Introduction to Perturbation Theory

Lecture 31

Physics 342
Quantum Mechanics I

Monday, April 21st, 2008

The program of time-independent quantum mechanics is straightforward – given a potential $V(x)$ (in one dimension, say), solve

$$-\frac{\hbar^2}{2\,m}\,\psi'' + V(x)\,\psi = E\,\psi, \tag{31.1}$$

for the eigenstates. These form a complete, orthogonal basis for all functions. Using this adapted basis, generate generic initial configurations and time evolve them according to

$$\begin{aligned} \Psi(x,t) &= \sum_{j=1}^{\infty} \alpha_j\,\psi_j(x)\,e^{-i\,\frac{E_j}{\hbar}\,t} \\ \alpha_j &= \Psi(x,0)\cdot\psi_j. \end{aligned} \tag{31.2}$$

At any given time, we have the probability density for the system, and can calculate various physically measurable properties. If we actually perform a measurement, the wavefunction takes on one of the eigenstates (of the operator associated with the measurement), and returns the value of the physical measurement for that state (part of our assumption).

The problem, as we have seen, is that solving (31.1) for all but the simplest potentials can be difficult. We turn now to the problem of approximating solutions – our first (and only, at this stage) tool will be perturbation theory. The technique is appropriate when we have a potential $V(x)$ that is closely related to a "simple" (read "solvable") potential $\bar{V}(x)$.

31.1 Perturbation – Polynomials

Before working on a full ODE like the time-independent Schrödinger equation, let's get the basic arguments down for a polynomial equation, where

some of the issues are simplified.

31.1.1 Distinct Roots

Consider the roots of the polynomial

$$a\,x^2 + \epsilon\,x + c = 0, \tag{31.3}$$

we know the solution here, just the quadratic formula

$$x = \frac{-\epsilon \pm \sqrt{\epsilon^2 - 4\,a\,c}}{2\,a}. \tag{31.4}$$

But suppose we didn't have/remember this. Further, suppose ϵ is itself a small parameter, so that the form of (31.3) is close to easily the solvable equation:

$$a\,x^2 + c = 0 \tag{31.5}$$

which has roots: $x = \pm i\,\sqrt{\frac{c}{a}}$.

The idea behind perturbation theory is to attempt to solve (31.3), given the solution to (31.5). Operationally, we take an ansatz for x:

$$x = x_0 + \epsilon\,x_1 + \epsilon^2\,x_2 + \ldots, \tag{31.6}$$

and insert that into (31.3). Note that an implicit assumption we are making here is that the coefficients a and c are order one, and that x itself is order one (meaning that these quantities do not scale with ϵ). Inputting our form gives:

$$\left(a\,x_0^2 + c\right) + \epsilon\,(x_0 + 2\,a\,x_0\,x_1) + \epsilon^2\left(x_1 + a\,x_1^2 + 2\,x_0\,x_2\right) + O(\epsilon^3) = 0. \tag{31.7}$$

Now for the simplifying trick – we assume that the terms of different order in ϵ do not talk to each other, that each order in ϵ must vanish separately – $5 + \epsilon = 0$ implies that $\epsilon = -5$ which is not small compared to 5. Using the independence of order, the above gives us three equations that we can solve separately to determine the set $\{x_0, x_1, x_2\}$:

$$\begin{aligned} \epsilon^0 &: \; a\,x_0^2 + c = 0 \\ \epsilon^1 &: \; x_0 + 2\,a\,x_0\,x_1 = 0 \\ \epsilon^2 &: \; x_1 + a\,x_1^2 + 2\,a\,x_0\,x_2 = 0. \end{aligned} \tag{31.8}$$

We can see how the $\epsilon = 0$ equation (31.5) plays a role here, it is the ϵ^0 equation that starts off the process by allowing us to solve for x_0. Notice the cascade here, knowing $x_0 = \pm\, i\sqrt{\frac{c}{a}}$, we can solve for x_1 (we don't actually need x_0 to find x_1 in the current case, but in general, we have a hierarchy of equations and perturbative dependence):

$$\pm i\sqrt{\frac{c}{a}}\,(1 + 2\,a\,x_1) = 0 \longrightarrow x_1 = -\frac{1}{2\,a}, \tag{31.9}$$

and by knowing x_0 and x_1, we can find x_2:

$$-\frac{1}{2\,a} + \frac{1}{4\,a} \pm 2\,i\,a\sqrt{\frac{c}{a}}\,x_2 = 0 \longrightarrow x_2 = \mp\frac{i}{8\,a^2\sqrt{\frac{c}{a}}}. \tag{31.10}$$

Putting it all together, we have, through order ϵ^2, the solution:

$$\boxed{x = x_0 + \epsilon\,x_1 + \epsilon^2\,x_2 = \pm i\sqrt{\frac{c}{a}} - \frac{\epsilon}{2\,a} \mp \frac{i\,\epsilon^2}{8\,a^2\sqrt{\frac{c}{a}}}.} \tag{31.11}$$

Here it is easy to compare with (31.4): Expanded in powers of ϵ via Taylor series (31.4) is

$$\begin{aligned} x = \frac{1}{2\,a}\left(-\epsilon \pm \sqrt{4\,a\,c}\,i\sqrt{1 - \frac{e^2}{4\,a\,c}}\right) &\sim \frac{1}{2\,a}\left(-\epsilon \pm i\sqrt{4\,a\,c}\left(1 - \frac{\epsilon^2}{8\,a\,c}\right)\right) \\ &= \pm\, i\sqrt{\frac{c}{a}} - \frac{\epsilon}{2\,a} \mp \frac{i\,\epsilon^2}{8\,a^2\sqrt{\frac{c}{a}}}. \end{aligned} \tag{31.12}$$

Here everything works out well. In some cases, we need to be more careful[1].

31.1.2 Degenerate Roots

In the preceding example, the roots of the "unperturbed" ($\epsilon = 0$) equation where separate, and we were effectively calculating corrections to each root separately. There is another case that has an analogue in our quantum mechanical calculations – suppose we had, for the unperturbed equation, the quadratic polynomial with degenerate roots:

$$x^2 - 2\,x + 1 = 0, \tag{31.13}$$

[1] Think of what happens when we have $\epsilon\,x^2 + b\,x + c = 0$, for example.

for which $x = 1$. Now if we introduce a perturbation, and define the perturbed equation via:

$$x^2 - 2\,x + (1 + \epsilon) = 0, \tag{31.14}$$

and take, again, $x = x_0 + \epsilon\, x_1$ (this time, we will find only the first correction), then:

$$0 = \left(x_0^2 - 2\,x_0 + 1\right) + \epsilon\,(2\,x_0\,x_1 - 2\,x_1 - 1) + O(\epsilon^2) \tag{31.15}$$

and our ϵ^0 equation reproduces (31.13), with $x_0 = 1$. But now, we have no way to satisfy the ϵ equation, which becomes $-1 = 0$. What happened? Think of the assumption we've made – we want x to be order unity, with corrections coming at order ϵ, but this ignores the quadratic term, which makes no contribution to ϵ order, we have, apparently, gone out too far in ϵ without taking into account potential corrective terms. Suppose we start, then, with

$$x = x_0 + \sqrt{\epsilon}\, x_1. \tag{31.16}$$

Then the perturbed equation becomes:

$$\left(x_0^2 - 2\,x_0 + 1\right) + \sqrt{\epsilon}\, x_1\,(2\,x_0 - 2) + \epsilon\left(1 + x_1^2\right) = 0 \tag{31.17}$$

and now the original solution $x_0 = 1$ satisfies the $\sqrt{\epsilon}$ equation, and we can move on to the ϵ equation, where we learn that $x_1 = \pm i$, giving us two perturbed solutions:

$$x_0 = 1 \pm i\,\sqrt{\epsilon}, \tag{31.18}$$

and splitting the degenerate root structure of the original equation (31.13).

31.2 Perturbation for ODEs

The same approach will work for ODEs, with similar caveats. Take the unperturbed equation:

$$\ddot{x}(t) + x(t) = 0 \quad x(0) = A \quad \dot{x}(0) = 0, \tag{31.19}$$

a harmonic oscillator that starts from rest. The solution is $x(t) = A\,\cos(t)$. Now suppose we want to solve

$$\ddot{x} + x - \epsilon\, x = 0 \quad x(0) = A \quad \dot{x}(0) = 0, \tag{31.20}$$

corresponding to a small frequency shift. We make the usual ansatz:

$$x(t) = x_0(t) + \epsilon\, x_1(t) \tag{31.21}$$

in order to generate the first order corrections. Then the ODE becomes:

$$(\ddot{x}_0 + x_0) + \epsilon\,(\ddot{x}_1 + x_1 - x_0) + O(\epsilon^2) = 0. \tag{31.22}$$

The solution to the ϵ^0 equation is just $x(t) = A\,\cos(t)$ as in the unperturbed case. Our ϵ^1 equation reads:

$$\ddot{x}_1 + x_1 - x_0 = 0 \tag{31.23}$$

which looks like a driven harmonic oscillator, with driving force $A\,\cos(t)$. We know how to solve this equation in general, but what should we do about the boundary conditions? In this case, the full boundary conditions are satisfied by the $x_0(t)$ solution, so we must have: $x_1(0) = 0$ and $\dot{x}_1(0) = 0$. Then the solution to the ϵ^1 ODE is:

$$x_1 = \frac{1}{2}\,A\,x_0\,t\,\sin(t) \tag{31.24}$$

and our full solution is

$$x(t) = A\,\cos(t) + \frac{1}{2}\,A\,\epsilon\,\sin(t)\,t \tag{31.25}$$

We can compare this with the Taylor expansion of the exact solution in this case:

$$A\,\cos\!\left(\sqrt{1-\epsilon}\,t\right) \sim A\,\cos(t) + \frac{1}{2}\,\epsilon\,A\,\sin(t)\,t. \tag{31.26}$$

So to order ϵ, we have the correct answer.

31.3 Perturbation for Eigenvalue Problem

We have seen how perturbation theory works, and what we need to do to get ODE solutions, the final element we need to consider to approach Schrödinger's equation perturbatively is to look at the perturbation of the eigenvalue equation itself. The twist is that we are looking for both eigenvectors and eigenvalues, and it is easiest to see how this will work out in the finite matrix case.

Take a symmetric real matrix, $\mathbb{A} = \mathbb{A}^T$ (so that we know the eigenvectors are complete and can be made orthonormal) in $\mathbb{R}^{N\times N}$. Suppose we know the eigenvectors and eigenvalues of this matrix:

$$\mathbb{A}\,\mathbf{x}_i = \lambda_i\,\mathbf{x}_i, \tag{31.27}$$

for $i = 1 \to N$, and we know that each eigenvalue has a distinct eigenvector (no degeneracy). We have constructed our eigenvectors so as to be orthonormal: $\mathbf{x}_i \cdot \mathbf{x}_j = \delta_{ij}$. Now we make a small perturbation to the matrix, $\mathbb{A} \longrightarrow \mathbb{A} + \epsilon\,\bar{\mathbb{A}}$, and we want to know how the eigenvalues and eigenvectors change under this perturbation. So introduce $\bar{\lambda}_i$ and $\bar{\mathbf{x}}$, that are themselves order ϵ corrections:

$$\mathbf{x}_i \longrightarrow \mathbf{x}_i + \epsilon\,\bar{\mathbf{x}}_i \qquad \lambda_i \longrightarrow \lambda_i + \epsilon\,\bar{\lambda}_i, \tag{31.28}$$

then the eigenvalue problem reads:

$$\left(\mathbb{A} + \epsilon\,\bar{\mathbb{A}}\right)\left(\mathbf{x}_i + \epsilon\,\bar{\mathbf{x}}_i\right) = \left(\lambda_i + \epsilon\,\bar{\lambda}_i\right)\left(\mathbf{x}_i + \epsilon\,\bar{\mathbf{x}}_i\right). \tag{31.29}$$

Expanding this, and keeping only those terms of order ϵ, we have:

$$\mathbb{A}\,\bar{\mathbf{x}}_i + \bar{\mathbb{A}}\,\mathbf{x}_i = \lambda_i\,\bar{\mathbf{x}}_i + \bar{\lambda}_i\,\mathbf{x}_i. \tag{31.30}$$

This seems like it will not be enough to determine *both* $\bar{\lambda}_i$ and $\bar{\mathbf{x}}_i$. But wait, since $\bar{\mathbf{x}}_i \in \mathbb{R}^N$, it has a decomposition in terms of the set $\{\mathbf{x}_i\}_{i=1}^N$:

$$\bar{\mathbf{x}}_i = \sum_{j=1}^{N} \alpha_j\,\mathbf{x}_j, \tag{31.31}$$

so that we can rewrite the first order equation as

$$\sum_{j=1}^{N} \alpha_j\,\lambda_j\,\mathbf{x}_j + \bar{\mathbb{A}}\,\mathbf{x}_i = \lambda_i \sum_{j=1}^{N} \alpha_j\,\mathbf{x}_j + \bar{\lambda}_i\,\mathbf{x}_i. \tag{31.32}$$

Take the dot product of this equation with $\mathbf{x}_i$, keeping in mind the presumed orthonormality:

$$\alpha_i\,\lambda_i + \mathbf{x}_i^T\,\bar{\mathbb{A}}\,\mathbf{x}_i = \lambda_i\,\alpha_i + \bar{\lambda}_i, \tag{31.33}$$

and we see that we can solve the above for $\bar{\lambda}_i$:

$$\boxed{\bar{\lambda}_i = \mathbf{x}_i^T\,\bar{\mathbb{A}}\,\mathbf{x}_i.} \tag{31.34}$$

Now for the eigenvector, we want to find, effectively, all of the coefficients α_j in the decomposition (31.31). Returning to the equation (31.32), we can take the dot-product w.r.t. *all* $\mathbf{x}_k$, not just the ith one – multiplying both sides by $\mathbf{x}_k^T$ gives

$$\alpha_k\,\lambda_k + \mathbf{x}_k^T\,\bar{\mathbb{A}}\,\mathbf{x}_i = \lambda_i\,\alpha_k, \tag{31.35}$$

and we can solve this for the α_k:

$$\alpha_k = \frac{\mathbf{x}_k^T \, \bar{\mathbb{A}} \, \mathbf{x}_i}{\lambda_i - \lambda_k}. \tag{31.36}$$

The only potential problem comes when we have $\lambda_i = \lambda_k$ – the equation corresponding to this case is the one we used to determine the correction to λ_i, and evidently, the equation itself is degenerate there. How, then, to find α_i itself? The whole perturbation analysis is neutral to the contribution of $\mathbf{x}_i$, that's what (31.33) says. And we are, in fact, free to *set* $\alpha_i = 0$.

The final eigenvector perturbation is

$$\boxed{\bar{\mathbf{x}}_i = \sum_{k=1 \neq i}^{N} \frac{\mathbf{x}_k^T \, \bar{\mathbb{A}} \, \mathbf{x}_i}{\lambda_i - \lambda_k} \, \mathbf{x}_k.} \tag{31.37}$$

Homework

Reading: Griffiths, pp. 249–254.

Problem 31.1

For the ODE:

$$\dot{x}(t) - x(t) + \epsilon \, x(t)^2 = 0 \qquad x(0) = A, \tag{31.38}$$

a. Take $x(t) = x_0(t) + \epsilon \, x_1(t)$ and solve for $x_0(t)$ and $x_1(t)$ to find the first-order perturbative solution.

b. Solve (31.38) directly (use Mathematica only as a last resort), and Taylor expand your result in ϵ, compare with your perturbative solution.

Problem 31.2

For the matrix:

$$\mathbb{A} \doteq \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{31.39}$$

find the eigenvalues and orthonormal eigenvectors.

Introduce the perturbation:

$$\bar{\mathbb{A}} \doteq \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}, \tag{31.40}$$

and calculate the first order correction to the eigenvalues (the matrix of interest, now, is $\mathbb{A} + \epsilon\,\bar{\mathbb{A}}$) and eigenvectors.