hat{P}\hat{H}\hat{P} = \sum_{i\in P}\sum_{j\in P} |\psi_i\rangle\langle\psi_i \,|\, \hat{H} \,|\, \psi_j\rangle\langle\psi_j| \tag{23}$$

On the other hand, the full Hamiltonian is (use Eq. 5)

$$\hat{H} = \hat{I}\hat{H}\hat{I} = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |\psi_i\rangle\langle\psi_i \,|\, \hat{H} \,|\, \psi_j\rangle\langle\psi_j| \tag{24}$$

When we compare these two equations, we see that $\hat{P}\hat{H}\hat{P}$ is a *restricted* version of the full Hamiltonian. It is constructed by *ignoring* the basis set functions $|\psi_i\rangle$ for which $i \in Q$. This restricted Hamiltonian is an approximate version of the full Hamiltonian. Similarly, the term $\hat{Q}\hat{H}\hat{Q}$ is the Hamiltonian restricted to the basis sub-set $\{|\psi_i\rangle\}_{i\in Q}$.

Another way of illustrating the same thing is to calculate the matrix element $\langle E_n|\hat{P}\hat{H}\hat{P}|E_m\rangle$ of $\hat{P}\hat{H}\hat{P}$. If the P-subspace is three-dimensional then $\langle E_n|\hat{P}\hat{H}\hat{P}|E_m\rangle$ is a 3×3 matrix. If the P-subspace is one-dimensional, the matrix is a number.

Let us look at how $\hat{P}\hat{H}\hat{P}$ acts on an arbitrary ket $|\psi\rangle = \hat{P}|\psi\rangle + \hat{Q}|\psi\rangle$. In $\hat{P}\hat{H}\hat{P}|\psi\rangle$, the first (rightmost) $\hat{P}$ obliterates the Q-components in $|\psi\rangle$. Then $\hat{H}$ acts on $\hat{P}|\psi\rangle$ and gives

$$\begin{aligned} \hat{H}\hat{P}|\psi\rangle &= \sum_{j=1}^{N}\sum_{i\in P} |E_j\rangle \langle E_j \,|\, \hat{H} \,|\, E_i\rangle \langle E_i \,|\, \psi\rangle \\ &= \sum_{j=1}^{N} |E_j\rangle \left[\sum_{i\in P} \langle E_j \,|\, \hat{H} \,|\, E_i\rangle \langle E_i \,|\, \psi\rangle\right] \end{aligned} \tag{25}$$

Note that $\hat{H}\hat{P}|\psi\rangle$ is a ket in the full space. Then the last (leftmost) $\hat{P}$ in $\hat{P}\hat{H}\hat{P}$ acts on that ket, chopping off the Q-component. The final result is a ket in the P subspace. The bottom line: $\hat{P}\hat{H}\hat{P}$ takes a vector $|\psi\rangle$ of the full space and returns a vector in the P subspace.

The term $\hat{P}\hat{H}\hat{Q}$ is interesting. Using the same method as above, you can show that

$$\hat{P}\hat{H}\hat{Q}|\psi\rangle = \sum_{j\in P} |E_j\rangle \left[\sum_{i\in Q} \langle E_j \,|\, \hat{H} \,|\, E_i\rangle \langle E_i \,|\, \psi\rangle\right] \tag{26}$$

$\hat{P}\hat{H}\hat{Q}$ takes a ket $|\psi\rangle$ represented in the full space, chops off its $\hat{P}$ part, acts with $\hat{H}$ to produce a ket in the full space, and then removes its components in the Q space. In Eq. 26 the Q-subspace components affect the components in the P subspace. We say that the term $\hat{P}\hat{H}\hat{Q}$ couples the Q subspace to the P subspace. If we were to neglect this term in Eq. 21, the resulting equation, $\hat{P}\hat{H}\hat{P}|\varepsilon\rangle = \varepsilon\hat{P}|\varepsilon\rangle$, would give eigenfunctions of the form $\hat{P}|\varepsilon\rangle = \sum_{j\in P} |\psi_j\rangle \langle\psi_j \,|\, \varepsilon\rangle$ and this is wrong unless the dimensionality of the P-subspace is very high; if this is the case we don't need perturbation theory since we can solve for the eigenvalue of $\hat{P}\hat{H}\hat{P}$ and get accurate results; it is this kind of massive calculation that perturbation theory is designed to avoid by keeping the dimensionality of the P-subspace small.

Section 17.3. Green's Functions (the Resolvent) and the Perturbation Expansion

§ 5 *Green's functions.* As a second step in our development of perturbation theory, you need to learn a few things about Green's functions.[2] They provide a method for rewriting differential equations and, sometimes, for solving them. Before explaining what they are, I remind you of the definition and a few properties of the *inverse* $\hat{A}^{-1}$ of an operator $\hat{A}$.

The operator $\hat{A}^{-1}$ is defined by the equation

$$\hat{A}^{-1}\hat{A}|\phi\rangle = \hat{A}\hat{A}^{-1}|\phi\rangle = |\phi\rangle \tag{27}$$

which must be satisfied for all states $|\phi\rangle$.

We can calculate the inverse of $\hat{A}$ if we know its eigenstates $|\lambda_\alpha\rangle$ and eigenvalues λ_α given by

$$\hat{A}|\lambda_\alpha\rangle = \lambda_\alpha|\lambda_\alpha\rangle, \quad \alpha = 1, 2, \ldots \tag{28}$$

The inverse operator $\hat{A}^{-1}$ is then

$$\hat{A}^{-1} = \sum_{\alpha=1}^{\infty} |\lambda_\alpha\rangle \frac{1}{\lambda_\alpha} \langle\lambda_\alpha| \tag{29}$$

[2] In 1828 Nottingam Review, published and advertisement which announced that for 7s.6d one can buy an "Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism" written by George Green. The ad did not say that Mr. Green was a baker who had one year of schooling when he was eight years old. At the age of forty, after the publication of the essay, he went to study undergraduate mathematics at Cambridge where he graduated fourth because of a "want of familiarity with ordinary boy's mathematics". The essay, which sold 61 copies, introduced the world to Green's functions and the two Green theorems from vector calculus.

If one of the eigenvalues of $\hat{A}$ is zero, *the operator does not have an inverse*, since division by zero is forbidden.

We have assumed here that the operator $\hat{A}$ either has only a discrete spectrum or that for the calculations that we plan to perform the continuous spectrum can be ignored. All but the most artificial model systems (i.e. a particle in a box with infinitely repulsive walls, or a harmonic oscillator) have a continuous spectrum, which corresponds to the fragmentation of the system into parts that fly away from each other. In the calculations that follow the continuous spectrum can be ignored because we deal with phenomena involving energies much lower than the one needed for fragmentation.

The Green's function of an operator $\hat{O}$ is defined by the expression[3]

$$G_{\hat{O}}(a) \equiv \left(a\hat{I} - \hat{O}\right)^{-1} \tag{30}$$

where a is a complex number with units of energy. Mathematicians call this the *resolvent* of the operator $\hat{O}$.

Entire books[4] have been dedicated to Green's functions since they are very useful in physics, chemistry, and engineering. Here we need one property: if

[3]The nomenclature is not uniform and different books may give definitions that appear very different. The definition given here is more abstract and more general than the definitions given for particular applications. For example, you might see definitions in which the Green's function is an integral operator; they turn out to be our definition expressed in coordinate representation.

[4]E. Butkov, *Mathematical Physics*, Addison-Wesley Co., Menlo Park, 1968; P. Morse and H. Feshback, *Methods of Theoretical Physics*, McGraw Hill, New York, 1953; G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, Elsevier, Amsterdam, 2005; I. Starkgold, *Green's Functions and Boundary Value Problems*, J. Wiley and Sons, New York, 1979; G. Barton, *Elements of Greens's Functions and Boundary Value Problems*, Clarendon Press, Oxford, 1989.

$\hat{H}$ is a Hermitian operator and E_i and $|\varepsilon_i\rangle$ satisfy the eigenvalue equation

$$\hat{H}|\varepsilon_i\rangle = E_i|\varepsilon_i\rangle, i = 1, 2, \ldots \tag{31}$$

where E_i are the discrete, isolated eigenvalues and if we can neglect the continuous spectrum, then the resolvent of $\hat{H}$ is

$$G_{\hat{H}}(E) \equiv \left(E\hat{I} - \hat{H}\right)^{-1} = \sum_{i=1}^{\infty} |\varepsilon_i\rangle \frac{1}{E - E_i} \langle\varepsilon_i| \tag{32}$$

This can be verified by direct calculation. Since

$$\hat{H} = \sum_{i=1}^{\infty} |\varepsilon_i\rangle E_i \langle\varepsilon_i|$$

we have (use the orthonormality and completeness conditions)

$$\begin{aligned}
\left(\sum_{i=1}^{\infty} |\varepsilon_i\rangle \frac{1}{E - E_i} \langle\varepsilon_i|\right)\left(E\hat{I} - \hat{H}\right) &= \left(\sum_{i=1}^{\infty} |\varepsilon_i\rangle \frac{1}{E - E_i} \langle\varepsilon_i|\right)\left(\sum_{j=1}^{\infty} |\varepsilon_j\rangle (E - E_j) \langle\varepsilon_j|\right) \\
&= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |\varepsilon_i\rangle \frac{1}{E - E_i}(E - E_j) \langle\varepsilon_i \,|\, \varepsilon_j\rangle \langle\varepsilon_j| \\
&= \sum_{i=1}^{\infty} |\varepsilon_i\rangle \langle\varepsilon_i| = \hat{I}
\end{aligned} \tag{33}$$

This derivation shows that Eq. 32 is correct.

Exercise 4 Show that

$$\left(E\hat{I} - \hat{H}\right)\left(\sum_{j=1}^{\infty} |\varepsilon_j\rangle \frac{1}{E - E_j} \langle\varepsilon_j|\right) = \hat{I} \tag{34}$$

Eq. 32 can be used only for values of E for which

$$E \neq E_j, \quad j = 1, 2, \ldots \tag{35}$$

If E is equal to one of the eigenvalues E_j, then the Green's function does not exist (the inverse $(E - \hat{H})^{-1}$ does not exist if E is equal to an eigenvalue).

§ 6 *A formula for the resolvent of* $\hat{H} = \hat{H}_0 + \hat{V}$. Very often a Hamiltonian $\hat{H}$ has the form

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{36}$$

where for now $\hat{H}_0$ and $\hat{V}$ are arbitrary operators. In this case the resolvent of $\hat{H}$ satisfies the equation

$$\hat{G}(E) = \hat{G}^0(E) + \hat{G}^0(E)\hat{V}\hat{G}(E) \tag{37}$$

where

$$\hat{G}(E) = \left(E\hat{I} - \hat{H}_0 - \hat{V}\right)^{-1} \equiv \left(E\hat{I} - \hat{H}\right)^{-1} \tag{38}$$

and

$$\hat{G}^0(E) = \left(E\hat{I} - \hat{H}_0\right)^{-1} \tag{39}$$

I verify below that Eq. 37 is correct. First I multiply Eq. 37 by $(E\hat{I} - \hat{H})$ on the right and obtain

$$\hat{G}(E)\left(E\hat{I} - \hat{H}\right) = \hat{G}^0(E)\left(E\hat{I} - \hat{H}\right) + \hat{G}^0\, V\, \hat{G}(E)\left(E\hat{I} - \hat{H}\right) \tag{40}$$

According to Eq. 38, $\hat{G}(E)(E\hat{I} - \hat{H}) = \hat{I}$ and using this in Eq. 40 gives

$$\hat{I} = \hat{G}^0(E)(E\hat{I} - \hat{H}) + \hat{G}^0(E)\hat{V}$$

Multiplying this by $E\hat{I} - \hat{H}_0$ on the left (remember that, in general, operators do not commute) gives

$$E\hat{I} - \hat{H}_0 = \left(E\hat{I} - \hat{H}\right) + \hat{V} = E\hat{I} - \hat{H}_0$$

(I have used $\hat{G}^0(E)(E\hat{I} - \hat{H}_0) = \hat{I}$.) This proves that Eq. 37 is an identity and therefore is correct.

Note however that Eq. 37 is possible only if E is *not* an eigenvalue of $\hat{H}$ or $\hat{H}_0$. If E is an eigenvalue of $\hat{H}_0$ then $\hat{G}^0(E)$ does not exist; if E is an eigenvalue of $\hat{H}$ then $\hat{G}(E)$ does not exist.

Exercise 5 Show that

$$\hat{G}(E) = \hat{G}^0(E) + \hat{G}(E)\hat{V}\hat{G}^0(E) \tag{41}$$

Exercise 6 Show that $\hat{G}(E) = (\hat{I} - \hat{G}^0(E)\hat{V})^{-1}\hat{G}_0 + \hat{A}$ where $\hat{A}$ satisfies $(\hat{I} - \hat{G}_0\hat{V})\hat{A} = 0$.

---

§ 7 *Perturbation theory for the resolvent.* Eq. 37 is an *implicit* formula for $\hat{G}(E)$. It is possible to solve it to obtain $\hat{G}(E)$, but this is done only in simple cases when the physics can be captured by using a small basis set.

Exercise 7 (a) Use Eq. 37 to obtain an explicit expression for $\hat{G}(E)$.

(b) Is that procedure simpler than diagonalizing $\hat{H}$ and using Eq. 32?

(c) Define a Hermitian matrix and calculate $\hat{G}(E)$ by the methods outlined in parts (a) and (b).

However, something interesting happens if I use Eq. 37 to replace $\hat{G}(E)$ in the right-hand side of Eq. 37:

$$\begin{aligned}\hat{G}(E) &= \hat{G}^0(E) + \hat{G}^0(E)\hat{V}\left[\hat{G}^0(E) + \hat{G}^0(E)\hat{V}\hat{G}(E)\right] \\ &= \hat{G}^0(E) + \hat{G}^0(E)\hat{V}\hat{G}^0(E) + \hat{G}^0(E)\hat{V}\hat{G}^0(E)\hat{V}\hat{G}(E)\end{aligned}$$

Unfortunately I still have $\hat{G}(E)$ in the right-hand side. I seem to turn the original expression into a more complicated one! But let us repeat this procedure over and over. The result is

$$\begin{aligned}\hat{G}(E) &= \hat{G}^0(E) + \hat{G}^0(E)\hat{V}\hat{G}^0(E) + \hat{G}^0(E)\hat{V}\hat{G}^0(E)\hat{V}\hat{G}^0(E) \\ &\quad + \hat{G}^0(E)\hat{V}\hat{G}^0(E)\hat{V}\hat{G}^0(E)\hat{V}\hat{G}^0(E) + \cdots \\ &= \hat{G}^0(E)\left[\hat{I} + \sum_{n=1}^{\infty}\left(\hat{V}\hat{G}^0(E)\right)^n\right] \qquad (42)\end{aligned}$$

Now $\hat{G}$ has disappeared from the right-hand side. Since $\hat{V}$ is known, if I know $\hat{G}^0$ then I can calculate $\hat{G}$. The price to pay is that I must evaluate an infinite number of terms! Can this be of any use?

Eq. 42 is a power series in $\hat{V}\hat{G}^0$. If $\hat{V}\hat{G}^0$ is small[5] this series may be convergent; moreover, I might *retain a few terms* and still get an accurate representation of $\hat{G}(E)$ in terms of $\hat{G}^0$ and $\hat{V}$(which we assume we know). Perturbation theory is the approximation that keeps a few terms in Eq. 42. An an example of using this expansion for three-dimensional matrices is given in the `Mathematica` file `resolvent.nb`.

[5]We will discuss later what this means.

Section 17.4. Combine Partitioning and Green's Functions

§ 8 *The problem.* Now that we are familiar with partitioning and with Green's functions let us return to perturbation theory. We want to solve the eigenvalue problem

$$\hat{H}|\varepsilon_i\rangle = \varepsilon_i|\varepsilon_i\rangle \tag{43}$$

where

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{44}$$

We know all the eigenvalues E_i and the eigenfunctions $|E_i\rangle$ of $\hat{H}_0$:

$$\hat{H}_0|E_i\rangle = E_i|E_i\rangle \tag{45}$$

Our physical intuition, or some experiments, tell us that "turning on" $\hat{V}$ changes the spectrum of $\hat{H}_0$ gradually into the spectrum of $\hat{H}$. This means that we have a one-to-one correspondence between ε_i and E_i and between $|\varepsilon_i\rangle$ and $|E_i\rangle$. Moreover, if $\hat{V}$ becomes smaller and smaller, $|\varepsilon_i\rangle$ and ε_i get closer and closer to $|E_i\rangle$ and E_i.

For example, $\hat{V}$ might be the interaction between the charges in a molecule and an external electric or magnetic field. The energy eigenvalues of the molecule change as the magnitude of the external field is changed (ε_i will differ from E_i). This means that the frequencies observed in the absorption and emission spectra of the molecule will change. In the absence of the field, they are

$$\omega_{ij} = \frac{|E_i - E_j|}{\hbar} \tag{46}$$

and when the field is turned on, they change to

$$\bar{\omega}_{ij} = \frac{|\varepsilon_i - \varepsilon_j|}{\hbar} \tag{47}$$

The difference $\omega_{ij} - \bar{\omega}_{ij}$ is called the *shift* caused by the external field. In some cases the effect of the field on the spectrum is richer: some of the peaks in the spectrum of the undisturbed system split into multiplets. This does not mean that the interaction with the field creates new states. Those states existed when the field was absent, but they were *degenerate* and gave the same peak in the spectrum. The field shifts the energies of the degenerate states unevenly and they produced several peaks at frequencies close to that of the original peak.

Our purpose here is to use the theoretical machinery developed so far to calculate how the eigenvalues and the eigenstates of an operator are changed by a *small* perturbation $\hat{V}$. In principle, we can do that by solving the eigenvalue problem $(\hat{H}_0 + \hat{V})|\varepsilon\rangle = \varepsilon|\varepsilon\rangle$. This can be done but it is laborious and we pursue here a different route. We take advantage of the smallness of $\hat{V}$ to express $|\varepsilon_i\rangle$ and ε_i in terms of $\hat{V}$ and the eigenstates and eigenvalues of $\hat{H}_0$. The method is useful only for those systems for which we know $|E_i\rangle$ and E_i.

§ 9 *Collect some of the results needed.* The method described in §3 partitioned the eigenvalue problem for $\hat{H}$ into (see Eqs. 21 and 22)

$$\hat{P}\hat{H}\hat{P}|\varepsilon\rangle + \hat{P}\hat{H}\hat{Q}|\varepsilon\rangle = \varepsilon\hat{P}|\varepsilon\rangle \tag{48}$$

and

$$\hat{Q}\hat{H}\hat{Q}|\varepsilon\rangle + \hat{Q}\hat{H}\hat{P}|\varepsilon\rangle = \varepsilon\hat{Q}|\varepsilon\rangle \tag{49}$$

We have chosen $|E_i\rangle$ so that $|E_i\rangle$ satisfies $\hat{H}_0|E_i\rangle = E_i|E_i\rangle$ and $\langle E_i \,|\, E_j\rangle = \delta_{ij}$, and this leads to some simplifications:

$$\hat{Q}\hat{H}\hat{P} \;=\; \hat{Q}\hat{V}\hat{P} \tag{50}$$

$$\hat{P}\hat{H}\hat{Q} \;=\; \hat{P}\hat{V}\hat{Q} \tag{51}$$

These equations are correct because $\hat{Q}\hat{H}_0\hat{P} = \hat{P}\hat{H}_0\hat{Q} = 0$. Therefore, we can rewrite Eqs. 48 and 49 as

$$\left(\varepsilon\hat{P} - \hat{P}\hat{H}_0\hat{P}\right)\hat{P}|\varepsilon\rangle = \hat{P}\hat{V}\hat{P}|\varepsilon\rangle + \hat{P}\hat{V}\hat{Q}|\varepsilon\rangle \tag{52}$$

and

$$\left(\varepsilon\hat{Q} - \hat{Q}\hat{H}\hat{Q}\right)\hat{Q}|\varepsilon\rangle = \hat{Q}\hat{V}\hat{P}|\varepsilon\rangle \tag{53}$$

Exercise 8 Derive Eqs. 50 and 51.

---

§ 10 *Use the resolvent to rewrite Eqs. 52 and 53.* I want to use the Green's function method to solve Eq.53 for $\hat{Q}|\epsilon\rangle$ and use the expression obtained in this way to eliminate $\hat{Q}|\epsilon\rangle$ for Eq. 52. The result is an exact equation for $\hat{P}|\epsilon\rangle$. To do this I define the Green's function of $\hat{Q}\hat{H}\hat{Q}$ to be

$$\hat{G}_Q(\varepsilon) = \hat{Q}\left(\varepsilon\hat{Q} - \hat{Q}\hat{H}\hat{Q}\right)^{-1}\hat{Q} \tag{54}$$

The first and last $\hat{Q}$ in the right-hand side are added to make sure that the inverse operator acts only on the states $\{|E_i\rangle\}_{i\in Q}$. They are superfluous because $(\varepsilon\hat{Q} - \hat{Q}\hat{H}\hat{Q})^{-1}$ is defined to act on the Q-subspace only.

Exercise 9 (a) Prove that $\hat{Q}\hat{H}\hat{Q}$ has zero eigenvalues in the P-subspace: if $|\phi\rangle$ belongs to the P-subspace then $\hat{Q}\hat{H}\hat{Q}|\phi\rangle = 0$.

(b) Prove that the elements of the matrix of $\hat{Q}\hat{H}_0\hat{Q}$ in the basis set $\{|E_i\rangle\}_{i=1}^{\infty}$ satisfy: $(\hat{Q}\hat{H}_0\hat{Q})_{ij} = 0$ if either i or j is in P.

Applying $\hat{G}_Q(\varepsilon)$ at the left of Eq. 53 gives

$$\hat{Q}|\varepsilon\rangle = \hat{G}_Q(\varepsilon)\hat{Q}\hat{V}\hat{P}|\varepsilon\rangle \tag{55}$$

Now use Eq. 55 to eliminate $\hat{Q}\epsilon\rangle$ from Eq. 52 and obtain

$$\left(\varepsilon\hat{P} - \hat{P}\hat{H}_0\hat{P}\right)\hat{P}|\varepsilon\rangle = \hat{P}\hat{V}\hat{P}|\varepsilon\rangle + \hat{P}\hat{V}\hat{Q}\hat{G}_Q(\varepsilon)\hat{Q}\hat{V}\hat{P}|\varepsilon\rangle \tag{56}$$

This equation no longer contains $\hat{Q}|\varepsilon\rangle$. This does not mean that the influence of the states $\{|E_i\rangle\}_{i\in Q}$ is eliminated (that would be wrong). They influence $\hat{P}|\varepsilon\rangle$ through the term $\hat{P}\hat{V}\hat{Q}\hat{G}_Q(\varepsilon)\hat{Q}\hat{V}\hat{P}|\varepsilon\rangle$. Note that if $\hat{V} = 0$ that influence vanishes.

Eq. 56 contains several parts. The term $\left(\varepsilon\hat{I} - \hat{P}\hat{H}_0\hat{P}\right)\hat{P}|\varepsilon\rangle = \hat{P}\hat{V}\hat{P}|\varepsilon\rangle$ is the eigenvalue equation $\hat{H}|\epsilon\rangle = \epsilon|\epsilon\rangle$ *restricted to the P-subspace*. Eq. 56 is more than just this restricted equation because the presence of the term $\hat{P}\hat{V}\hat{Q}\hat{G}_Q(\varepsilon)\hat{Q}\hat{V}\hat{P}|\varepsilon\rangle$ gives exactly how the existence of the Q-space modifies the states in the P-subspace. All these manipulations are exact as long as $\hat{G}_Q(\varepsilon)$ exists (i.e. $\varepsilon\hat{Q} - \hat{Q}\hat{H}\hat{Q}$ has an inverse).

§ 11 *Expand the resolvent in powers of $\hat{V}$.* We are now ready to profit from the smallness of $\hat{V}$. We start from Eq. 42 for $\hat{G}_Q(\varepsilon)$:

$$\hat{G}_Q(\varepsilon) = \hat{G}^0_Q(\varepsilon) + \hat{G}^0_Q(\varepsilon)\sum_{n=1}^{\infty}\left[\hat{Q}\hat{V}\hat{Q}\hat{G}^0_Q(\varepsilon)\right]^n \tag{57}$$

where

$$\hat{G}^0_Q(\varepsilon) = \hat{Q}\left(\varepsilon\hat{Q} - \hat{Q}\hat{H}_0\hat{Q}\right)^{-1}\hat{Q} \tag{58}$$

This expansion can be used only if the sum is convergent, which means that $\hat{G}_Q^0\hat{V}$ must be small. We will establish later what 'small' means.

By deriving Eq. 57 we have achieved quite a bit: we have expressed $\hat{G}_Q(\varepsilon)$, which is difficult to calculate because it involves the full Hamiltonian $\hat{H}$, in terms of $\hat{G}_Q^0(\varepsilon)$ which we can calculate because we know the eigenstates and eigenvalues of $\hat{H}^0$.

I am not interested in anything happening in the Q-subspace except for knowing how the Q-subspace affects what is happening with the states in the P-subspace. Let me repeat here Eq. 56, which describes how the states in the P-subspace depend on the states on the Q-subspace:

$$\left(\varepsilon\hat{P} - \hat{P}\hat{H}_0\hat{P}\right)\hat{P}|\varepsilon\rangle = \hat{P}\hat{V}\hat{P}|\varepsilon\rangle + \hat{P}\hat{V}\hat{Q}\hat{G}_Q(\varepsilon)\hat{Q}\hat{V}\hat{P}|\varepsilon\rangle \tag{59}$$

The trouble with this exact equation is practical: the equation contains $\hat{G}_Q(\varepsilon)$ which depends on the full Hamiltonian $\hat{H}$. Equation 57 expresses the troublesome $\hat{G}_Q(\varepsilon)$ with an expression that depends only on $\hat{G}_Q^0(\varepsilon)$, which I know how to compute because it is defined in terms of the eigenstates of $\hat{H}_0$ which I am supposed to know. Therefore I replace $\hat{G}_Q(\varepsilon)$ in Eq. 59 with the expression given by Eq. 57. The result is

$$\begin{aligned}\left(\varepsilon\hat{P} - \hat{P}\hat{H}_0\hat{P}\right)\hat{P}|\varepsilon\rangle &= \hat{P}\hat{V}\hat{P}|\varepsilon\rangle + \hat{P}\hat{V}\hat{G}_Q(\varepsilon)\hat{V}\hat{P}|\varepsilon\rangle \\ &= \hat{P}\hat{V}\hat{P}|\varepsilon\rangle + \hat{P}\hat{V}\hat{G}_Q^0(\varepsilon)\sum_{n=0}^{\infty}\left[\hat{V}\hat{G}_Q^0(\varepsilon)\right]^n\hat{V}\hat{P}|\varepsilon\rangle \end{aligned} \tag{60}$$

From now on, we will work with Eq. 60. Note that this equation has a serious problem: its right-hand side depends on $\hat{P}|\epsilon\rangle$ which I don't know. A further approximation is needed if we want to use it to calculated $\hat{P}|\epsilon\rangle$. This will

be done in the next two lectures: one deals with the perturbation of a state that is not degenerate; the other with the case when we are interested in the perturbation of a group of degenerate or nearly degenerate states.