Degenerate Eigenvalue Problem

Lecture 32

Physics 342
Quantum Mechanics I

Wednesday, April 23rd, 2008

We have the matrix form of the first order perturbative result from last time. This carries over pretty directly to the Schrödinger equation, with only minimal replacement (the inner product and finite vector space change, but notationally, the results are identical). Because there are a variety of quantum mechanical systems with degenerate spectra (like the Hydrogen eigenstates, each E_n has n^2 associated eigenstates) and we want to be able to predict the energy shift associated with perturbations in these systems, we can copy our arguments for matrices to cover matrices with more than one eigenvector per eigenvalue. The punch line of that program is that we can use the non-degenerate perturbed energies, provided we start with the "correct" degenerate linear combinations.

32.1 Degenerate Perturbation

Going back to our symmetric matrix example, we have $\mathbb{A} \in \mathbb{R}^{N\times N}$, and again, a set of eigenvectors and eigenvalues: $\mathbb{A}\,\mathbf{x}_i = \lambda_i\,\mathbf{x}_i$. This time, suppose that the eigenvalue λ_i has a set of M associated eigenvectors – that is, suppose a set of eigenvectors $\mathbf{y}_j$ satisfy:

$$\mathbb{A}\,\mathbf{y}_j = \lambda_i\,\mathbf{y}_j \quad j = 1 \longrightarrow M \tag{32.1}$$

(so this represents M separate equations) that are themselves orthonormal[1]. Clearly, any linear combination of these vectors is also an eigenvector:

$$\mathbb{A} \sum_{k=1}^{M} \beta_k \, \mathbf{y}_k = \lambda_i \sum_{k=1}^{M} \beta_k \, \mathbf{y}_k. \tag{32.2}$$

Define the general combination of $\{\mathbf{y}_i\}_{i=1}^{M}$ to be $\mathbf{z} \equiv \sum_{k=1}^{M} \beta_k \, \mathbf{y}_k$, also an eigenvector of $\mathbb{A}$ with eigenvalue λ_i. Now going back to our perturbation: We're going to perturb the matrix $\mathbb{A}$ via $\mathbb{A} + \epsilon \, \bar{\mathbb{A}}$, and we will perturb the linear combination $\mathbf{z} \longrightarrow \mathbf{z} + \epsilon \, \bar{\mathbf{x}}_i$, so we do not yet know what particular linear combination we are referring to, but nevertheless, we can do them all at once. Our perturbed eigenvalue equation reads, through first order in ϵ:

$$\mathbb{A} \, \mathbf{z} + \epsilon \, \mathbb{A} \, \bar{\mathbf{x}}_i + \epsilon \, \bar{\mathbb{A}} \, \mathbf{z} = \lambda_i \, \mathbf{z} + \epsilon \, \lambda_i \, \bar{\mathbf{x}}_i + \epsilon \, \bar{\lambda}_i \, \mathbf{z}. \tag{32.3}$$

Regardless of our decomposition within the degenerate subspace (i.e. the particular values of β_k), the e^0 terms cancel. That leaves us with the usual:

$$\mathbb{A} \, \bar{\mathbf{x}}_i + \bar{\mathbb{A}} \, \mathbf{z} = \lambda_i \, \bar{\mathbf{x}}_i + \bar{\lambda}_i \, \mathbf{z}. \tag{32.4}$$

We once again assume that the perturbed eigenvector can be written as a linear combination of the original eigenvectors. We will split the contributions

$$\bar{\mathbf{x}}_i = \sum_{k=1}^{N-M} \alpha_k \, \mathbf{x}_k + \sum_{k=1}^{M} \gamma_k \, \mathbf{y}_k. \tag{32.5}$$

Inputting this into (32.4), together with the decomposition of $\mathbf{z}$, we have

$$(\mathbb{A} - \lambda_i \, \mathbb{I}) \left(\sum_{k=1}^{N-M} \alpha_k \, \mathbf{x}_k + \sum_{k=1}^{M} \gamma_k \, \mathbf{y}_k \right) + \left(\bar{\mathbb{A}} - \bar{\lambda}_i \, \mathbb{I}\right) \sum_{k=1}^{M} \beta_k \, \mathbf{y}_k = 0, \tag{32.6}$$

and we can act on the first term with $\mathbb{A}$

$$\left(\sum_{k=1}^{N-M} \alpha_k \, (\lambda_k - \lambda_i) \, \mathbf{x}_k \right) + \left(\bar{\mathbb{A}} - \bar{\lambda}_i \, \mathbb{I}\right) \sum_{k=1}^{M} \beta_k \, \mathbf{y}_k = 0. \tag{32.7}$$

[1] What we have is a subspace of $\mathbb{R}^N$, defined by the degenerate eigenvectors. These are all orthogonal to the rest of the eigenvectors, and we can find a basis spanning the subspace that will be orthogonal within the subspace. Think of partitioning the eigenvectors so that $\{\mathbf{x}_j\}_{j=1}^{N-M}$ are non-degenerate, and then $\{\mathbf{y}_i\}_{i=1}^{M} = \{\mathbf{x}_{i+(N-M)}\}_{i=1}^{M}$ are the eigenvectors sharing the eigenvalue λ_i.

Finally, if we take the dot product w.r.t. $\mathbf{y}_j$, we have

$$\sum_{k=1}^{M} \mathbf{y}_j^T \, \bar{\mathbb{A}} \, \mathbf{y}_k \, \beta_k = \bar{\lambda}_i \, \beta_j. \tag{32.8}$$

The above can be written as a matrix equation – define the matrix $\mathbb{W}$ by its entries: $W_{jk} = \mathbf{y}_j^T \, \bar{\mathbb{A}} \, \mathbf{y}_k$ – notice that this is symmetric, then we have

$$\mathbb{W} \, \boldsymbol{\beta} = \bar{\lambda}_i \, \boldsymbol{\beta}. \tag{32.9}$$

Now we see the point – if you chose a linear combination of the degenerate vectors, $\mathbf{z}$ defined in terms of coefficients β_k, then if, as a vector, $\boldsymbol{\beta}$ was an eigenvector of $\mathbb{W}$, we have a "good" choice of initial unperturbed vector $\mathbf{z}$[2], *and* we know the correction to the energy that results from perturbation. This is an example of "lifting the degeneracy" of the initial set $\mathbf{y}_k$.

32.1.1 Example

As an example, take the matrix

$$\mathbb{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{32.10}$$

this has eigenvalues $\lambda_1 = 2$, $\lambda_2 = 0$, and $\lambda_3 = 0$, and eigenvectors:

$$\mathbf{x}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{y}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{y}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{32.11}$$

In this case, $\mathbf{y}_1^T \, \mathbf{y}_2 = 0$ already, so we do not have to go through a process of orthogonalization.

Take the perturbation matrix to be

$$\bar{\mathbb{A}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \tag{32.12}$$

[2]Of course, this begs the question: "When can you find a simultaneous eigenvector of the matrices $\mathbb{W}$, $\mathbb{A}$ and $\bar{\mathbb{A}}$" ...

then we have the matrix $\mathbb{W}$ given by

$$\mathbb{W} = \begin{pmatrix} \mathbf{y}_1^T \bar{\mathbb{A}} \mathbf{y}_1 & \mathbf{y}_1^T \bar{\mathbb{A}} \mathbf{y}_2 \\ \mathbf{y}_2^T \bar{\mathbb{A}} \mathbf{y}_1 & \mathbf{y}_2^T \bar{\mathbb{A}} \mathbf{y}_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & \sqrt{2} \end{pmatrix}. \tag{32.13}$$

We are trying to find the linear combination $\mathbf{z} = \beta_1 \mathbf{y}_1 + \beta_2 \mathbf{y}_2$, so what we want to do, referring to (32.9) is find the eigenvectors and eigenvalues of $\mathbb{W}$ – the eigenvalues themselves are the perturbations to the original degenerate eigenvalue of $\mathbb{A}$ (i.e. $\lambda_2 = \lambda_3 = 0$) – for $\mathbb{W}$, we have

$$\bar{\lambda}_i = \frac{1}{2}\left(1 \pm \sqrt{3}\right), \tag{32.14}$$

with eigenvectors:

$$\boldsymbol{\beta}_\pm \sim \begin{pmatrix} \frac{1}{\sqrt{2}}\left(-1 \pm \sqrt{3}\right) \\ 1 \end{pmatrix} \tag{32.15}$$

(we have left the eigenvectors unnormalized). This tells us that the particular linear combinations we should have started with, $\mathbf{z}$ are

$$\mathbf{z}_\pm = \frac{1}{\sqrt{2}}\left(-1 \pm \sqrt{3}\right)\mathbf{y}_1 + \mathbf{y}_2, \tag{32.16}$$

while these are both eigenvectors of $\mathbb{A}$ with eigenvalue 0, they become separate eigenvectors of $\mathbb{A} + \epsilon\,\bar{\mathbb{A}}$. Notice that if we are smart, we can choose "good" $\mathbf{z}$ from the start (i.e. without going through all of this), but it would have been difficult to motivate the choice of starting vectors: $\mathbf{x}_1$, $\mathbf{z}_+$, and $\mathbf{z}_-$.

Finishing the job, we have, for the matrix $\mathbb{A} + \epsilon\,\bar{\mathbb{A}}$ (using primes to denote these new eigenvalues):

$$\lambda_2' = 0 + \frac{1}{2}\,\epsilon\left(1 + \sqrt{3}\right), \quad \lambda_3' = 0 + \frac{1}{2}\,\epsilon\left(1 - \sqrt{3}\right). \tag{32.17}$$

Returning to the original perturbation, we have

$$\mathbb{A}\,\bar{\mathbf{z}}_\pm + \bar{\mathbb{A}}\,\mathbf{z}_\pm = \lambda_2\,\bar{\mathbf{z}}_\pm + \bar{\lambda}_i\,\mathbf{z}_\pm. \tag{32.18}$$

To find the corrections $\bar{\mathbf{z}}_\pm$, we can dot both sides of the above into $\mathbf{x}_1$:

$$\lambda_1\,\mathbf{x}_1^T\,\bar{\mathbf{z}}_\pm + \mathbf{x}_1^T\,\bar{\mathbb{A}}\,\mathbf{z}_\pm = \lambda_2\,\mathbf{x}_1^T\,\bar{\mathbf{z}}_\pm. \tag{32.19}$$

Since, by assumption, the eigenvector perturbation has no contribution in the $\mathbf{z}_\pm$ directions, we have:

$$\bar{\mathbf{z}}_\pm = \frac{\mathbf{x}_1^T \,\bar{\mathbb{A}}\, \mathbf{z}_\pm}{\lambda_2 - \lambda_1}\, \mathbf{x}_1, \tag{32.20}$$

so we expect to get the new eigenvectors:

$$\begin{aligned} \mathbf{z}'_+ &= \mathbf{z}_+ + \epsilon\, \frac{\mathbf{x}_1^T \,\bar{\mathbb{A}}\, \mathbf{z}_+}{\lambda_2 - \lambda_1}\, \mathbf{x}_1 \\ \mathbf{z}'_- &= \mathbf{z}_- + \epsilon\, \frac{\mathbf{x}_1^T \,\bar{\mathbb{A}}\, \mathbf{z}_-}{\lambda_2 - \lambda_1}\, \mathbf{x}_1. \end{aligned} \tag{32.21}$$

The key element in all of this was our ability to diagonalize $\mathbb{W}$. For the quantum mechanical analogue of this procedure, we often have a clue about "good" combinations within the degenerate subspace – if there was, here, a matrix $\mathbb{B}$ that commuted with *both* $\mathbb{A}$ and $\bar{\mathbb{A}}$, then we know that $\mathbb{B}$ has eigenvectors that are shared with $\mathbb{A}$ and $\bar{\mathbb{A}}$, and it is these eigenvectors that will guide our choice of $\mathbf{z}_\pm$, rather than the matrix $\mathbb{W}$ (although we are always free to work from there).

32.2 Perturbation Theory and Quantum Mechanics

All of our discussion so far carries over to quantum mechanical perturbation theory – we could have developed all of our formulae in terms of bra-ket notation, and there would literally be no difference between our finite real matrices and the Hermitian operator eigenvalue problem. For example, take the non-degenerate, first order perturbative result: For a matrix $\mathbb{A}$ with eigenvectors/values $\mathbb{A}\,\mathbf{x}_i = \lambda_i\,\mathbf{x}_i$, and a perturbing matrix $\bar{\mathbb{A}}$, the solution to

$$\left(\mathbb{A} + \epsilon\,\bar{\mathbb{A}}\right)\left(\mathbf{x}_i + \epsilon\,\bar{\mathbf{x}}_i\right) = \left(\lambda_i + \epsilon\,\bar{\lambda}_i\right)\left(\mathbf{x}_i + \epsilon\,\bar{\mathbf{x}}_i\right) \tag{32.22}$$

is

$$\begin{aligned} \lambda_i + \epsilon\,\bar{\lambda}_i &= \lambda_i + \epsilon\,\mathbf{x}_i^T\,\bar{\mathbb{A}}\,\mathbf{x}_i \\ \mathbf{x}_i + \epsilon\,\bar{\mathbf{x}}_i &= \mathbf{x}_i + \epsilon \sum_{k=1\neq i}^{N} \frac{\mathbf{x}_k^T\,\bar{\mathbb{A}}\,\mathbf{x}_i}{\lambda_i - \lambda_k}\,\mathbf{x}_k \end{aligned} \tag{32.23}$$

If we rewrite this entire problem in bra-ket notation, what we have is: $\mathbb{A}\,|x_i\rangle = \lambda_i\,|x_i\rangle$, and the perturbation results are:

$$\begin{aligned} \lambda_i + \epsilon\,\bar{\lambda}_i &= \lambda_i + \epsilon\,\langle x_i|\,\bar{\mathbb{A}}\,|x_i\rangle \\ |x_i\rangle + \epsilon\,|\bar{x}_i\rangle &= |x_i\rangle + \epsilon\,\sum_{k=1\neq i}^{N} \frac{\langle x_k|\,\bar{\mathbb{A}}\,|x_i\rangle}{\lambda_i - \lambda_k}\,|x_k\rangle\,. \end{aligned} \tag{32.24}$$

Our particular interest is in a Hamiltonian H and a perturbing Hamiltonian H', so that for a complete set of wavefunctions with $H\,|\psi_n\rangle = E_n\,|\psi_n\rangle$, we have the first order corrections forming the approximation to the eigenstates/energies of $H + \epsilon\,H'$:

$$\begin{aligned} E_i' \approx E_i + \epsilon\,\bar{E}_i &= E_i + \epsilon\,\langle\psi_i|\,H'\,|\psi_i\rangle \\ \left|\psi_i'\right\rangle \approx \left|\psi_i'\right\rangle + \epsilon\,\left|\bar{\psi}_i'\right\rangle &= |\psi_i\rangle + \epsilon\,\sum_{k=1\neq i}^{\infty} \frac{\langle\psi_k|\,H'\,|\psi_k\rangle}{E_i - E_k}\,|\psi_k\rangle\,. \end{aligned} \tag{32.25}$$

There is, functionally, no difference in the argument here. One has to be a little careful about inner products (which are now complex), and convergence for the infinite sum, but those are details for particular cases.

32.3 Example

Let's calculate the eigenstates and energies of the following potential:

$$V(x) = \begin{cases} \epsilon\,x\,(x-a) & 0 < x < a \\ \infty & x < 0 \;\; x > a \end{cases}\,. \tag{32.26}$$

Our first question: What are H and H'? We want to choose H so that the unperturbed states are easy to calculate. In this case, the natural choice for the unperturbed problem is the infinite square well. We know the solution here, we have:

$$\psi''(x) = -\frac{2\,m\,E}{\hbar^2}\,\psi(x) \tag{32.27}$$

so the solutions are

$$\psi_n(x) = \sqrt{\frac{2}{a}}\,\sin\Big(\frac{n\,\pi\,x}{a}\Big) \qquad n \in \mathbb{Z} \tag{32.28}$$

with

$$E_n = \frac{\hbar^2\,n^2\,\pi^2}{2\,a^2\,m} \tag{32.29}$$

as usual.

Now for the perturbation – for the energies, we need the appropriate perturbing Hamiltonian– in this case, $H' = x\,(x-a)$, and (32.25) read:

$$\begin{aligned} E'_n &= E_n + \epsilon \int_{-a}^{a} \psi_n(x)\, x\,(x-a)\ \psi_n(x)\, dx \\ &= \frac{\hbar^2\, n^2\, \pi^2}{2\, a^2\, m} + \epsilon \int_{-a}^{a} x\,(x-a) \left(\frac{2}{a} \sin^2\!\left(\frac{n\,\pi\,x}{a} \right) \right) dx \\ &= \frac{\hbar^2\, n^2\, \pi^2}{2\, a^2\, m} - \epsilon \left(\frac{a^2}{6} + \frac{a^2}{2\,\pi^2\, n^2} \right) \end{aligned} \tag{32.30}$$

This works out well – using a simple numerical solution, we can plot the difference in the first twenty eigenvalues of the "exact" (but numerical) solution to the perturbed problem, and compare with our theoretical prediction, the result is shown in Figure 32.1 (for $a = 1$, $\epsilon = 10$, in units where $\hbar^2 = 1$, $m = \frac{1}{2}$).

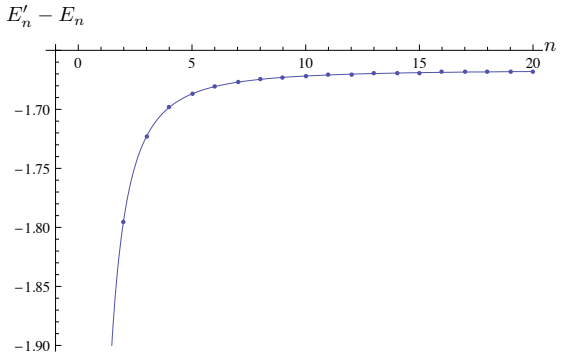

Figure 32.1: The difference $E'_n - E_n$ computed numerically (using 100 grid points) is shown as points, and the predicted function (32.30) is shown as a continuous line.

The eigenvector perturbations are harder to compute, and not as dramatic.

Homework

Reading: Griffiths, pp. 255–266.

Problem 32.1

Griffiths 6.1 a. Calculating the energy adjustments for a delta perturbation in an infinite square well.

Problem 32.2

Griffiths 6.2. Perturbing the harmonic oscillator. Remember to write the operator x^2 in terms of raising and lowering operators.

Problem 32.3

Here, we will look at the Stern-Gerlach Hamiltonian perturbatively. Ignoring the kinetic term, we can write:

$$H = -\gamma\, \mathbf{B} \cdot \mathbf{S} \tag{32.31}$$

with

$$\mathbf{B} = -\alpha\, x\, \hat{\mathbf{x}} + (B_0 + \alpha\, z)\, \hat{\mathbf{z}}. \tag{32.32}$$

a. Assuming α "small", split the Hamiltonian into a piece of order α^0 and one of order α (identifying H_0 and H', effectively). Find the eigenvalues and eigenvectors of the unperturbed Hamiltonian.

b. Calculate the correction to the two energies (associated with the spin-up and spin-down states) implied by the H' term.

Problem 32.4

For the matrix:

$$\mathbb{A} \doteq \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{32.33}$$

with perturbation

$$\bar{\mathbb{A}} \doteq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{32.34}$$

we argued in class that the perturbed eigenvalues of $\mathbb{A} + \epsilon\,\bar{\mathbb{A}}$ depended on our choice of eigenvectors for the null space of $\mathbb{A}$ (which is degenerate). In particular, we used a third matrix $\mathbb{Q}$ that commutes with *both* $\mathbb{A}$ and $\bar{\mathbb{A}}$ and hence shared their eigenvectors to *choose* a linear combination that gave the correct answer. In this problem, you will show that this process is highly constraining.

a. Construct a matrix $\mathbb{Q}$ with

$$[\mathbb{A}, \mathbb{Q}] = [\bar{\mathbb{A}}, \mathbb{Q}] = 0 \tag{32.35}$$

and $\mathbb{Q} \neq \alpha\,\mathbb{A}$ (i.e. $\mathbb{Q}$ is not a simple multiple of $\mathbb{A}$).

b. Find the eigenvectors of the matrix $\mathbb{Q}$ and using these, compute $\bar{\lambda}_1$ and $\bar{\lambda}_2$, the first order corrections to the $\lambda = 0$ eigenvalue of $\mathbb{A}$ associated with the perturbed matrix $\mathbb{A} + \epsilon\,\bar{\mathbb{A}}$.