Lecture 33

Relevant sections in text: §5.1

Approximation methods

We now begin studying various approximation methods in quantum mechanics. Approximation methods have a practical and a conceptual value. On the practical side, because all but the simplest systems resist attempts to find closed form analytical solutions to eigenvalue problems, the Schrödinger equation, *etc. ,* we use such methods to get useful, approximations to wave functions, energies, spectra, as well as transition probabilities and other dynamical quantities. On the conceptual side we shall see that some of our most cherished ways of thinking about energy levels and dynamics stem principally from the point of view of approximation methods.

The need for approximation methods arises from the simple fact that almost all realistic physical systems one wants to study are too complicated for explicit analytic solutions to be available. (This is no less true in classical mechanics.) So, for example, while we can analytically handle the hydrogen atom (when modeled as a charged particle in a Coulomb field), we cannot handle helium or more complicated atoms in the same fashion – let alone dynamical processes involving the interaction of these atoms with electromagnetic fields. In fact, even more realistic models of the hydrogen atom (including things like spin-orbit coupling, hyperfine interaction, finite size of nucleus, *etc.*) are not exactly soluble. Thus the *only* way we can understand these systems is to find methods of approximation.

We shall study two of several possible approximation techniques. First we shall look at what is usually called "time independent perturbation theory" (TIPT), which gives approximate solutions to eigenvalue problems. But this is also called "stationary state perturbation theory" (since one is usually studying the eigenvalue problem for the Hamiltonian). Then we shall study time-dependent perturbation theory (TDPT), which is designed to give approximate solutions to the (time-dependent) Schrödinger equation.

For the most part I am going to explain the results of the theory with very little in the way of derivations. Then we will look at some important applications.

Time independent perturbation theory

This approximation method is designed to approximate the eigenvalues and eigenvectors of a given observable. This observable is usually the Hamiltonian (whence the alternate name "stationary state perturbation theory"), but the techniques and results are not restricted to just the Hamiltonian; any observable will do.

The basic idea is that one is attempting to view a given observable of interest as in

some sense "close" to a simpler, well-understood observable. One then approximates the eigenvalues and eigenvectors of the given observable in terms of those of the simpler observable. For example, one could be interested in the energies of an anharmonic oscillator, with Hamiltonian

$$H = \frac{P^2}{2m} + \frac{1}{2}m\omega^2 X^2 + \beta X^4.$$

Assuming that the anharmonicity (described by β) is suitably "small", one can usefully approximate the eigenvalues and eigenvectors of H in terms of those of the harmonic oscillator. Let us now make this more precise.

We suppose that the observable of interest H admits eigenvalues and can be expressed as

$$H = H_0 + V,$$

where V is a small "perturbation" of H_0, *e.g.,* its matrix elements in the basis of eigenvectors of H_0 are small compared to the eigenvalues of H_0. For simplicity, we assume that H_0 has discrete spectrum. We assume that all the operators in question are sufficiently well behaved such that the eigenvectors and eigenvalues of H can be obtained via a 1-parameter family of operators* beginning with H_0:

$$H(\lambda) = H_0 + \lambda V, \quad H(0) = H_0, \quad H(1) = H.$$

For each value of λ we have

$$H(\lambda)|E(\lambda)\rangle = E(\lambda)|E(\lambda)\rangle.$$

The idea is that if the effect of V is small compared to that of H_0 then one should be able to approximate the eigenvectors and eigenvalues of H by expanding them in a power series in λ, keeping just the first term or so, and then evaluating at $\lambda = 1$. This is equivalent to approximating the eigenvectors and eigenvalues using an expansion in powers of the matrix elements of the potential (which can be taken to be λV). Thus we assume that

$$E_n(\lambda) = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots,$$

$$|E_n(\lambda)\rangle = |E_n\rangle^{(0)} + \lambda|E_n\rangle^{(1)} + \lambda^2|E_n\rangle^{(2)} + \cdots.$$

The plan is to solve the equation

$$(H_0+\lambda V)(|E_n\rangle^{(0)} + \lambda|E_n\rangle^{(1)} + \lambda^2|E_n\rangle^{(2)} + \cdots)$$
$$= (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots)(|E_n\rangle^{(0)} + \lambda|E_n\rangle^{(1)} + \lambda^2|E_n\rangle^{(2)} + \cdots)$$

---

* The parameter λ is not essential; it is just a convenient means of bookkeeping, as you will see.

order by order in λ to derive the corresponding perturbative approximations to the eigenvectors and eigenvalues. Let us note that eigenvectors are only determined up to multiplication by a scalar. This is, of course, what allows us to normalize them and view them as state vectors. Therefore, *when we solve the eigenvalue problem perturbatively we will still need to normalize the result.*

At zeroth order, it is easy to see by inspection that the relevant equation is

$$H_0|E_n\rangle^{(0)} = E_n^{(0)}|E_n\rangle^{(0)},$$

which just says that the zeroth vector and scalar are the unperturbed eigenvector and eigenvalue. This should come as no surprise. The higher order equations constitute a "triangular" system of linear, inhomogeneous equations. Each set of equations depends only upon the results of the lower order equations. So one can successively solve these equations from lower to higher order. In the end, any perturbative correction can be expressed completely in terms of the unperturbed eigenvalues and eigenvectors. Of course, the idea is that for a sufficiently small perturbation a good approximation can be obtained by just sticking to relatively low orders. The details of the solution process (which constitutes a very nice application of linear algebra techniques) can be found in your text. Here I will just state the results for first-order perturbation theory ($\mathcal{O}(\lambda)$ corrections) and show you how to use them.

As it turns out, the solutions of the first-order equations are simplest when the unperturbed ($\lambda = 0$) eigenvalue is non-degenerate. We begin with that case. *If $E_n^{(0)}$ is a non-degenerate eigenvalue of H_0,* then we have

$$E_n^{(1)} = V_{nn} \equiv {}^{(0)}\langle E_n|V|E_n\rangle^{(0)},$$

and

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

where

$$V_{kn} = {}^{(0)}\langle E_k|V|E_n\rangle^{(0)}.$$

The first-order approximation to the energy eigenvalue is then

$$E_n \approx E_n^{(0)} + V_{nn}.$$

The *un-normalized* eigenvector is, to first-order,

$$|E_n\rangle \approx |E_n\rangle^{(0)} + \sum_{k \neq n} |E_k\rangle^{(0)} \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}}.$$

3

Thus the correct eigenvector is a superposition of the old eigenvectors, which form a basis.

If we wish to approximate the state of the system using this vector, we must normalize it. It is a straightforward exercise to see that, to first-order in perturbation theory,

$$|E_n\rangle_{normalized} = \sqrt{Z_n}|E_n\rangle,$$

where

$$Z_n = 1 - \sum_{k \neq n} \frac{|V_{kn}|^2}{(E_n^{(0)} - E_k^{(0)})^2}.$$

You can easily see that Z_n can be interpreted as the probability for getting $E_n^{(0)}$ when measuring H_0 in the state where H is known to have the value E_n with certainty. Put differently, we can say that Z_n is the probability for finding the state $|E_n\rangle^{(0)}$ when the system is in the state $|E_n\rangle$. In general, assuming that at least one of the $V_{kn} \neq 0$, we have that $Z_n < 1$.