

## Chapter 8

# Rayleigh–Schrödinger Perturbation Theory

In this chapter we turn to an important practical and conceptual question in quantum theory, which is how to approximate the energy levels of systems that are *close to* being described by a system that we can solve. We will say a little (but not a lot) more below about what we mean by “close”, but intuitively, we can imagine (for example) modelling an atom with several electrons as being approximated by such an atom where the electrons do not interact among themselves except through their fermionic statistics. Alternatively, when we consider a Hydrogen-like, single-electron atom in the real world, we think of it as being well-described by the usual Hamiltonian with Coulomb potential, but there are actually small corrections to this due to the effects of (for instance) special relativity.

### 8.1 Formal perturbation theory

We imagine a scenario in which the Hamiltonian (for now we take this to be time-independent) of a quantum system can be written as a perturbation of a reference Hamiltonian  $H^{(0)}$ , so taking the form

$$H = H^{(0)} + \delta H. \quad (8.1)$$

We have in mind (though it need not be the case to develop the abstract formalism) that  $H^{(0)}$  is a Hamiltonian whose spectral problem we have understood exactly, such as that of the harmonic oscillator or the Hydrogen atom (restricting to bound states). Now we want to consider  $\delta H$  as being small in an appropriate sense, and we make this explicit by writing

$$\delta H = \varepsilon H^{(1)}, \quad (8.2)$$

where  $\varepsilon$  represents a small parameter, either in the sense of being numerically small (with  $H'$  being somehow fixed in size and so not scaling like an inverse power of  $\varepsilon$ ), or more accurately for our immediate purposes, in the sense of being *formally small*. So we will take our energy levels and stationary states to be formal series expansions in this parameter  $\varepsilon$ ,

$$\begin{aligned} \psi_\varepsilon &= \psi^{(0)} + \varepsilon \psi^{(1)} + \varepsilon^2 \psi^{(2)} + \dots + \varepsilon^n \psi^{(n)} + \dots, \\ E_\varepsilon &= E^{(0)} + \varepsilon E^{(1)} + \varepsilon^2 E^{(2)} + \dots + \varepsilon^n E^{(n)} + \dots, \end{aligned} \quad (8.3)$$

and ask that these satisfy the time-independent Schrödinger equation as formal series,

$$H\psi_\varepsilon = E_\varepsilon \psi_\varepsilon, \quad \text{order by order in } \varepsilon. \quad (8.4)$$

The issue of normalisation for this formal solution can be a little bit subtle, since the normalisation of our formal solution will itself in principle be a formal series in  $\varepsilon$ . In particular, given one formal series solution of (8.4), we can produce another by multiplying both sides by (the same) formal series in  $\varepsilon$  with numerical coefficients.

It turns out that a convenient normalisation is, rather than demanding  $\psi_\varepsilon$  be unit normalised, to require

$$\langle \psi^{(0)} | \psi_\varepsilon \rangle = 1, \quad (8.5)$$

which is equivalent to

$$\langle \psi^{(0)} | \psi^{(n)} \rangle = \delta_{n,0}. \quad (8.6)$$

It may be useful to ponder the enforcement of this normalisation if given a formal solution with an arbitrary normalisation.

Writing out (8.4) as an expansion in  $\varepsilon$ , we have,

$$\begin{aligned} \left( H^{(0)} + \varepsilon H^{(1)} \right) \left( \psi^{(0)} + \varepsilon \psi^{(1)} + \varepsilon^2 \psi^{(2)} + \dots \right) = \\ \left( E^{(0)} + \varepsilon E^{(1)} + \varepsilon^2 E^{(2)} + \dots \right) \left( \psi^{(0)} + \varepsilon \psi^{(1)} + \varepsilon^2 \psi^{(2)} + \dots \right), \end{aligned} \quad (8.7)$$

and equating coefficients of order  $\varepsilon^n$  gives an infinite sequence of relations

$$H^{(0)} \psi^{(n)} + H^{(1)} \psi^{(n-1)} = \sum_{m=0}^n E^{(m)} \psi^{(n-m)}. \quad (8.8)$$

The first of these relations, where  $n = 0$  (by convention we let  $\psi^{(-1)} = 0$ ) just tells us that zeroth term in each expansion corresponds to an eigenstate and energy level of the unperturbed system,

$$H^{(0)} \psi^{(0)} = E^{(0)} \psi^{(0)}. \quad (8.9)$$

The higher order terms will then be our object of focus. However, first we should make a short comment about the interpretation of this formal series.

*Remark 8.1.1 (Analytic considerations).* The analytic status of these formal series solutions is an interesting subject. Natural questions include whether we can assign some numerical value to  $\varepsilon$  so that the resultant series expansions for the energies and the eigenstates converge, and if so what the radius of convergence might be. The following theorem due to Kato gives a sense of under what conditions we can guarantee convergence.

*Theorem 8.1.2.* If there exist real constants  $a, b \geq 0$  such that for any  $\psi$  in the domain of  $H^{(0)}$ , we have

$$\|H^{(1)} \psi\| \leq a \|\psi\| + b \|H^{(0)} \psi\|, \quad (8.10)$$

then the formal series for  $\psi_\varepsilon$  and  $E_\varepsilon$  will have a finite radius of convergence.

Estimating this radius of convergence is a more subtle issue. Observe that for the case of a finite dimensional Hilbert space, this is always satisfied since all of the quantities appearing in (8.10) will be bounded. In the infinite dimensional case, the simplest scenario is again when the left hand side is simply bounded for all  $\psi \in \mathcal{H}$ ; in such a case there is always a non-zero radius of convergence.

The cases that are most often of interest in a physical setting are of the more complicated variety, where  $\mathcal{H}$  is infinite dimensional and the operators in question lead to unbounded left hand side of (8.10). In these cases establishing the bound in question might be a difficult problem. However, even when the series have zero radius of convergence, these expansions can often be interpreted as asymptotic series, and can be used to do computations that give good agreement with experiment (sometimes extraordinarily good agreement) by simply truncating the formal series after a fixed number of terms. In our applications this will be our *modus operandi*, and we will never have to grapple with the analytic subtleties of convergence.

## 8.2 First order perturbation theory (nondegenerate)

Now we return to our equation in (8.8) and look at the first correction term ( $m = 1$ ),

$$H^{(1)} \psi^{(0)} + H^{(0)} \psi^{(1)} = E^{(1)} \psi^{(0)} + E^{(0)} \psi^{(1)}. \quad (8.11)$$

Rewriting this gives

$$\left( H^{(0)} - E^{(0)} \right) \psi^{(1)} = - \left( H^{(1)} - E^{(1)} \right) \psi^{(0)}. \quad (8.12)$$

We want this equation to determine both  $E^{(1)}$  and  $\psi^{(1)}$ , given the unperturbed (zeroth order) information. For now we assume that  $E^{(0)}$  is a *nondegenerate energy level*, so  $\psi^{(0)}$  is the only state vector (up to rescaling) with eigenvalue  $E^{(0)}$ . An important practical fact is that we can determine the energy correction  $E^{(1)}$  without worrying about  $\psi^{(1)}$ . To do

this, we take the inner product of both sides of equation (8.12) with the state vector  $\psi^{(0)}$ ,

$$\langle \psi^{(0)} | (H^{(0)} - E^{(0)}) | \psi^{(1)} \rangle = - \langle \psi^{(0)} | (H^{(1)} - E^{(1)}) | \psi^{(0)} \rangle . \quad (8.13)$$

The left hand side vanishes, since  $H^{(0)}$  can act to the left and just give  $E^{(0)}$ , and we get the equation for our energy correction

$$E^{(1)} = \langle \psi^{(0)} | H^{(1)} | \psi^{(0)} \rangle . \quad (8.14)$$

This simple expression is extremely important and extraordinarily powerful; in a slogan it tells us that the leading correction to the energy of a stationary state is given by the expectation value of the perturbation in that state.

We then turn to the problem of determining  $\psi^{(1)}$ . We will find an expression for the expansion of this correction vector in terms of a basis of stationary state vectors for the unperturbed Hamiltonian, which we denote by  $\{\psi_n^{(0)}\}$ , where

$$H^{(0)} | \psi_n^{(0)} \rangle = E_n^{(0)} | \psi_n^{(0)} \rangle . \quad (8.15)$$

We label these states so that the state whose perturbation we are studying is the  $m$ 'th state ( $n = m$ ), with energy  $E_m^{(0)}$ . We write such an expansion of the correction vector using a resolution of the identity,

$$| \psi_m^{(1)} \rangle = \sum_n | \psi_n^{(0)} \rangle \langle \psi_n^{(0)} | \psi_m^{(1)} \rangle = \sum_{n \neq m} \left( \langle \psi_n^{(0)} | \psi_m^{(1)} \rangle \right) | \psi_n^{(0)} \rangle . \quad (8.16)$$

where the  $n = m$  term drops out of the summation due to our normalisation condition (8.5). Now we can compute the coefficients by taking the inner product of both sides of (8.12) with  $\psi_n^{(0)}$ , which yields

$$\begin{aligned} \langle \psi_n^{(0)} | (H^{(0)} - E_m^{(0)}) | \psi_m^{(1)} \rangle &= - \langle \psi_n^{(0)} | (H^{(1)} - E_m^{(1)}) | \psi_m^{(0)} \rangle , \\ \implies (E_n^{(0)} - E_m^{(0)}) \langle \psi_n^{(0)} | \psi_m^{(1)} \rangle &= - \langle \psi_n^{(0)} | H^{(1)} | \psi_m^{(0)} \rangle , \\ \implies \langle \psi_n^{(0)} | \psi_m^{(1)} \rangle &= \frac{\langle \psi_n^{(0)} | H^{(1)} | \psi_m^{(0)} \rangle}{E_m^{(0)} - E_n^{(0)}} . \end{aligned} \quad (8.17)$$

This allows us to write a general expression for the first correction to the  $m$ th energy eigenstate,

$$\psi_m^{(1)} = \sum_{n \neq m} \left( \frac{\langle \psi_n^{(0)} | H^{(1)} | \psi_m^{(0)} \rangle}{E_m^{(0)} - E_n^{(0)}} \right) \psi_n^{(0)} . \quad (8.18)$$

*Remark 8.2.1.* There is a nice way to understand this expression for the first correction to energy eigenstates that is somewhat slicker and provides a nice heuristic that we will reuse later. In principle, we would like to take (8.12) and solve for  $\psi_m^{(1)}$  by *inverting* the operator appearing on the left hand side,

$$| \psi_m^{(1)} \rangle = \frac{-1}{H^{(0)} - E_m^{(0)}} (H^{(1)} - E_m^{(1)}) | \psi_m^{(0)} \rangle . \quad (8.19)$$

However, the operator  $H^{(0)} - E_m^{(0)}$  is definitely *not* invertible, since it has a kernel (spanned by  $\psi_m^{(0)}$  in this non-degenerate case) so is not injective. A related fact is that the range of this operator doesn't include all of  $\mathcal{H}$ ;  $\psi_m^{(0)}$  is absent, and indeed it is the only one of our basis vectors that is absent.

Consequently, we can only hope to define this inverse operator on the subspace of  $\mathcal{H}$  that is orthogonal to  $\psi_m^{(0)}$ , and even then its action can only be defined only up to the possible addition of multiples of  $\psi_m^{(0)}$ . The first of these requirements gives us a solvability condition: the right hand side of (8.12) must lie in the orthogonal complement to  $\psi_m^{(0)}$ . This is precisely the condition that we used to determine  $E_m^{(1)}$ . The second issue means that we must adopt a convention for the component of  $\psi_m^{(1)}$  in the direction of  $\psi_m^{(0)}$ . This is exactly what our normalisation condition (8.5) accomplishes.

With both of those issues square away, we can get an expression for  $\psi_m^{(1)}$  upon inserting a resolution of the identity for the orthogonal complement  $\mathcal{H}^\perp$  to  $\psi_m^{(0)}$ . (This must act as the identity given that we have ensured that the state vector on the right of (8.12) is lies in  $\mathcal{H}^\perp$ .)

$$\begin{aligned} |\psi_m^{(1)}\rangle &= \frac{-1}{H^{(0)} - E_m^{(0)}} \sum_{n \neq m} |\psi_n^{(0)}\rangle \langle \psi_n^{(0)} | (H^{(1)} - E_m^{(1)}) |\psi_m^{(0)}\rangle, \\ &= \sum_{n \neq m} \left( \frac{\langle \psi_n^{(0)} | H^{(1)} | \psi_m^{(0)} \rangle}{E_m^{(0)} - E_n^{(0)}} \right) |\psi_n^{(0)}\rangle. \end{aligned} \quad (8.20)$$

In passing to the second line, we have defined the action of the inverted operator using the usual rule that on eigenvectors of an observable  $A$  with eigenvalue  $a$ , we let a function of the observable  $f(A)$  act by  $f(a)$ . We observe that this (admittedly formal) procedure neatly reproduces the more piecemeal analysis carried out above.

**Example 8.2.2** (Energy of the Helium atom ground state). A classic application of first order perturbation theory is to the study of a di-electronic atom (an atom with two electrons) like Helium. For now we will ignore the effects of spin and fermionic statistics. We are therefore studying wavefunctions  $\psi(\mathbf{x}_1, \mathbf{x}_2)$  of two position vectors, with a Hamiltonian of the form

$$H = H_1 + H_2 + H_{\text{int}}, \quad (8.21)$$

where

$$H_i = -\frac{\hbar^2}{2m} \nabla_i^2 - \frac{Zq_e^2}{|\mathbf{x}_i|}, \quad i = 1, 2, \quad H_{\text{int}} = \frac{q_e^2}{|\mathbf{x}_1 - \mathbf{x}_2|}. \quad (8.22)$$

Here the nuclear charge  $Z$  would be two for Helium,  $H_1$  and  $H_2$  are the standard single-electron Hamiltonians including a Coulomb potential from the nucleus, and  $H_{\text{int}}$  is the *interaction Hamiltonian* that encodes the repulsion between the two electrons.

If we could ignore the interaction term and had as our Hamiltonian just  $H_1 + H_2$ , then this system would be the tensor product of two copies of the Hydrogen atom system, and we would have stationary states that were separable,

$$\psi_{n_1, n_2}(\mathbf{x}_1, \mathbf{x}_2) = \psi_{n_1}(\mathbf{x}_1) \psi_{n_2}(\mathbf{x}_2). \quad (8.23)$$

In particular, the ground state is non-degenerate and given by<sup>48</sup>

$$\begin{aligned} \psi_{1,1}(\mathbf{x}_1, \mathbf{x}_2) &= \left( \frac{Z^3}{\pi a_0^3} \right)^{\frac{1}{2}} \exp\left(-\frac{Zr_1}{a_0}\right) \times \left( \frac{Z^3}{\pi a_0^3} \right)^{\frac{1}{2}} \exp\left(-\frac{Zr_2}{a_0}\right), \\ &= \left( \frac{Z^3}{\pi a_0^3} \right) \exp\left(-\frac{Z(r_1 + r_2)}{a_0}\right). \end{aligned} \quad (8.24)$$

Though it is certainly not clear that the interaction Hamiltonian can be thought of as being *small*, we can nevertheless proceed with a formal perturbative analysis setting  $\delta H = H_{\text{int}}$ .<sup>49</sup> The first order correction to the ground state energy using our machinery from above is then given by the expectation value,

$$E_1^{(1)} = \langle \psi^{(0)} | H_{\text{int}} | \psi^{(0)} \rangle = \left( \frac{Z^3}{\pi a_0^3} \right)^2 \int_{\mathbb{R}^6} \frac{q_e^2 \exp\left(\frac{-2Z(r_1 + r_2)}{a_0}\right)}{|\mathbf{x}_1 - \mathbf{x}_2|} d^3\mathbf{x}_1 d^3\mathbf{x}_2. \quad (8.25)$$

<sup>48</sup>In more detail, not ignoring spin, the ground state for a single electron is two dimensional and can be represented as  $\psi_1(\mathbf{x})(\alpha |\frac{1}{2}\rangle + \beta |-\frac{1}{2}\rangle)$  with  $\alpha, \beta$  some arbitrary complex numbers. With two particles in the ground state, we cannot make the wave function antisymmetric in the positions  $\mathbf{x}_1$  and  $\mathbf{x}_2$  because the ground state is unique, but we can make the total state antisymmetric using the spins, giving the two-particle ground state

$$|\psi\rangle = \psi_1(\mathbf{x}_1)\psi_1(\mathbf{x}_2) \left( |\frac{1}{2}; -\frac{1}{2}\rangle - |-\frac{1}{2}; \frac{1}{2}\rangle \right),$$

where  $|\frac{1}{2}; -\frac{1}{2}\rangle$  is the state in which the first particle has intrinsic spin  $m_s = \frac{1}{2}$  and the second  $m_s = -\frac{1}{2}$ , etc.. At the level of spins, or course, this is a simple case of addition of angular momentum where  $\mathcal{H}_{1/2} \otimes \mathcal{H}_{1/2} = \mathcal{H}_1 \oplus \mathcal{H}_0$  and we take the (unique) spin-0 state.

<sup>49</sup>One way one could contrive a more convincing setting for this perturbative analysis would be to take the large  $Z$  limit. Then the Coulomb potential due to the nucleus becomes arbitrarily strong compared to the inter-electronic interaction, so that the approximation of ignoring the interaction seems plausible.

Evaluating integrals like these is (unfortunately) largely unavoidable when it comes time to turn the abstract algebra of perturbation theory into actual numbers for systems like the Helium atom. We can perform this integral by first noting that spherical symmetry allows us to fix  $\mathbf{x}_1$  to point in the  $\hat{z}$  direction (we then pick up a factor of  $4\pi$  from the angular integration for  $\mathbf{x}_1$ ). Introducing spherical polar coordinates for  $\mathbf{x}_2$ , we then have

$$\mathbf{x}_1 \cdot \mathbf{x}_2 = r_1 r_2 \cos \theta, \quad |\mathbf{x}_1 - \mathbf{x}_2| = (r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta)^{1/2}. \quad (8.26)$$

The  $\varphi$  integral for  $\mathbf{x}_2$  is trivial and gives an extra factor of  $2\pi$ . Thus the original six-dimensional integral reduces to

$$\left( \frac{8Z^6 q_e^2}{a_0^6} \right) \int_0^\infty \int_0^\infty \int_0^\pi \exp\left( \frac{-2Z(r_1 + r_2)}{a_0} \right) \frac{r_1^2 r_2^2 \sin \theta}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta}} d\theta dr_1 dr_2. \quad (8.27)$$

The angular integral can be done immediately,

$$\int_0^{2\pi} \frac{\sin \theta d\theta}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta}} = \frac{1}{r_1 r_2} \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta} \Big|_{\theta=0}^{\theta=\pi}. \quad (8.28)$$

$$= \frac{1}{r_1 r_2} ((r_1 + r_2) - |r_1 - r_2|), \quad (8.29)$$

$$= \begin{cases} \frac{2}{r_1} & r_1 > r_2, \\ \frac{2}{r_2} & r_2 > r_1. \end{cases} \quad (8.30)$$

By symmetry the integration regions where  $r_1 > r_2$  and that with  $r_2 > r_1$  contribute equally, so we can write just take twice the part where  $r_1 > r_2$  and the resulting radial integral becomes

$$E_1^{(1)} = \left( \frac{32Z^6 q_e^2}{a_0^6} \right) \int_0^\infty \int_{r_2}^\infty \exp\left( -\frac{2Z(r_1 + r_2)}{a_0} \right) r_1 r_2^2 dr_1 dr_2. \quad (8.31)$$

This leaves radial integrations, and these can be performed with the help of the following useful integral identity

**Lemma 8.2.3.** For any non-negative integer  $n$ , we have

$$\int_R^\infty e^{-kr} r^n dr = \frac{n! e^{-kR}}{k^{n+1}} \sum_{j=0}^n \frac{(kR)^j}{j!}. \quad (8.32)$$

This is proved by induction on  $n$  by differentiating with respect to  $k$ .

Applying this lemma twice in (8.31) gives the final result,

$$E_1^{(1)} = \frac{5}{8} \frac{Z q_e^2}{a_0}. \quad (8.33)$$

Selecting  $Z = 2$  as is appropriate for Helium and performing a crude truncation of the perturbation series to first order, this gives an estimation for the ground state energy as

$$E_1 \approx E_1^{(0)} + E_1^{(1)} = -\frac{2q_e^2}{a_0} \left( 2 - \frac{5}{8} \right) = -\frac{2.75q_e^2}{a_0}. \quad (8.34)$$

The experimental result for the Helium atom's ground state energy is

$$E_1^{(\text{exp})} \approx -\frac{2.92q_e^2}{a_0}. \quad (8.35)$$

Though we don't have systematic control of higher order corrections, it is gratifying to see that the estimate with a coefficient of 2.75 is much closer to the experimental value than the zeroth order estimate with a coefficient of 4.

### 8.3 First order perturbation theory (degenerate)

We need to be wary of the case where the energy level whose corrections we are considering is degenerate in the unperturbed theory. For example, we may be considering a Hydrogen-like ion where the energy levels all have some degeneracy. In this case, the analysis of the previous subsection needs to be improved a bit.

This is most readily observed by considering the calculation that yielded the first correction to the energy. Suppose that both  $\psi^{(0)}$  and  $\varphi^{(0)}$  both have unperturbed energy  $E^{(0)}$ . Taking the inner product of both sides of (8.12) with both  $\psi^{(0)}$  and  $\varphi^{(0)}$ , we find

$$E^{(1)} = \langle \psi^{(0)} | H^{(1)} | \psi^{(0)} \rangle, \quad 0 = \langle \varphi^{(0)} | H^{(1)} | \psi^{(0)} \rangle. \quad (8.36)$$

The first of these is the same relation we had before and could be thought to determine the first energy correction, but the second may or may not be satisfied depending on whether we have been lucky in our choice of basis vectors  $\psi^{(0)}$  and  $\varphi^{(0)}$ . Indeed, the problem is precisely that there is an ambiguity in the basis that we choose for the degenerate  $E^{(0)}$  eigenspace, whereas after perturbation there is a preferred choice of basis. This resolution proceeds as follows.

**Theorem 8.3.1.** Let  $\varphi_1, \dots, \varphi_d$  be an orthonormal basis for the  $E^{(0)}$  eigenspace of  $H^{(0)}$ . Then (8.12) can be solved if and only if  $E^{(1)}$  and  $\psi^{(0)}$  are chosen so that  $E^{(1)}$  is a solution to

$$\det \left( \langle \varphi_r | H^{(1)} | \varphi_s \rangle - E^{(1)} \delta_{rs} \right) = 0, \quad (8.37)$$

i.e.,  $E^{(1)}$  should be an eigenvalue of the matrix whose with entries  $\langle \varphi_r | H^{(1)} | \varphi_s \rangle$ , and  $\psi^{(0)} = \sum_{r=1}^d c_r \varphi_r$  is the corresponding eigenvector

$$\sum_s \langle \varphi_r | H^{(1)} | \varphi_s \rangle c_s = E^{(1)} c_r. \quad (8.38)$$

In a slogan, this says that before setting up the first order perturbation problem, one should choose a basis for the degenerate  $E^{(0)}$  eigenspace that diagonalises the action of the restriction of  $H^{(1)}$  to that subspace. In this basis, the rules are the same as in the non-degenerate setting.

*Proof.* This follows directly from the requirement that we avoid the potential contradictions arising as in (8.36), but we will argue more abstractly along the lines of Remark 8.2.1.

We need to be able to solve (8.12), but again  $H^{(0)} - E^{(0)}$  has a (now  $d$ -dimensional) kernel and is not invertible, and we need to make our choices to ensure that the right hand side lies in the range of  $H^{(0)} - E^{(0)}$ . There is a helpful proposition whose proof will not be important, though we give most of it (omitting a technical detail).

**Proposition 8.3.2.** The range  $\text{Ran}(A)$  of a self-adjoint operator  $A$  on a Hilbert space  $\mathcal{H}$  (for which zero is not an accumulation point of the spectrum) coincides with the orthogonal complement of its kernel.

*Proof.* We first establish that  $\text{Ran}(A) \subseteq \ker(A)^\perp$ . Let  $\psi \in \text{Ran}(A)$ , so we can write  $\psi = A\psi'$ . Then we have for any  $\phi \in \ker(A)$ ,

$$\langle \phi | \psi \rangle = \langle \phi | A\psi' \rangle = \langle A\phi | \psi' \rangle = 0, \quad (8.39)$$

which gives the inclusion. We next establish that  $\text{Ran}(A)^\perp \subseteq \ker(A)$ . First, note that for  $\phi \in \text{Ran}(A)^\perp$ , we have that for all  $\psi \in \mathcal{H}$ , we have

$$0 = \langle \phi | A\psi \rangle = \langle A\phi | \psi \rangle, \quad (8.40)$$

so  $A\phi$  must be the zero vector, and thus  $\phi \in \ker(A)$ , and taking orthogonal complements,  $\ker(A)^\perp \subseteq (\text{Ran}(A)^\perp)^\perp$ . All that remains is make the identification<sup>50</sup>

$$(\text{Ran}(A)^\perp)^\perp \stackrel{?}{=} \text{Ran}(A). \quad (8.41)$$

It turns out this is not true automatically; it requires the operator  $A$  to have a *closed range*. This, in turn, follows (via a straightforward but uninteresting-for-our-purposes proof) from the condition that zero not be an accumulation point in the spectrum of  $A$ . With both inclusions in place, we get then recover the stated identification. ■

<sup>50</sup>It should be a familiar fact that the iterated orthogonal complement is the identity for finite-dimensional inner product spaces; the novelty here is that in infinite dimensions it isn't necessarily true.

Thus what we need to ensure is that the the right hand side of (8.12) is orthogonal to the operator's kernel. Taking the inner product with any one of the  $\phi_r$  we find,

$$0 = \langle \phi_r | (H^{(1)} - E^{(1)}) | \psi^{(0)} \rangle = \sum_s \langle \phi_r | H^{(1)} | \phi_s \rangle c_s - E^{(1)} \delta_{rs} c_s, \quad (8.42)$$

where  $\psi^{(0)}$  is expanded in terms of the  $\phi_i$  as in the statement of the theorem. This is precisely the condition that  $E^{(1)}$  be an eigenvalue of the matrix with matrix elements  $\langle \phi_r | H^{(1)} | \phi_s \rangle$  with  $\psi^{(0)}$  (thought of as an element of the  $d$ -dimensional  $E^{(0)}$  eigenspace) the corresponding eigenvector. ■

In practice, what one should usually do is to choose the basis  $\phi_1, \dots, \phi_d$  to already diagonalise the restriction to the  $E^{(0)}$  eigenspace of the perturbation  $H^{(1)}$  (i.e., to be the eigenvectors appearing in the above theorem). Using these as the unperturbed state vectors when setting up the formal perturbation problem (say, setting  $\psi^{(0)} = \phi_1$ ) then gives for the first correction to such an eigenstate,

$$\psi^{(1)} = \sum_{E_n^{(0)} \neq E^{(0)}} \left( \frac{\langle \psi_n^{(0)} | H^{(1)} | \psi^{(0)} \rangle}{E^{(0)} - E_n^{(0)}} \right) | \psi_n^{(0)} \rangle + \sum_{j=2}^d \lambda_j | \phi_j \rangle, \quad (8.43)$$

where the parameters  $\lambda_j$  parameterise the ambiguity in defining the preimage of the right hand side of (8.12) due to the enlarged kernel of  $H^{(0)} - E^{(0)}$ . At this stage these parameters represent genuine ambiguities. You will see on the third problem sheet that these parameters are fixed upon continuing to higher orders in perturbation theory.

**Example 8.3.3** (Strong field Zeeman effect). A simple example of degenerate perturbation theory arises in what is known as the Zeeman effect (or Zeeman splitting) for Hydrogen-like atoms. If we apply a constant magnetic field  $\mathbf{B}$  to such an atom, there turns out to be a correction to the Hamiltonian of the following form,

$$\delta H = \frac{q_e}{2m_e c} \mathbf{B} \cdot (\mathbf{L} + g_e \mathbf{S}). \quad (8.44)$$

Here,  $g_e \approx 2$  is a numerical constant known as the *gyromagnetic ratio* of the electron. Without loss of generality, we will take  $\mathbf{B}$  to be oriented in the  $x_3$ -direction.

When considering the effect of this perturbation, we must account for the degeneracy of the Hydrogen energy levels, and as we understood in the previous chapter, there are at least two natural bases for the degenerate energy eigenspaces of the Hydrogen atom, depending on whether we want to diagonalise the total angular momentum ( $|n, \ell, j, m_j\rangle$  states) or the  $x_3$  component of both orbital and spin angular momentum ( $|n, \ell, m_\ell; m_s\rangle$  states). For this perturbation, then, we should use the latter states since they are actually already simultaneous eigenvectors of  $L_3$  and  $S_3$ .

If we restrict to states with principle quantum number  $n$  (so unperturbed energy  $E_n$ ), then the first order corrections are given by

$$E_{n,\ell,m_\ell,m_s}^{(1)} = \langle n, \ell, m_\ell; m_s | \delta H | n, \ell, m_\ell; m_s \rangle = \frac{q_e B}{2m_e c} (m_\ell + g_e m_s). \quad (8.45)$$

The above example is a bit artificial as an example of perturbation theory, because the  $|n, \ell, m_\ell; m_s\rangle$  states actually are *exact eigenstates* of the perturbed Hamiltonian, as they diagonalise the  $J_3$  and  $S_3$  operators appearing in  $\delta H$  on the nose. Correspondingly, you should be able to convince yourself that the correction (8.43) to the energy eigenstates vanishes exactly. Therefore this analysis is valid for *large* values of  $B$ . In fact, it is *only* valid for sufficiently large values of  $B$ , because for small  $B$  there are competing corrections that need to be accounted for.

#### 8.4 Higher order perturbation theory

We will briefly comment on the extension of this approach to higher orders in the perturbative expansion. To begin, let us just proceed directly to second order, where the  $O(\varepsilon^2)$  term in (8.7) amounts to the equation

$$H^{(0)} \psi^{(2)} + H^{(1)} \psi^{(1)} = E^{(0)} \psi^{(2)} + E^{(1)} \psi^{(1)} + E^{(2)} \psi^{(0)}. \quad (8.46)$$

We reorganise this suggestively as

$$\left(H^{(0)} - E^{(0)}\right) \psi^{(2)} = -\left(H^{(1)} - E^{(1)}\right) \psi^{(1)} + E^{(2)} \psi^{(0)}, \quad (8.47)$$

which again tells us that to be able to determine  $\psi^{(2)}$  we have to find the preimage under  $H^{(0)} - E^{(0)}$  of the vector on the right hand side. If for now we assume that the unperturbed energy level in question is non-degenerate, then as at first order, our first task will be to establish a solvability condition that the right hand side is orthogonal to the kernel of  $H^{(0)} - E^{(0)}$ , which is to say that it is orthogonal to  $\psi^{(0)}$ . This determines  $E^{(2)}$ :

$$E^{(2)} = \left\langle \psi^{(0)} \left| \left( H^{(1)} - E^{(1)} \right) \psi^{(1)} \right\rangle = \left\langle \psi^{(0)} \left| H^{(1)} \psi^{(1)} \right\rangle, \quad (8.48)$$

which, plugging in our expression for  $\psi^{(1)}$ , gives

$$E^{(2)} = \sum_{\psi_n^{(0)} \neq \psi^{(0)}} \frac{\langle \psi^{(0)} | H^{(1)} | \psi_n^{(0)} \rangle \langle \psi_n^{(0)} | H^{(1)} | \psi^{(0)} \rangle}{E^{(0)} - E_n^{(0)}} = \sum_{\psi_n^{(0)} \neq \psi^{(0)}} \frac{|\langle \psi_n^{(0)} | H^{(1)} | \psi^{(0)} \rangle|^2}{E^{(0)} - E_n^{(0)}}. \quad (8.49)$$

Let us make a few comments.

- If  $\psi^{(0)}$  is the unperturbed ground state, then the denominator in (8.49) is always negative and so the second order energy correction is necessarily negative.
- This formula is especially useful if  $H^{(1)} \psi^{(0)}$  is itself an energy eigenstate, so that we only get one nonvanishing term in the sum due to the orthogonality of the different energy eigenstates.
- If energy levels are well separated, then we expect the terms corresponding to mixing with nearby energy eigenstates to dominate the correction due to suppression in the denominator by energy difference.

Continuing to the second-order correction to the energy eigenstate (subject to our normalisation condition), we will have an expansion,

$$\psi^{(2)} = \sum_{\psi_n^{(0)} \neq \psi^{(0)}} \left\langle \psi_n^{(0)} \left| \psi^{(2)} \right\rangle \psi_n^{(0)}, \quad (8.50)$$

and we can solve for the coefficients using (8.46). Ultimately, we find<sup>51</sup>

$$\begin{aligned} \psi^{(2)} = \sum_{\psi_{n,m}^{(0)} \neq \psi^{(0)}} \frac{\langle \psi_n^{(0)} | H^{(1)} | \psi_m^{(0)} \rangle \langle \psi_m^{(0)} | H^{(1)} | \psi^{(0)} \rangle}{\left(E_n^{(0)} - E^{(0)}\right) \left(E_m^{(0)} - E^{(0)}\right)} \psi_n^{(0)} - \\ \sum_{\psi_n^{(0)} \neq \psi^{(0)}} \frac{\langle \psi_n^{(0)} | H^{(1)} | \psi^{(0)} \rangle \langle \psi^{(0)} | H^{(1)} | \psi^{(0)} \rangle}{\left(E_n^{(0)} - E^{(0)}\right)^2} \psi_n^{(0)}. \end{aligned} \quad (8.52)$$

Clearly, at this point (and even more at higher orders) things get quite cluttered in these general expressions.

Often it is more important to have expressions for higher order corrections to energy levels than it is to have the high order expressions for the actual state vectors. From our general perturbative expansion, this would naively require obtaining  $\psi^{(k-1)}$  if we want to compute  $E^{(k)}$  because expanding  $H_\varepsilon \psi_\varepsilon = E_\varepsilon \psi_\varepsilon$  to order  $k$  gives from the coefficient of  $\varepsilon^k$

$$H^{(0)} \psi^{(k)} + H^{(1)} \psi^{(k-1)} = E^{(0)} \psi^{(k)} + E^{(1)} \psi^{(k-1)} + \dots + E^{(k)} \psi^{(0)}, \quad (8.53)$$

<sup>51</sup>An alternate way to organise this calculation is to write

$$\psi^{(2)} = -\frac{1}{\left(H^{(0)} - E^{(0)}\right)} \Pi_{\mathcal{H}_{E^{(0)}}^\perp} \left( H^{(1)} - E^{(1)} \right) \psi^{(1)}, \quad (8.51)$$

where the projection operator  $\Pi_{\mathcal{H}_{E^{(0)}}^\perp}$  serves to eliminate components in the direction of  $\psi^{(0)}$ , which is the result of appropriately fixing the value of  $E^{(2)}$ .



and taking the inner product with  $\psi^{(0)}$  (imposing the solvability condition in the nondegenerate case), using  $\langle \psi^{(0)} | \psi^{(l)} \rangle = 0$  that follows from the normalization condition, we obtain

$$E^{(k)} = \left\langle \psi^{(0)} \left| H^{(1)} \right| \psi^{(k-1)} \right\rangle, \quad (8.54)$$

In fact, one can avoid calculating all the way up to  $\psi^{(k-1)}$  using a clever trick that draws out relations between the different perturbative corrections.

**Lemma 8.4.1.** As a formal expansion in  $\varepsilon_1$  and  $\varepsilon_2$ , we have

$$\left\langle \psi_{\varepsilon_1} \left| H^{(1)} \right| \psi_{\varepsilon_2} \right\rangle = \frac{E_{\varepsilon_1} - E_{\varepsilon_2}}{\varepsilon_1 - \varepsilon_2} \langle \psi_{\varepsilon_1} | \psi_{\varepsilon_2} \rangle. \quad (8.55)$$

*Proof.* We have  $H_{\varepsilon_1} - H_{\varepsilon_2} = (\varepsilon_1 - \varepsilon_2)H^{(1)}$  so we can compute,

$$\langle \psi_{\varepsilon_1} | H_{\varepsilon_1} - H_{\varepsilon_2} | \psi_{\varepsilon_2} \rangle = \langle \psi_{\varepsilon_1} | H_{\varepsilon_1} | \psi_{\varepsilon_2} \rangle - \langle \psi_{\varepsilon_1} | H_{\varepsilon_2} | \psi_{\varepsilon_2} \rangle, \quad (8.56)$$

$$(\varepsilon_1 - \varepsilon_2) \langle \psi_{\varepsilon_1} | H^{(1)} | \psi_{\varepsilon_2} \rangle = E_{\varepsilon_1} \langle \psi_{\varepsilon_1} | \psi_{\varepsilon_2} \rangle - E_{\varepsilon_2} \langle \psi_{\varepsilon_1} | \psi_{\varepsilon_2} \rangle. \quad (8.57)$$

which leads to the conclusion. ■

Note that  $E_{\varepsilon_1} - E_{\varepsilon_2}$  is proportional to  $\varepsilon_1 - \varepsilon_2$ , so the right hand side can still be taken as a formal expansion in the two variables.

By expanding both sides of (8.55), and considering each monomial in the  $\varepsilon$ 's, we generate various identities for the perturbative corrections. Indeed, our first naive higher order formula (8.54) represents the coefficient of  $\varepsilon_2^{k-1}$ . However, we can be a bit more efficient if we are clever.

**Corollary 8.4.2.**  $E^{(2k+1)}$  can be expressed in terms of only the  $\psi^{(r)}$  with  $r \leq k$  and their matrix elements with respect to  $H^{(1)}$ .

*Proof.* The coefficient on the right hand side of (8.55) can be expanded as

$$\frac{E_{\varepsilon_1} - E_{\varepsilon_2}}{\varepsilon_1 - \varepsilon_2} = \sum_{n=1}^{\infty} \frac{\varepsilon_1^n - \varepsilon_2^n}{\varepsilon_1 - \varepsilon_2} E^{(n)} \quad (8.58)$$

$$= \sum_{n=1}^{\infty} E^{(n)} \left( \sum_{j=0}^{n-1} \varepsilon_1^j \varepsilon_2^{n-1-j} \right), \quad (8.59)$$

so at order  $\varepsilon_1^k \varepsilon_2^k$  we have  $E^{(2k+1)}$ . This is then equal to the coefficient of  $\varepsilon_1^k \varepsilon_2^k$  in the expansion of

$$\frac{\langle \psi_{\varepsilon_1} | H^{(1)} | \psi_{\varepsilon_2} \rangle}{\langle \psi_{\varepsilon_1} | \psi_{\varepsilon_2} \rangle}. \quad (8.60)$$

That term clearly depends only on  $\psi^{(r)}$  for  $r \leq k$ . ■

The simplest example of this gives us an economical expression for the third order energy correction,

$$E^{(3)} = \left\langle \psi^{(1)} \left| H^{(1)} \right| \psi^{(1)} \right\rangle - E^{(1)} \left\langle \psi^{(1)} \left| \psi^{(1)} \right\rangle, \quad (8.61)$$

which we can calculate using only the results from our previous first-order computations.