

## Quantum numbers and complete sets of commuting observables

In recent lectures we have seen that it is convenient to choose basis states for the Hilbert space that correspond to eigenstates of the Hamiltonian  $\mathbf{H}$ :

$$\begin{aligned} \sum_j |\Psi_j\rangle\langle\Psi_j| &= \mathbf{1}, \\ \langle\Psi_j|\Psi_k\rangle &= \delta_{jk}, \\ \mathbf{H}|\Psi_j\rangle &= \varepsilon_j|\Psi_j\rangle. \end{aligned} \tag{1}$$

We can do this because  $\mathbf{H}$  is a normal (in fact, Hermitian) operator. When arbitrary states  $|\Phi\rangle$  are expressed as linear combinations of the  $|\Psi_j\rangle$ , it becomes trivial to evolve them forward in time:

$$\begin{aligned} |\Phi\rangle &= \sum_j c_j(t)|\Psi_j\rangle, \\ c_j(t) &= \exp(-i\varepsilon_j t/\hbar)c_j(0). \end{aligned} \tag{2}$$

Accordingly, it's often convenient to label energy basis states directly by their corresponding eigenvalues,  $\{|\varepsilon_j\rangle\}$ :

$$\mathbf{H}|\varepsilon_j\rangle = \varepsilon_j|\varepsilon_j\rangle. \tag{3}$$

Of course, this labeling can be unambiguous only if all the eigenvalues of  $\mathbf{H}$  are distinct!

Generally speaking, any physical quantity  $q$  for which the eigenvalues of  $\mathbf{O}_q$  can be used to specify members of an orthonormal basis can be thought of as a **quantum number**. Given the central importance of the energy basis in quantum dynamics, it is natural to focus on quantum numbers that specify eigenvectors of  $\mathbf{H}$ .

In practice, it is often the case that the Hamiltonian is proportional to some observable corresponding to a different physical quantity. For example, the Hamiltonian for a spin- $\frac{1}{2}$  particle in a magnetic field  $B_z$  (whose direction we use to define the  $z$  axis) may be written

$$\mathbf{H} = \omega_L \mathbf{S}_z, \tag{4}$$

where  $\omega_L = -\gamma B_z$  and  $\mathbf{S}_z$  is the observable corresponding to projection of spin along the  $z$  axis,

$$\mathbf{S}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5}$$

As we have written the matrix representation for  $\mathbf{S}_z$  in the energy eigenbasis, we clearly see that its eigenvectors are the same as those of  $\mathbf{H}$  and that its eigenvalues are  $\pm\hbar/2$ . We could therefore just as well use eigenvalues of  $\mathbf{S}_z$  to label the energy basis states:

$$\begin{aligned}\mathbf{H}|\hbar/2\rangle &= \frac{1}{2}\hbar\omega_L, \\ \mathbf{H}|-\hbar/2\rangle &= -\frac{1}{2}\hbar\omega_L.\end{aligned}\tag{6}$$

Hence  $z$  projection of spin represents a natural quantum number for this system.

Given a particular Hamiltonian  $\mathbf{H}$ , a physical quantity  $q$  is said to be a 'good' or 'conserved' quantum number if every eigenvector of  $\mathbf{O}_q$  remains an eigenvector with the **same** eigenvalue, as time evolves. Hence, if

$$\mathbf{O}_q|j_q\rangle = \mathbf{O}_q \sum_k c_k(0)|\varepsilon_k\rangle = q_j|j_q\rangle,\tag{7}$$

then we require

$$\mathbf{O}_q \sum_k c_k(0) \exp(-i\varepsilon_k t/\hbar)|\varepsilon_k\rangle = q_j \sum_k c_k(0) \exp(-i\varepsilon_k t/\hbar)|\varepsilon_k\rangle\tag{8}$$

for all eigenvectors  $|j_q\rangle$  in order to say that  $q$  is a good quantum number.

Note that a necessary and sufficient condition for  $q$  to be a good quantum number is that  $\mathbf{O}_q$  commute with  $\mathbf{H}$ . Assume that  $[\mathbf{O}_q, \mathbf{H}] = 0$ . If  $|\Psi(0)\rangle$  is an eigenvector of  $\mathbf{O}_q$

$$\mathbf{O}_q|\Psi(0)\rangle = q_j|\Psi(0)\rangle,\tag{9}$$

then

$$\begin{aligned}\mathbf{O}_q|\Psi(t)\rangle &= \mathbf{O}_q \mathbf{T}(t,0)|\Psi(0)\rangle \\ &= \mathbf{O}_q \exp(-i\mathbf{H}t/\hbar)|\Psi(0)\rangle \\ &= \mathbf{O}_q \sum_{n=0}^{\infty} \frac{1}{n!} (-i\mathbf{H}t/\hbar)^n |\Psi(0)\rangle \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} (-i\mathbf{H}t/\hbar)^n \mathbf{O}_q |\Psi(0)\rangle \\ &= q_j |\Psi(t)\rangle.\end{aligned}\tag{10}$$

So for a two-dimensional system whose Hamiltonian is given by

$$\mathbf{H} = \omega_L \mathbf{S}_z,\tag{11}$$

the  $x$  projection of spin

$$\mathbf{s}_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\tag{12}$$

is **not** a good quantum number since

$$[\mathbf{S}_x, \mathbf{H}] \neq 0.\tag{13}$$

Things are pretty trivial in two dimensions, but let's now consider a three-dimensional system with the following Hamiltonian:

$$\mathbf{H} = \varepsilon_0 \mathbf{O}_\alpha,$$

$$\mathbf{O}_\alpha = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad 14$$

This Hamiltonian is **degenerate**, by which we mean that it has repeated eigenvalues  $(\varepsilon_0, 2\varepsilon_0, \varepsilon_0)$ . Hence we cannot use the eigenvalues of  $\mathbf{H}$  (or  $\mathbf{O}_\alpha$ ) to uniquely label the energy basis states. But note that if we are careful, we can specify an additional observable that ‘lifts’ this degeneracy without introducing any inconsistencies to the labeling scheme. An obvious choice would be

$$\mathbf{O}_\gamma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad 15$$

whose eigenvalues are  $(1, 1, 2)$ . Clearly, we can uniquely specify any particular member of the energy basis by listing its eigenvalues with respect to both  $\mathbf{H}$  and  $\mathbf{O}_\gamma$ :

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \leftrightarrow |\varepsilon_0, 1\rangle, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \leftrightarrow |2\varepsilon_0, 1\rangle, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \leftrightarrow |\varepsilon_0, 2\rangle. \quad 16$$

Note that we had to be sure to pick an observable whose eigenstates could be chosen in common with  $\mathbf{H}$ , or else it would not have made sense to specify both eigenvalues simultaneously!

Another example of an observable that lifts the degeneracy of  $\mathbf{H}$  would be

$$\mathbf{O}_\beta = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}. \quad 17$$

In this case, the eigenvalues are  $(0, 0, 2)$  with corresponding eigenvectors

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \leftrightarrow |\varepsilon_0, 0\rangle, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \leftrightarrow |2\varepsilon_0, 0\rangle, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \leftrightarrow |\varepsilon_0, 2\rangle. \quad 18$$

The difference in choice of  $\mathbf{O}_\beta$  rather than  $\mathbf{O}_\alpha$  corresponds to a different choice of basis within the degenerate subspace of  $\mathbf{H}$ .

Recall that the necessary and sufficient condition for two observables  $\mathbf{O}_1$  and  $\mathbf{O}_2$  to have simultaneous eigenstates is  $[\mathbf{O}_1, \mathbf{O}_2] = 0$ . For this reason, commuting observables are said to be ‘compatible.’

A set of mutually compatible observables  $\{\mathbf{O}_j\}$  whose simultaneous eigenvalues **uniquely** specify a particular member of a complete orthonormal basis is called a ‘Complete Set of Commuting Observables’ (CSCO). Hence, a CSCO is associated with a complete orthonormal basis  $\{|k\rangle\}$  for the relevant Hilbert space, such that each  $|k\rangle$  is a simultaneous

eigenstate of every  $\mathbf{O}_j$  and such that a list of eigenvalues  $\{q_j\}$  uniquely specifies  $|k\rangle$ .

Generally speaking, the most useful kind of CSCO is one in which each observable corresponds to a conserved quantum number. Such a CSCO specifies the energy basis, so in fact the Hamiltonian alone can suffice *if* it is nondegenerate. Otherwise we need to add additional observables to arrive at a CSCO, and different choices will lead to different complete orthonormal bases, corresponding to different basis within the degenerate subspace(s) of  $\mathbf{H}$ .

## Stationary perturbation theory

Let's now return to the perturbation scenario that we discussed in general for two-level systems. We start with a given Hamiltonian  $\mathbf{H}_0$ , with eigenvalues and eigenstates

$$\begin{aligned}\mathbf{H}_0|1\rangle &= E_1|1\rangle, \\ \mathbf{H}_0|2\rangle &= E_2|2\rangle.\end{aligned}\tag{19}$$

Even if  $E_1 = E_2$ , we assume that the eigenstates have been chosen such that  $\langle 1|2\rangle = 0$ . The given energy basis is thus complete and orthonormal, and the matrix representation of  $\mathbf{H}_0$  is simply

$$\mathbf{H}_0 \leftrightarrow \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}.\tag{20}$$

Now we want to consider the effect of a 'perturbation' to  $\mathbf{H}_0$ . By this we mean that we want to consider a scenario in which a new energy term  $\mathbf{W}$  is added to produce a new Hamiltonian

$$\mathbf{H}_{tot} = \mathbf{H}_0 + \mathbf{W}.\tag{21}$$

Specifically, we want to determine the new eigenvalues and eigenvectors in terms of  $E_{1,2}$  and  $\{|1\rangle, |2\rangle\}$  and the matrix elements of  $\mathbf{W}$ . Notationally, we introduce

$$\begin{aligned}\mathbf{H}_{tot}|\Psi_-\rangle &= E_-|\Psi_-\rangle, \\ \mathbf{H}_{tot}|\Psi_+\rangle &= E_+|\Psi_+\rangle.\end{aligned}\tag{22}$$

For a two-level quantum system, one can solve the general case analytically (see Cohen-Tannoudji *et al* complement B<sub>IV</sub>)! Writing

$$\mathbf{W} \leftrightarrow \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix},\tag{23}$$

with  $W_{11}$  and  $W_{22}$  real and  $W_{12} = W_{21}^*$ ,

$$\begin{aligned}
E_+ &= \frac{1}{2}(E_1 + W_{11} + E_2 + W_{22}) + \frac{1}{2}\sqrt{(E_1 + W_{11} - E_2 - W_{22})^2 + 4|W_{12}|^2}, \\
E_- &= \frac{1}{2}(E_1 + W_{11} + E_2 + W_{22}) - \frac{1}{2}\sqrt{(E_1 + W_{11} - E_2 - W_{22})^2 + 4|W_{12}|^2}, \\
|\Psi_+\rangle &= \cos\frac{\theta}{2}\exp(-i\varphi/2)|1\rangle + \sin\frac{\theta}{2}\exp(+i\varphi/2)|2\rangle, \\
|\Psi_-\rangle &= -\sin\frac{\theta}{2}\exp(-i\varphi/2)|1\rangle + \cos\frac{\theta}{2}\exp(+i\varphi/2)|2\rangle, \\
\tan\theta &= \frac{2|W_{12}|}{E_1 + W_{11} - E_2 - W_{22}} \quad (0 \leq \theta < \pi), \\
W_{21} &= |W_{21}|\exp(i\varphi).
\end{aligned} \tag{24}$$

These are obtained by explicit diagonalization of

$$\mathbf{H}_{tot} \leftrightarrow \begin{pmatrix} E_1 + W_{11} & W_{12} \\ W_{21} & E_2 + W_{22} \end{pmatrix}. \tag{25}$$

In the last lecture (see also Cohen-Tannoudji Ch. IV section C) we saw that  $W_{11}$  and  $W_{22}$  contributed simple level shifts while  $W_{12}$  is responsible for coupling the eigenstates of  $\mathbf{H}_0$ .

A simple example of such a perturbation comes from the general Hamiltonian for a spin- $\frac{1}{2}$  particle in a magnetic field:

$$\begin{aligned}
\mathbf{H}_{tot} &= -\gamma\vec{\mathbf{S}} \cdot \vec{\mathbf{B}} \\
&= -\gamma(\mathbf{S}_x B_x + \mathbf{S}_y B_y + \mathbf{S}_z B_z).
\end{aligned} \tag{26}$$

Here  $\gamma$  is the particle's gyromagnetic ratio and  $\mathbf{S}_x, \mathbf{S}_y, \mathbf{S}_z$  are observables corresponding to projections of the particle's spin along the  $x, y, z$  directions. If we start with a magnetic field that points purely along  $z$ ,

$$\mathbf{H}_0 = -\gamma\mathbf{S}_z B_z. \tag{27}$$

The addition of a magnetic field component along  $x$  could then be considered as a perturbation,

$$\mathbf{W} = -\gamma\mathbf{S}_x B_x. \tag{28}$$

What one finds from the above analytic expressions is that the new eigenstates correspond to particle spin pointing parallel or antiparallel to the vector  $\vec{\mathbf{B}} = (B_x, 0, B_z)$  and a change in the energy splitting from  $|E_1 - E_2| = \hbar\gamma B_z$  to  $|E_+ - E_-| = \hbar\gamma\sqrt{B_x^2 + B_z^2}$ .

This is all great, but unfortunately it's really only for an isolated two-dimensional system that the effects of a general perturbation can be computed exactly! In more complicated scenarios, one must often resort to **perturbation theory**, which gives prescriptions for how to approximate the effects of a perturbation up to some fixed order in the size of the perturbation.

As we've been learning, 'solving' the Schrödinger Equation with a time-independent Hamiltonian pretty much amounts to determining the eigenvalues and eigenvectors of  $\mathbf{H}$ . Accordingly, **stationary** perturbation theory (which deals with time-independent  $\mathbf{H}_{tot}$ ) provides methods to approximate changes in the eigenvalues and eigenvectors. In the notation used above, this means that we will be computing leading orders in expansions such as

$$E_+ = E_1 + \sum_{n=1}^{\infty} a_n \lambda^n, \quad 29$$

where  $\lambda$  is a measure of the ‘size’ of the perturbation. In any power series of this sort, it should be clear that the  $a_n$  should have the same units as  $E_+$  and that  $\lambda$  must be dimensionless. In order for a low-order approximation to be valid we must have  $\lambda \ll 1$ , which limits the range of problems for which perturbation theory is useful. In practice it often is very useful!

But what is this parameter  $\lambda$ ? Generally speaking, it corresponds to the ratio of ‘relevant’ matrix elements of  $\mathbf{W}$  to relevant matrix elements of  $\mathbf{H}_0$ . It’s hard to be much more precise at this point. To illustrate what we mean, however, consider the exact expression for perturbation of a two-level system (with  $W_{11} = W_{22} = 0$ ):

$$E_+ = \frac{1}{2}(E_1 + E_2) + \frac{1}{2}\sqrt{(E_1 - E_2)^2 + 4|W_{12}|^2}. \quad 30$$

We can clearly make an approximation to this expression by assuming  $|W_{12}| \ll (E_1 - E_2)$ ,

$$\begin{aligned} E_+ &= \frac{1}{2}(E_1 + E_2) + \frac{1}{2}(E_1 - E_2) \sqrt{1 + \left(\frac{2|W_{12}|}{E_1 - E_2}\right)^2} \\ &\approx \frac{1}{2}(E_1 + E_2) + \frac{1}{2}(E_1 - E_2) \left(1 + \frac{1}{2} \left(\frac{2|W_{12}|}{E_1 - E_2}\right)^2\right) \\ &= E_1 + \frac{|W_{12}|^2}{E_1 - E_2}. \end{aligned} \quad 31$$

A typical application of stationary perturbation theory would be to arrive at this approximate expression directly, without first having to derive the exact answer. In fact, the above expression is second-order in matrix elements of  $W$  – it so happens in the limit we chose ( $|W_{12}| \ll (E_1 - E_2)$ ) the first-order term is zero.

## General formulation of perturbative expansions

(Following Cohen-Tannoudji *et al*, Ch. XI).

We assume that the total Hamiltonian for a finite-dimensional system is given by

$$\mathbf{H} = \mathbf{H}_0 + \mathbf{W},$$

where the eigenvalues and eigenstates of the ‘bare’ Hamiltonian  $\mathbf{H}_0$  are already known:

$$\mathbf{H}_0 |\varphi_n^i\rangle = E_n^0 |\varphi_n^i\rangle.$$

(Here the additional index  $i$  in  $|\varphi_n^i\rangle$  allows us when necessary to specify a set of basis states within the subspace corresponding to a degenerate eigenvalue.) Our goal is to compute the effect of the perturbation  $\mathbf{W}$ .

Perturbation theory focuses on the situation where  $\mathbf{W}$  is much smaller than  $\mathbf{H}_0$ . In particular, it is useful to define

$$\mathbf{W} = \lambda \bar{\mathbf{W}},$$

where the matrix elements of  $\bar{\mathbf{W}}$  are comparable to those of  $\mathbf{H}_0$  (in particular, certain ones must be smaller than the differences between the eigenvalues of  $\mathbf{H}_0$ ) and  $\lambda \ll 1$ . We can then think of the total Hamiltonian as being a function of the dimensionless parameter  $\lambda$ ,

and formulate the general eigenvalue relation

$$\mathbf{H}(\lambda)|\Psi(\lambda)\rangle = E(\lambda)|\Psi(\lambda)\rangle.$$

We assume that  $E(\lambda)$  and  $|\Psi(\lambda)\rangle$  can be expanded in powers of  $\lambda$ ,

$$E(\lambda) = \varepsilon_0 + \lambda\varepsilon_1 + \lambda^2\varepsilon_2 + O(\lambda^3),$$

$$|\Psi(\lambda)\rangle = |0\rangle + \lambda|1\rangle + \lambda^2|2\rangle + O(\lambda^3).$$

(Note that  $|0\rangle, |1\rangle, |2\rangle, \dots$  here refer to terms in an expansion and are not supposed to refer to any pre-defined basis set, unlike our usual notation in these lectures). Recalling  $\mathbf{H}(\lambda) = \mathbf{H}_0 + \lambda\bar{\mathbf{W}}$ , we can then substitute the expansions into the eigenvalue relation:

$$(\mathbf{H}_0 + \lambda\bar{\mathbf{W}}) \left[ \sum_{q=0}^{\infty} \lambda^q |q\rangle \right] = \left[ \sum_{q'=0}^{\infty} \lambda^{q'} \varepsilon_{q'} \right] \left[ \sum_{q=0}^{\infty} \lambda^q |q\rangle \right].$$

Equating powers of  $\lambda$ , we get a hierarchy of equations:

$$\mathbf{H}_0|0\rangle = \varepsilon_0|0\rangle,$$

$$\mathbf{H}_0|1\rangle + \bar{\mathbf{W}}|0\rangle = \varepsilon_0|1\rangle + \varepsilon_1|0\rangle,$$

$$\mathbf{H}_0|2\rangle + \bar{\mathbf{W}}|1\rangle = \varepsilon_0|2\rangle + \varepsilon_1|1\rangle + \varepsilon_2|0\rangle,$$

$\vdots$

Note how this hierarchy looks essentially like recurrence relations in which  $(q+1)^{th}$ -order terms are given in terms of the  $q^{th}$  order terms and  $\mathbf{H}_0$  and  $\bar{\mathbf{W}}$ . These equations are our basic tools in perturbation theory!

When we set  $\lambda$  strictly equal to zero, each  $|\Psi(\lambda)\rangle$  and  $E(\lambda)$  correspond exactly to some eigenstate  $|\varphi_n\rangle$  and eigenvalue  $E_n^0$  of  $\mathbf{H}_0$ . Different procedures must be used to compute leading terms in the  $\lambda$  expansions, depending on whether  $E_n^0$  is a degenerate eigenvalue.

## Non-degenerate case

Recall our definition of the perturbation expansions,

$$\mathbf{H}(\lambda)|\Psi(\lambda)\rangle = E(\lambda)|\Psi(\lambda)\rangle,$$

$$E(\lambda) = \varepsilon_0 + \lambda\varepsilon_1 + \lambda^2\varepsilon_2 + O(\lambda^3),$$

$$|\Psi(\lambda)\rangle = |0\rangle + \lambda|1\rangle + \lambda^2|2\rangle + O(\lambda^3).$$

Clearly we can begin by setting  $\varepsilon_0 = E_n^0$  and  $|0\rangle = |\varphi_n\rangle$ . Our next step will be to demand that the normalization of  $|\Psi(\lambda)\rangle$  be independent of  $\lambda$ . To first order, this gives us the relation

$$\langle\Psi(\lambda)|\Psi(\lambda)\rangle = \langle 0|0\rangle + \lambda[\langle 0|1\rangle + \langle 1|0\rangle] + O(\lambda^2).$$

Since  $\langle 0|0\rangle = \langle\varphi_n|\varphi_n\rangle = 1$ , we infer that  $\langle 0|1\rangle = \langle 1|0\rangle = 0$ , *i.e.*, the first-order eigenstate correction must be orthogonal to  $|\varphi_n\rangle$ . Moving on to second order in  $\lambda$ , we get

$$\begin{aligned} \langle\Psi(\lambda)|\Psi(\lambda)\rangle &= \langle 0|0\rangle + \lambda[\langle 0|1\rangle + \langle 1|0\rangle] + \lambda^2\langle 1|1\rangle \\ &\quad + \lambda^2[\langle 0|2\rangle + \langle 2|0\rangle] + O(\lambda^3) \\ &= 1 + \lambda^2[\langle 1|1\rangle + \langle 0|2\rangle + \langle 2|0\rangle] + O(\lambda^3). \end{aligned}$$

Hence we must have

$$\langle 0|2 \rangle + \langle 2|0 \rangle = -\langle 1|1 \rangle.$$

Although we won't actually need to use the latter relation here, this type of reasoning can be used to infer normalizations and overall phase factors for the higher expansion terms  $|q\rangle$ .

We can now use the 'recurrence relation'

$$\mathbf{H}_0|1\rangle + \bar{\mathbf{W}}|0\rangle = \varepsilon_0|1\rangle + \varepsilon_1|0\rangle$$

to compute  $\varepsilon_1$  and  $|1\rangle$ . For the eigenvalue correction, we simply project onto  $\langle \varphi_n |$  yielding

$$\langle \varphi_n | \mathbf{H}_0 | 1 \rangle + \langle \varphi_n | \bar{\mathbf{W}} | 0 \rangle = \varepsilon_0 \langle \varphi_n | 1 \rangle + \varepsilon_1 \langle \varphi_n | 0 \rangle.$$

Since we know that  $\langle \varphi_n | \mathbf{H}_0 = \varepsilon_0 \langle \varphi_n |$ ,  $\langle \varphi_n | 0 \rangle = 1$ , and  $\langle \varphi_n | 1 \rangle = 0$ , we immediately have

$$\varepsilon_1 = \langle \varphi_n | \bar{\mathbf{W}} | 0 \rangle.$$

Thus to first order in  $\lambda$  (recall  $\lambda \bar{\mathbf{W}} = \mathbf{W}$ ),

$$E_n(\lambda) = E_n^0 + \langle \varphi_n | \mathbf{W} | \varphi_n \rangle + O(\lambda^2).$$

Moving on to the eigenstate correction, we project the recurrence relation onto other eigenstates of  $\mathbf{H}_0$ :

$$\langle \varphi_p^i | \mathbf{H}_0 | 1 \rangle + \langle \varphi_p^i | \bar{\mathbf{W}} | 0 \rangle = \varepsilon_0 \langle \varphi_p^i | 1 \rangle + \varepsilon_1 \langle \varphi_p^i | 0 \rangle, \quad (p \neq n)$$

$$\langle \varphi_p^i | (\mathbf{H}_0 - \varepsilon_0) | 1 \rangle = -\langle \varphi_p^i | (\bar{\mathbf{W}} - \varepsilon_1) | 0 \rangle,$$

$$(E_p^0 - E_n^0) \langle \varphi_p^i | 1 \rangle = -\langle \varphi_p^i | \bar{\mathbf{W}} | \varphi_n \rangle.$$

Since we already know  $\langle \varphi_n | 1 \rangle = 0$  and since the eigenstates of  $\mathbf{H}_0$  form a basis, this gives us a complete expansion of  $|1\rangle$ :

$$|1\rangle = \sum_{p \neq n} \sum_i \frac{\langle \varphi_p^i | \bar{\mathbf{W}} | \varphi_n \rangle}{(E_n^0 - E_p^0)} |\varphi_p^i\rangle.$$

Hence, to first order in  $\lambda$ ,

$$|\Psi_n(\lambda)\rangle = |\varphi_n\rangle + \sum_{p \neq n} \sum_i \frac{\langle \varphi_p^i | \mathbf{W} | \varphi_n \rangle}{(E_n^0 - E_p^0)} |\varphi_p^i\rangle + O(\lambda^2).$$

Hence we see that  $\mathbf{W}$  generally induces a mixing of  $|\varphi_n\rangle$  with all other bare eigenstates except those for which  $\langle \varphi_p^i | \mathbf{W} | \varphi_n \rangle = 0$ , and that the contributions are strongest for those nearby in energy (so that  $|E_n^0 - E_p^0|$  is small).

We can easily derive the second order correction to  $E_n(\lambda)$  by projecting the second recurrence relation onto  $\langle \varphi_n |$ ,

$$\langle \varphi_n | \mathbf{H}_0 | 2 \rangle + \langle \varphi_n | \bar{\mathbf{W}} | 1 \rangle = \varepsilon_0 \langle \varphi_n | 2 \rangle + \varepsilon_1 \langle \varphi_n | 1 \rangle + \varepsilon_2 \langle \varphi_n | 0 \rangle,$$

$$\varepsilon_0 \langle \varphi_n | 2 \rangle + \langle \varphi_n | \bar{\mathbf{W}} | 1 \rangle = \varepsilon_0 \langle \varphi_n | 2 \rangle + \varepsilon_2,$$

$$\varepsilon_2 = \langle \varphi_n | \bar{\mathbf{W}} | 1 \rangle.$$

Substituting our earlier expression for  $|1\rangle$ ,

$$\begin{aligned} \varepsilon_2 &= \sum_{p \neq n} \sum_i \langle \varphi_n | \bar{\mathbf{W}} \frac{\langle \varphi_p^i | \bar{\mathbf{W}} | \varphi_n \rangle}{(E_n^0 - E_p^0)} |\varphi_p^i\rangle \\ &= \sum_{p \neq n} \sum_i \frac{|\langle \varphi_p^i | \bar{\mathbf{W}} | \varphi_n \rangle|^2}{(E_n^0 - E_p^0)}. \end{aligned}$$

Hence to second order in  $\lambda$ ,

$$E_n(\lambda) = E_n^0 + \langle \varphi_n | \mathbf{W} | \varphi_n \rangle + \sum_{p \neq n} \sum_i \frac{|\langle \varphi_p^i | \mathbf{W} | \varphi_n \rangle|^2}{(E_n^0 - E_p^0)} + O(\lambda^3).$$

Regarding the latter two expressions we have derived, note that we must have relations like

$$|\langle \varphi_p^i | \mathbf{W} | \varphi_n \rangle| \ll |E_n^0 - E_p^0|,$$

in order for the expansions to converge. That is, off-diagonal elements of  $W$  (in the energy basis) must be smaller than the bare energy differences  $|E_n^0 - E_p^0|$ .

An interesting inequality can be derived from the above general expression for  $\varepsilon_2$ . If we define a quantity  $\Delta E$  such that

$$|E_n^0 - E_p^0| \geq \Delta E$$

for all  $p$ , then

$$\begin{aligned} |\varepsilon_2| &\leq \frac{1}{\Delta E} \sum_{p \neq n} \sum_i |\langle \varphi_p^i | \bar{\mathbf{W}} | \varphi_n \rangle|^2 \\ &= \frac{1}{\Delta E} \sum_{p \neq n} \sum_i \langle \varphi_n | \bar{\mathbf{W}} | \varphi_p^i \rangle \langle \varphi_p^i | \bar{\mathbf{W}} | \varphi_n \rangle \\ &= \frac{1}{\Delta E} \langle \varphi_n | \bar{\mathbf{W}} \left[ \sum_{p \neq n} \sum_i |\varphi_p^i \rangle \langle \varphi_p^i| \right] \bar{\mathbf{W}} | \varphi_n \rangle. \end{aligned}$$

Since the energy basis is complete, we have

$$|\varphi_n \rangle \langle \varphi_n| + \sum_{p \neq n} \sum_i |\varphi_p^i \rangle \langle \varphi_p^i| = \mathbf{1},$$

so

$$\begin{aligned} |\varepsilon_2| &\leq \frac{1}{\Delta E} \langle \varphi_n | \bar{\mathbf{W}} [\mathbf{1} - |\varphi_n \rangle \langle \varphi_n|] \bar{\mathbf{W}} | \varphi_n \rangle \\ &= \frac{1}{\Delta E} \left[ \langle \varphi_n | \bar{\mathbf{W}}^2 | \varphi_n \rangle - \langle \varphi_n | \bar{\mathbf{W}} | \varphi_n \rangle^2 \right]. \end{aligned}$$

Hence,

$$|\lambda^2 \varepsilon_2| \leq \frac{(\Delta \mathbf{W})^2}{\Delta E}.$$

This gives us a bound on the second-order energy correction, in terms of  $\Delta E$  and the uncertainty of the perturbation  $\mathbf{W}$  with respect to the bare eigenstate  $|\varphi_n \rangle$ .

In particular,  $\lambda^2 \varepsilon_2 \rightarrow 0$  if  $|\varphi_n \rangle$  is an eigenstate of  $\mathbf{W}$ . Indeed, note that

$$(\mathbf{H}_0 + \mathbf{W})|\varphi_n \rangle = (E_n^0 + \lambda \varepsilon_1)|\varphi_n \rangle,$$

so all higher terms vanish from the expansion.

As an example, let's return to the spin- $\frac{1}{2}$  particle in a magnetic field. We consider an initial Hamiltonian

$$\mathbf{H}_0 = \omega_L \mathbf{S}_z$$

with eigenstates and eigenvalues

$$\mathbf{H}_0|+\rangle = +\frac{1}{2}\hbar\omega_L|+\rangle,$$

$$\mathbf{H}_0|-\rangle = -\frac{1}{2}\hbar\omega_L|-\rangle.$$

If the perturbation is

$$\mathbf{W} = \lambda\omega_L\mathbf{S}_x$$

$$= \frac{1}{2}\lambda\hbar\omega_L(|-\rangle\langle+| + |+\rangle\langle-|),$$

where  $\lambda = B_x/B_z \ll 1$ , then we can easily compute the leading-order energy corrections using the method we have just outlined:

$$E_n(\lambda) = E_n^0 + \langle\varphi_n|\mathbf{W}|\varphi_n\rangle + \sum_{p \neq n} \sum_i \frac{|\langle\varphi_p^i|\mathbf{W}|\varphi_n\rangle|^2}{(E_n^0 - E_p^0)} + O(\lambda^3).$$

For the correction to  $E_-^0$ , we have

$$E_-(\lambda) = -\frac{1}{2}\hbar\omega_L + \lambda\omega_L\langle-|\mathbf{S}_x|-\rangle + \frac{|\langle+|\lambda\omega_L\mathbf{S}_x|-\rangle|^2}{-\hbar\omega_L} + O(\lambda^3).$$

Since  $\mathbf{S}_x$  is purely off-diagonal, the first-order term vanishes and we are left with

$$E_-(\lambda) \approx E_-^0 + \frac{|\langle+|\mathbf{W}|-\rangle|^2}{E_-^0 - E_+^0}$$

$$= -\frac{1}{2}\hbar\omega_L - \lambda^2\frac{1}{4}\hbar\omega_L$$

$$= -\frac{1}{2}\hbar\gamma B_z \left(1 + \frac{1}{2}B_x^2B_z^{-2}\right)$$

$$\approx -\frac{1}{2}\hbar\gamma\sqrt{B_z^2 + B_x^2}.$$

Hence we can recover the expressions that we derived last time by Taylor-expanding the exact solution.

## Degenerate case

Now we suppose that  $E(\lambda)$  limits to a degenerate eigenvalue as  $\lambda$  goes to zero. Let  $g_n$  denote the 'degree of degeneracy' of  $E_n^0$ , by which we mean the number of times this eigenvalue is repeated in the spectrum of  $\mathbf{H}_0$ . Of course,  $g_n$  is also the dimension of the associated subspace of eigenstates, so

$$\mathbf{H}_0|\varphi_n^i\rangle = E_n^0|\varphi_n^i\rangle,$$

$$\langle\varphi_n^i|\varphi_n^j\rangle = \delta_{ij},$$

for  $i = 1..g_n$ .

Now in the perturbation expansions

$$E(\lambda) = \varepsilon_0 + \lambda\varepsilon_1 + \lambda^2\varepsilon_2 + O(\lambda^3),$$

$$|\Psi(\lambda)\rangle = |0\rangle + \lambda|1\rangle + \lambda^2|2\rangle + O(\lambda^2),$$

we can still set  $\varepsilon_0 = E_n^0$ , but it is no longer clear specifically how to choose  $|0\rangle$ . We can once again project the first recurrence relation onto the state  $\langle\varphi_n^i|$ ,

$$\begin{aligned}\langle \varphi_n^i | \mathbf{H}_0 | 1 \rangle + \langle \varphi_n^i | \bar{\mathbf{W}} | 0 \rangle &= \varepsilon_0 \langle \varphi_n^i | 1 \rangle + \varepsilon_1 \langle \varphi_n^i | 0 \rangle, \\ \langle \varphi_n^i | \bar{\mathbf{W}} | 0 \rangle &= \varepsilon_1 \langle \varphi_n^i | 0 \rangle.\end{aligned}$$

Here we have used  $\langle \varphi_n^i | \mathbf{H}_0 = \varepsilon_0 \langle \varphi_n^i |$ . Next we insert the closure relation for the  $\{|\varphi_p^i\rangle\}$  basis between  $\bar{\mathbf{W}}$  and the state  $|0\rangle$ , yielding

$$\sum_p \sum_{i'} \langle \varphi_n^i | \bar{\mathbf{W}} | \varphi_p^{i'} \rangle \langle \varphi_p^{i'} | 0 \rangle = \varepsilon_1 \langle \varphi_n^i | 0 \rangle.$$

Since  $|0\rangle$  lies inside the  $E_n^0$  subspace, only terms with  $p = n$  contribute to the sum on the LHS:

$$\sum_{i'} \langle \varphi_n^i | \bar{\mathbf{W}} | \varphi_n^{i'} \rangle \langle \varphi_n^{i'} | 0 \rangle = \varepsilon_1 \langle \varphi_n^i | 0 \rangle.$$

Next we must note that this summation defines an eigenvalue relation for  $|0\rangle$  and the **restriction** of  $\bar{\mathbf{W}}$  to the  $E_n^0$  subspace. One way to see this is to write out the summation in matrix form (over the indices  $i, i' = 1..g_n$ ):

$$\begin{pmatrix} \bar{\mathbf{W}}_{11} & \bar{\mathbf{W}}_{12} & \cdots & \bar{\mathbf{W}}_{1g_n} \\ \bar{\mathbf{W}}_{21} & \bar{\mathbf{W}}_{22} & \cdots & \bar{\mathbf{W}}_{2g_n} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\mathbf{W}}_{g_n 1} & \bar{\mathbf{W}}_{g_n 2} & \cdots & \bar{\mathbf{W}}_{g_n g_n} \end{pmatrix} \begin{pmatrix} \langle \varphi_n^1 | 0 \rangle \\ \langle \varphi_n^2 | 0 \rangle \\ \vdots \\ \langle \varphi_n^{g_n} | 0 \rangle \end{pmatrix} = w \begin{pmatrix} \langle \varphi_n^1 | 0 \rangle \\ \langle \varphi_n^2 | 0 \rangle \\ \vdots \\ \langle \varphi_n^{g_n} | 0 \rangle \end{pmatrix},$$

where  $\bar{\mathbf{W}}_{ii'} \equiv \langle \varphi_n^i | \bar{\mathbf{W}} | \varphi_n^{i'} \rangle$ . We denote the  $g_n \times g_n$  matrix of  $\bar{\mathbf{W}}_{ii'}$ 's as  $\bar{\mathbf{W}}^{(n)}$ , and note that the problem of determining  $|0\rangle$  corresponds precisely to that of diagonalizing  $\bar{\mathbf{W}}^{(n)}$ .

In general,  $\bar{\mathbf{W}}^{(n)}$  could still have degeneracies! Let's say that it has  $f_n$  distinct eigenvalues, where  $1 \leq f_n \leq g_n$ . Then corresponding to the given bare energy  $E_n^0$  there will be  $f_n$  distinct zeroth-order perturbed eigenstates corresponding to linear combinations of the states  $|\varphi_n^i\rangle$  :

$$|0_j\rangle = \sum_{i=1}^{g_n} c_{ij} |\varphi_n^i\rangle.$$

The coefficients  $c_{ij} = \langle \varphi_n^i | 0_j \rangle$  (with fixed  $j$  and  $i = 1..g_n$ ) will correspond to eigenvectors of the matrix  $\bar{\mathbf{W}}^{(n)}$ .

We are now finally ready to apply our first recurrence relation

$$\mathbf{H}_0 |1_j\rangle + \bar{\mathbf{W}} |0_j\rangle = \varepsilon_0 |1_j\rangle + \varepsilon_1 |0_j\rangle,$$

by projecting it onto the state  $\langle 0_j |$ :

$$\langle 0_j | \mathbf{H}_0 | 1_j \rangle + \langle 0_j | \bar{\mathbf{W}} | 0_j \rangle = \varepsilon_0 \langle 0_j | 1_j \rangle + \varepsilon_1.$$

Since  $|0_j\rangle$  still lies entirely within the  $E_n^0$  subspace  $\langle 0_j | \mathbf{H}_0 = E_n^0 \langle 0_j | = \varepsilon_0 \langle 0_j |$ . Thus,

$$\varepsilon_1 = \langle 0_j | \bar{\mathbf{W}} | 0_j \rangle = w_j,$$

where  $w_j$  is the eigenvalue of  $|0_j\rangle$  with respect to  $\bar{\mathbf{W}}^{(n)}$ , and

$$E_n^j(\lambda) = E_n^0 + \lambda w_j.$$

Let's consider the simple example

$$\mathbf{H}_0 = \varepsilon_0 \mathbf{O}_\alpha = \varepsilon_0 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

with

$$\mathbf{W} = \lambda \varepsilon_0 \mathbf{O}_\delta = \lambda \varepsilon_0 \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 2 \\ 1 & 2 & 1 \end{pmatrix},$$

where we would like to compute the zeroth-order eigenvector and first-order eigenvalue corrections in  $\lambda$  to the degenerate subspace  $E_n^0 = \varepsilon_0$ . This subspace is two-dimensional, and is spanned by the vectors

$$|\varphi_n^1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |\varphi_n^2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

The restriction of  $\mathbf{W}$  to this subspace is

$$\mathbf{W}^{(n)} = \begin{pmatrix} \bar{\mathbf{W}}_{11} & \bar{\mathbf{W}}_{12} \\ \bar{\mathbf{W}}_{21} & \bar{\mathbf{W}}_{22} \end{pmatrix} = \varepsilon_0 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

whose eigenvalues we know to be 0 and  $2\varepsilon_0$  with corresponding eigenvectors  $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \end{pmatrix}$  and  $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix}$ , respectively.

Thus, our degenerate perturbation theory tells us that  $\mathbf{W}$  lifts the degeneracy within the subspace corresponding to  $E_n^0 = \varepsilon_0$ , yielding

$$|0^1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad E_n^1 = \varepsilon_0$$

$$|0^2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad E_n^2 = \varepsilon_0 + 2\lambda\varepsilon_0.$$

Note that these are still only approximations to the true eigenvectors and eigenvalues of  $\mathbf{H} = \mathbf{H}_0 + \mathbf{W}$ , even though we have performed a diagonalization of  $\mathbf{W}^{(n)}$ . Looking at  $\mathbf{H}$ ,

$$\mathbf{H} = \varepsilon_0 \begin{pmatrix} 1 + \lambda & 0 & \lambda \\ 0 & 2 & 2\lambda \\ \lambda & 2\lambda & 1 + \lambda \end{pmatrix},$$

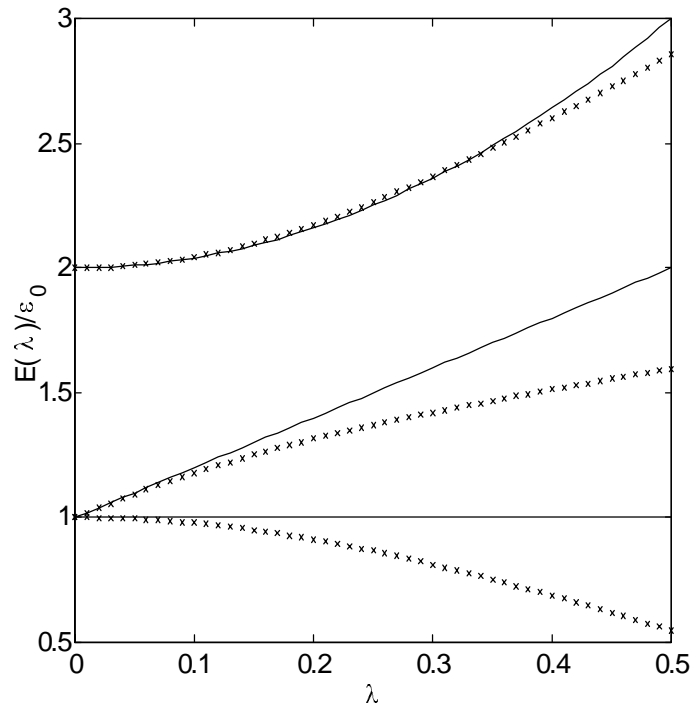
we see that our  $|0^1\rangle$  and  $|0^2\rangle$  are not true eigenstates because of the terms in  $2\lambda$ . The idea here is that we have only performed a ‘restricted’ diagonalization within a subspace, and therefore got away with diagonalizing a  $2 \times 2$  matrix instead of a  $3 \times 3$  one. In general, degenerate perturbation theory tells us how to compute leading-order corrections by

diagonalizing a  $g_n \times g_n$  matrix instead of the full  $N \times N$  Hamiltonian (where  $N$  is the dimension of the Hilbert space).

Just for completeness, let's continue by computing leading corrections to the other eigenvalue of  $\mathbf{H}_0$ . This is  $E_n^0 = 2\varepsilon_0$ , with corresponding eigenvector  $\langle \varphi_n | = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$ .

$$\begin{aligned}
 E_p(\lambda) &= E_n^0 + \langle \varphi_n | \mathbf{W} | \varphi_n \rangle + \sum_{p \neq n} \sum_i \frac{|\langle \varphi_p^i | \mathbf{W} | \varphi_n \rangle|^2}{(E_n^0 - E_p^0)} + O(\lambda^3) \\
 &= 2\varepsilon_0 + \lambda\varepsilon_0 \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 2 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\
 &\quad + \frac{\lambda^2 \varepsilon_0^2}{\varepsilon_0} \left| \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 2 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right|^2 \\
 &\quad + \frac{\lambda^2 \varepsilon_0^2}{\varepsilon_0} \left| \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 2 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right|^2 \\
 &= 2\varepsilon_0 + \lambda^2 \varepsilon_0 \left( \left| \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \right|^2 + \left| \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \right|^2 \right) \\
 &= 2\varepsilon_0 + 4\lambda^2 \varepsilon_0.
 \end{aligned}$$

We can now make a plot of the perturbative expressions for all three eigenvalues  $E(\lambda)$ , together with numerical calculations of the exact values:



Clearly, our second-order approximation stays valid for larger  $\lambda$  than our first-order approximations!