## Chapter 9

# Variational Methods

It isn't uncommon to encounter a system that is not particularly close to being solvable, rendering perturbation theory at least unreliable and at worst completely unhelpful. There are an important class of techniques that go by the name of *variational methods* that can be applied quite generally in these circumstances, though they often require a bit of creativity to exploit well. In this chapter we present the essential idea behind these methods and study (again!) the Helium atom as a nice example.

### 9.1 Rayleigh quotients for observables

The main tool behind our variational methods will be the so-called *Rayleigh quotient.*

**Definition 9.1.1** (Rayleigh Quotient)**.** For an observable $A$, the (real) function

$$\begin{aligned} f_A : \mathcal{H} &\longrightarrow \mathbb{R}\,, \\ \psi &\longmapsto \mathbb{E}_\psi(A) = \frac{\langle\psi|A|\psi\rangle}{\langle\psi|\psi\rangle}\,. \end{aligned} \tag{9.1}$$

is known as the *Rayleigh quotient* for $A$.

It may seem a bit overwrought to introduce a new name for what we already know as the expectation value of $A$ in the state $\psi$. The point here is to accentuate the fact that this is now being thought of as a (smooth) function on $\mathcal{H}$, which is not necessarily how we contextualised expectation values of observables previously.

A key property of the Rayleigh quotient is that the stationary values/vectors for $f_A(\psi)$ are precisely eigenvalue/vectors for $A$. This can be demonstrated as follows.

**Lemma 9.1.2.** Given a linear subspace $\mathcal{K} \subseteq \mathcal{H}$, then

$$\frac{\mathrm{d}}{\mathrm{d}t} f_A(\psi + t\phi)\bigg|_{t=0} = 0\,, \quad \forall\phi\in\mathcal{K} \quad \Longleftrightarrow \quad \langle\phi|A - f_A(\psi)|\psi\rangle = 0 \quad \forall\phi\in\mathcal{K}\,. \tag{9.2}$$

*Proof.* For $\phi \in \mathcal{K}$, also $i\phi \in \mathcal{K}$, so stationarity with respect to adding $\phi$ implies

$$\frac{\mathrm{d}}{\mathrm{d}t} f_A(\psi + t\phi) = 0 \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}t} f_A(\psi + it\phi) = 0\,. \tag{9.3}$$

By direct computation, the first of these relations gives

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\langle\psi+t\phi|A|\psi+t\phi\rangle}{\langle\psi+t\phi|\psi+t\phi\rangle}\bigg|_{t=0}\,, \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\langle\psi|A|\psi\rangle + t\,\langle\phi|A|\psi\rangle + t\,\langle\psi|A|\phi\rangle + O(t^2)}{\langle\psi|\psi\rangle + t\,\langle\phi|\psi\rangle + t\,\langle\psi|\phi\rangle + O(t^2)}\bigg|_{t=0}\,, \\ &= \frac{\langle\phi|A|\psi\rangle + \langle\psi|A|\phi\rangle}{\langle\psi|\psi\rangle} - \frac{\langle\psi|A|\psi\rangle\,(\langle\phi|\psi\rangle + \langle\psi|\phi\rangle)}{\langle\psi|\psi\rangle^2}\,, \\ &= 2\Re\left(\frac{\langle\phi|(A - f_A(\psi))|\psi\rangle}{\langle\psi|\psi\rangle}\right)\,. \end{aligned}$$

This gives the real part of the desired equation. If we repeat the argument with $\phi \to i\phi$ we get the imaginary part as well, and thus deduce that for states that are stationary against adding vectors $\phi \in \mathcal{K}$,

$$\langle\phi|(A - f_A(\psi))|\psi\rangle = 0\,. \tag{9.4}$$

Thus $(A - f_A(\psi))\,|\psi\rangle \in \mathcal{K}^{\perp}$. ■

If we set $\mathcal{K} = \mathcal{H}$, then $\mathcal{K}^{\perp} = \{0\}$ and we have the strong relation

$$(A - f_A(\psi))\,|\psi\rangle = 0\,. \tag{9.5}$$

This means that the critical points of $f_A(\psi)$ are precisely the eigenvectors of $A$ and the critical values are the eigenvalues of $A$.

### 9.2 The virial theorem

Using our analysis of the Rayleigh quotient, we can already obtain useful, general information about energy eigenstates by considering variations within one-parameter families of states. An important example is the following important theorem, which generalises an analogous result in classical mechanics.

**Theorem 9.2.1** (Virial theorem). Assume the Hamiltonian for a quantum mechanical system whose Hilbert space is identified with a space of wave functions in $d$ dimensions has the conventional form $H = T + V$ with

$$T = \frac{\mathbf{P}^2}{2m} = -\frac{\hbar^2}{2m}\nabla^2\,, \qquad V = V(\mathbf{X})\,. \tag{9.6}$$

Then for any stationary state $\psi$ (so $H\psi = E\psi$), the following condition holds:

$$2\mathbb{E}_{\psi}(T) = \mathbb{E}_{\psi}(\mathbf{x}\cdot\nabla V)\,. \tag{9.7}$$

If $V$ is homogeneous of degree $N$ (*i.e.*, $V(\lambda\mathbf{x}) = \lambda^N V(\mathbf{x})$), then we have the following stronger result,

$$\mathbb{E}_{\psi}(T) = \frac{N}{N+2}E\,, \qquad \mathbb{E}_{\psi}(V) = \frac{2}{N+2}E\,. \tag{9.8}$$

*Proof.* The idea of the proof is to perform a variational analysis for the family of wave functions of the form $\psi_\lambda(\mathbf{x}) = \lambda^{d/2}\psi(\lambda\mathbf{x})$ for a given reference wave function $\psi(\mathbf{x})$. The factor of $\lambda^{d/2}$ is included to ensure that all of these wave functions are normalised equally; this is only done for convenience so that in computing Rayleigh quotients we can ignore the denominator.

Now suppose that $\psi(\mathbf{x})$ is a (normalised) stationary state; then the Rayleigh quotient of $\psi_\lambda(\mathbf{x})$ must be stationary as a function of $\lambda$ at $\lambda = 1$,[52]

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} f_H(\psi_\lambda)\bigg|_{\lambda=1} = 0\,. \tag{9.9}$$

Proceeding by direct computation, we have by the chain rule $\nabla\psi_\lambda(\mathbf{x}) = \lambda^{\frac{d+2}{2}}\left(\nabla\psi\right)(\lambda\mathbf{x})$, which gives us for the Rayleigh quotient,

$$\begin{aligned} f_H(\psi_\lambda) &= \int\limits_{\mathbb{R}^d} \left(\frac{-\hbar^2\lambda^{d+2}}{2m}\,|(\nabla\psi)(\lambda\mathbf{x})|^2 + \lambda^d V(\mathbf{x})|\psi(\lambda\mathbf{x})|^2\right)\mathrm{d}^d\mathbf{x}\,, \\ &= \int\limits_{\mathbb{R}^d} \left(\frac{-\lambda^2\hbar^2}{2m}|\nabla'\psi(\mathbf{x}')| + V(\lambda^{-1}\mathbf{x}')|\psi(\mathbf{x}')|^2\right)\mathrm{d}^d\mathbf{x}'\,, \\ &= \lambda^2\mathbb{E}_{\psi}(T) + \mathbb{E}_{\psi}(V(\lambda^{-1}\mathbf{x}))\,. \end{aligned} \tag{9.10}$$

In the passing to the second line we have defined $\mathbf{x}' = \lambda\mathbf{x}$. With this, (9.9) gives the condition

$$0 = 2\mathbb{E}_{\psi}(T) + \frac{\mathrm{d}}{\mathrm{d}\lambda}\mathbb{E}_{\psi}(V(\lambda^{-1}\mathbf{x}))\bigg|_{\lambda=1} = 2\mathbb{E}_{\psi}(T) - \mathbb{E}_{\psi}(\mathbf{x}\cdot\nabla V(\mathbf{x}))\,, \tag{9.11}$$

---

[52] Note that while Lemma (9.1.2) is formulated for linear subspaces of $\mathcal{H}$, the conclusion means that in any parameterised family of states, the stationary states must give critical points.

which reproduces (9.7). For $V$ homogeneous of degree $N$, $\mathbf{x}\cdot\nabla V = NV$ and we get the simpler result

$$2\mathbb{E}_{\psi}(T) = N\mathbb{E}_{\psi}\left(V(\mathbf{x})\right) . \tag{9.12}$$

We also have for an energy eigenstate

$$E = \mathbb{E}_{\psi}(T) + \mathbb{E}_{\psi}(V(\mathbf{x})) , \tag{9.13}$$

and putting these together gives

$$\mathbb{E}_{\psi}(T) = \frac{N}{N+2}E , \qquad \mathbb{E}_{\psi}(V) = \frac{2}{N+2}E , \tag{9.14}$$

as required. ■

*Remark* 9.2.2. We offer some additional comments here.

- We see that for the Coulomb potential, for which $N = -1$, we must have $E < 0$ because $V < 0$ and $\mathbb{E}_{\psi}(V)$ is twice the size of $\mathbb{E}_{\psi}(T)$. (Alternatively, $T$ is a positive operator, so the fact that $\mathbb{E}_{\psi}(T) = -E$ gives the negativity of $E$.)
- For the harmonic oscillator, kinetic and potential energies are equal and balanced $\mathbb{E}_{\psi}(V) = \mathbb{E}_{\psi}(T) = \frac{1}{2}E$.
- This *quantum virial theorem* has a classical counterpart, which is the original virial theorem. The classical theorem has to do with time-averages of energies along classical trajectories.

### 9.3 Approximating the ground state

The flagship application of variational methods in quantum theory is to the study of the ground state and ground state energy of complicated quantum systems. To this end, one observes the following.

**Proposition 9.3.1.** If $f_H$ is bounded below and achieves its minimum $E_0 := \inf_{\mathbb{P}(\mathcal{H})} f_H$, then $E_0$ is the ground state energy (minimum eigenvalue) and any state $\psi$ for which $f_H(\psi) = E_0$ is a ground state.

*Proof.* $f_H$ is automatically stationary at its minimum, and this will necessarily correspond to an eigenstate. Its eigenvalue will be the minimal one because the other eigenvalues are realised as values of $f_H$ as well. ■

Conversely, when a system does have a normalisable ground state $\psi_0$, the function $f_H$ achieves its lower bound at $\psi_0$.

Though this is a relatively simple observation, this result is actually very powerful. The idea is that we can try to find an approximate ground state by finding the minimum $E_{0,\text{approx}}$ of $f_H$ restricted to some cleverly chosen subset of $\mathcal{H}$. We are guaranteed that $E_{0,\text{approx}} \geqslant E_0$ by the above, so this procedure produces rigorous upper bounds for $E_0$. If we are lucky and creative, these upper bounds will also be good approximations (though to estimate errors would be beyond the scope of the discussion here).

**Example 9.3.2** (Helium again). We return to the two-electron Helium atom, with Hamiltonian

$$H = \frac{\mathbf{P}_1^2}{2m} + \frac{\mathbf{P}_2^2}{2m} - 2q_e^2\left(\frac{1}{|\mathbf{x}_1|} + \frac{1}{|\mathbf{x}_2|}\right) + \frac{q_e^2}{|\mathbf{x}_1 - \mathbf{x}_2|} = T + V + \delta H , \tag{9.15}$$

where

$$T = \frac{\mathbf{P}_1^2}{2m} + \frac{\mathbf{P}_2^2}{2m} , \quad V = -2q_e^2\left(\frac{1}{|\mathbf{x}_1|} + \frac{1}{|\mathbf{x}_2|}\right) , \quad \delta H = \frac{q_e^2}{|\mathbf{x}_1 - \mathbf{x}_2|} . \tag{9.16}$$

As we saw in the previous chapter, we can try to treat $\delta H$ as a small correction and use first-order perturbation theory to get decent results for the ground state energy. However, the approximation in that case seemed quite uncontrolled; $\delta H$ doesn't actually seem particularly negligible compared to other terms in the potential. This is a perfect case in which to attempt a variational estimate, and indeed we can do quite well.

**Proposition 9.3.3.** The Helium ground state energy $E_0$ is bounded above according to

$$E_0 \leqslant -\left(\frac{27}{16}\right)^2 \frac{q_e^2}{a_0} \approx 2.85\frac{q_e^2}{a_0}\,. \tag{9.17}$$

This should be compared to our estimate of $-2.75q_e^2/a_0$ from first-order perturbation theory and the experimental value of approximately $-2.92q_e^2/a_0$; the variational upper bound is well below the first order estimate, and indeed is within three percent of the experimental value.

*Proof.* We derive this using a one-parameter family of wavefunctions like we did with the virial theorem. To choose our *variational Ansatz*, we use a physical argument. The idea is that if we want to model the dynamics of two electrons in the Helium atom as being non-interacting, we should adjust the effective value of the nuclear charge to account for screening, *i.e.*, each electron should on average see less than the full charge of the nucleus since the other electron is producing an electric field with the opposite sign. In other circumstances, this kind of an approximation is sometimes called a *mean field approximation*, and it can be quite effective.

To put this idea into practice, we adopt trial wave functions of the form

$$\psi_Z(\mathbf{x}_1,\mathbf{x}_2) = \left(\frac{Z^3}{\pi a_0^3}\right)\exp\left(-\frac{Z(r_1+r_2)}{a_0}\right)\,, \tag{9.18}$$

where $Z$ is the parameter we will minimise over. This is the exact ground state for the effective Hamiltonian

$$H_Z = \frac{\mathbf{P}_1^2}{2m} + \frac{\mathbf{P}_2^2}{2m} - Zq_e^2\left(\frac{1}{r_1}+\frac{1}{r_2}\right) = T + \frac{Z}{2}V\,, \tag{9.19}$$

satisfying $H_Z\psi_Z = E_Z\psi_Z$ with $E_Z = -Z^2q_e^2/a_0$, so indeed these are the ground state wave functions for systems of non-interacting electrons with an adjustable nuclear charge. We then want to minimise the Rayleigh quotient associated to the true Hamiltonian $H$ evaluated on these states as a function of $Z$, so we need to compute

$$f_H(\psi_Z) = \mathbb{E}_{\psi_Z}(T) + \frac{2}{Z}\mathbb{E}_{\psi_Z}\left(\frac{ZV}{2}\right) + \mathbb{E}_{\psi_Z}(\delta H)\,. \tag{9.20}$$

The first two terms can be evaluated using the virial theorem with respect to the effective Hamiltonian $H_Z$, which gives us

$$\mathbb{E}_{\psi_Z}(T) = -E_Z = \frac{Z^2q_e^2}{a_0}\,, \qquad \mathbb{E}_{\psi_Z}\left(\frac{ZV}{2}\right) = 2E_Z = -\frac{2Z^2q_e^2}{a_0}\,. \tag{9.21}$$

The last term in (9.20) is precisely what we computed in our first-order perturbation theory calculation previously, and gives

$$\langle\psi_Z|\delta H|\psi_Z\rangle = q_e^2\left\langle\psi_Z\middle|\frac{1}{|\mathbf{x}_1-\mathbf{x}_2|}\middle|\psi_Z\right\rangle = \frac{5}{8}\frac{Zq_e^2}{a_0}\,. \tag{9.22}$$

Putting everything together, we have

$$f_H(\psi_Z) = \frac{q_e^2}{a_0}\left(Z^2 - \frac{27}{8}Z\right) = \frac{q_e^2}{a_0}\left[\left(Z-\frac{27}{16}\right)^2 - \left(\frac{27}{16}\right)^2\right]\,. \tag{9.23}$$

As a function of $Z$, this is minimised at $Z = \frac{27}{16}$, reflecting the extent to which each electron shields the charge of the nucleus to the other. From this we extract the upper bound

$$E_0 \leqslant f_H(\psi_{\frac{27}{16}}) = -\frac{q_e^2}{a_0}\left(\frac{27}{16}\right)^2\,. \tag{9.24}$$

■

As we noted above, this is not only an upper bound but turns out to be a very good estimate for the ground state energy.

The accuracy can be improved by considering a more general Ansatz for trial wave functions—indeed, an accuracy to within three parts in $10^8$ has been obtained for this calculation using a family of *393 basis functions*.

The above result represents a marked improvement over what we got using elementary perturbation theory. The fact that we do better is not a coincidence at all. Indeed, we have the following.

**Proposition 9.3.4.** Let $H = H_0 + \delta H$ as before and let the family of states $\{\psi_\lambda\}$ over which we will minimise be chosen so that $\psi_0^{(0)}$ (the ground state of $H_0$) is contained in the family. Let $E_0^{(0)} + E_0^{(1)}$ be the first order perturbation theoretic estimate of the ground state energy found before. Then for $E_{\text{var}} = \inf_{\{\psi_\lambda\}} f_H$ we have

$$E_0^{(0)} + E_0^{(1)} \geqslant E_{\text{var}} \geqslant E_{\text{ground}} \,. \tag{9.25}$$

*Proof.* Recall that $E_0^{(1)} = \left\langle \psi_0^{(0)} \middle| \delta H \middle| \psi_0^{(0)} \right\rangle$ so that

$$E_0^{(0)} + E_0^{(1)} = \left\langle \psi_0^{(0)} \middle| H_0 + \delta H \middle| \psi_0^{(0)} \right\rangle = f_H\left(\psi_0^{(0)}\right) \,, \tag{9.26}$$

and since $\psi_0^{(0)} \in \{\psi_\lambda\}$, this is an upper bound for the infimum of $f_H$ on $\{\psi_\lambda\}$. ■

## 9.4 Approximating excited states

Though the ground state energy tends to be of particular interest, we may also want to approximate the energies of excited states. If we were to have perfect knowledge of the first $k$ eigenstates $\psi_0, \ldots \psi_{k-1}$, with energies $E_0, E_1, \ldots, E_{k-1}$, say, then it would be straightforward to use the same variational ideas as we did above to approximate the $k+1$st state. Letting $\mathcal{H}_k = \text{Span}\{\psi_0, \ldots, \psi_{k-1}\}$, we would have

**Proposition 9.4.1.** If $\inf_{\mathcal{H}_k^\perp} f_H$ is attained for some $\psi_k \in \mathcal{H}_k^\perp$, then this is the $k+1$st lowest energy eigenstate, and $f_H(\psi_k) \geqslant E_j, j = 0, \ldots, k-1$.

*Proof.* Since $H : \mathcal{H}_k \to \mathcal{H}_k$, self adjointness implies that $H : \mathcal{H}_k^\perp \to \mathcal{H}_k^\perp$. Applying the main variational result for $f_H$ on $\mathcal{H}_k^\perp$ gives that if $f_H$ achieves its infimum at $\psi_k$, then $(H - f_H(\psi_k)) \left|\psi_k\right\rangle = 0$ and $\psi_k$ is the eigenstate with the lowest eigenvalue in $\mathcal{H}_k^\perp$. Since the lowest $k$ eigenvalues are in $\mathcal{H}_k$, this must be the $k+1$st. ■

If this was where things stopped, we would be in the unfortunate situation of needing to have an exact result for lower lying states before approximating the higher ones. It turns out that we can proceed without knowledge of the first $k$ eigenvectors/eigenvalues using the following important theorem.

**Theorem 9.4.2** (Min-max). If the infimum

$$\inf\{\max\{f_H(\psi)\,,\ \ \psi \in \mathcal{K}\}\,,\ \ \dim \mathcal{K} = k\} \tag{9.27}$$

is attained, where the infimum is taken over $k$-dimensional subspaces $\mathcal{K} \subset \mathcal{H}$, then it is the $k$th lowest eigenvalue and the state on which it is attained the corresponding eigenvector. Thus, the maximum value of $f_H$ on $\mathcal{K}$ is an upper bound for the $k$th energy level of the system.

We will give the proof for the case where $\mathcal{H}$ has a basis of normalisable energy eigenstates. In the general case (involving generalised eigenstates), both the statement of the theorem and the proof become slightly more technical.

*Proof.* Let $E_0 \leqslant E_1 \leqslant E_2 \leqslant \ldots$ be the ordered energy eigenvalues and let $\psi_0, \psi_1, \ldots$ be the corresponding energy eigenstates. Let $\mathcal{H}_{k-1} = \text{Span}\{\psi_0, \ldots, \psi_{k-2}\}$. Now for any $k$-dimensional subspace $\mathcal{K} \subset \mathcal{H}$, we will have $\dim\left(\mathcal{K} \cap \mathcal{H}_{k-1}^\perp\right) \geqslant 1$, so we can find a (normalised) vector $\psi \in \mathcal{K}$ that can be expressed in terms of the $\psi_n$ for $n \geqslant k-1$.

$$\psi = \sum_{n=k-1}^{\dim \mathcal{H}} a_n \psi_n \,, \qquad \sum_{n=k-1}^{\dim \mathcal{H}} |a_n|^2 = 1 \,. \tag{9.28}$$

For which we have

$$f_H(\psi) = \sum_{n=k-1}^{d} |a_n|^2 E_n \geqslant \sum_{n=k-1}^{d} |a_n|^2 E_{k-1} = E_{k-1} \,, \tag{9.29}$$

so $\max_{\psi \in \mathcal{K}} \{f_H(\psi)\} \geqslant E_{k-1}$.

On the other hand, for $\mathcal{K} = \operatorname{span}\{\psi_0, \psi_1, \ldots, \psi_{k-1}\}$ we have that $\max_{\psi \in \mathcal{K}} f_H(\psi) = E_{k-1}$, and the result follows. ■

By considering families of choices of $\mathcal{K}$ we can try to improve the bound. Indeed, this leads to a nice approximation scheme. To do so, we choose trial subspaces $\mathcal{K}$ (perhaps varying with respect to some parameters), and for a given trial subspace of dimension $k$, the maximum of $f_H$ will be the largest eigenvalue of the restriction/projection of $H$ to $\mathcal{K}$ (*i.e.*, $H_{\mathcal{K}} = \Pi_{\mathcal{K}} \circ H \circ \Pi_{\mathcal{K}}$). So we can produce bounds (and potentially estimates) for $E_{k-1}$ by solving finite-dimensional eigenvalue problems, *i.e.*, by solving (perhaps numerically) for the largest roots of the characteristic polynomials of $k \times k$ matrices. This is known as the *Rayleigh–Ritz method.*

**Example 9.4.3** (Angular momentum with Rayleigh–Ritz)**.** The variational methods introduced in this section can be used not just for the Hamiltonian of a system, but for any self-adjoint operator with bounded-below spectrum. Indeed, consider the case of a particle moving on the sphere (the *rigid rotor*). The total angular momentum operator in spherical polar coordinates (and atomic units: $\hbar = 1$), has the form

$$L^2 = -\frac{1}{\sin^2\theta}\left(\sin\theta \frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{\partial^2}{\partial\varphi^2}\right) . \tag{9.30}$$

We can obtain Rayleigh–Ritz estimates for the first two eigenvalues of $L^2$ by taking, as a basis for a space of trial functions $\psi_1(\theta, \varphi) = 1$ and $\psi_2(\theta, \varphi) = \cos^2\theta$. We then need to solve the characteristic equation for the projection of the $L^2$ operator to this trial space,

$$\det\left(\left\langle \psi_j \middle| L^2 \middle| \psi_k \right\rangle - \lambda \left\langle \psi_j \middle| \psi_k \right\rangle\right) = 0 \,. \tag{9.31}$$

Clearly $L^2\psi_1 = 0$, and we calculate that

$$\begin{aligned} L^2\psi_2 &= -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\cos^2\theta\right) = \frac{2}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin^2\theta\cos\theta\right) \,, \\ &= \frac{2}{\sin\theta}\left(2\sin\theta\cos^2\theta - \sin^3\theta\right) = 2\left(2\cos^2\theta - \sin^2\theta\right) \,, \\ &= 2\left(3\cos^2\theta - 1\right) = 2\left(3\psi_2 - \psi_1\right) \,. \end{aligned} \tag{9.32}$$

We also need the following normalisations and inner products,

$$\begin{aligned} \|\psi_1\|^2 &= \int \sin\theta \; \mathrm{d}\theta \, \mathrm{d}\phi = 4\pi \,, \\ \|\psi_2\|^2 &= \int \cos^4\theta \sin\theta \; \mathrm{d}\theta \, \mathrm{d}\phi = 2\pi\left(-\frac{1}{5}\cos^5\theta\right)\Bigg|_0^\pi = \frac{4\pi}{5} \,, \\ \langle\psi_1|\psi_2\rangle &= \int \cos^2\theta \sin\theta \; \mathrm{d}\theta \, \mathrm{d}\phi = 2\pi\left(-\frac{1}{3}\cos^3\theta\right)\Bigg|_0^\pi = \frac{4\pi}{3} \,. \end{aligned} \tag{9.33}$$

We clearly have $\left\langle \psi_j \middle| L^2\psi_1 \right\rangle = 0 = \left\langle \psi_1 \middle| L^2\psi_j \right\rangle$, so the only remaining element is

$$\langle\psi_2|L^2\psi_2\rangle = \langle\psi_2|6\psi_2 - 2\psi_1\rangle = \frac{24\pi}{5} - \frac{8\pi}{3} = \frac{32\pi}{15}. \tag{9.34}$$

Our eigenvalue equation therefore gives

$$0 = \begin{vmatrix} -4\lambda\pi & -\frac{4\pi\lambda}{3} \\ -\frac{4\pi\lambda}{3} & \frac{32\pi}{15} - \frac{4\pi\lambda}{5} \end{vmatrix} = \frac{62\pi^2}{45}\lambda(\lambda - 6) \,. \tag{9.35}$$

so $\lambda = 0$ or $\lambda = 6$, which correspond to $\ell = 0$ and $\ell = 2$, respectively, in the standard formula $\ell(\ell + 1)$ for the eigenvalues of $L^2$ (with $\hbar = 1$). The first case gives the eigenvector $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, that is $\psi_1$; the second gives $\left(\begin{smallmatrix} -1 \\ 3 \end{smallmatrix}\right)$ or $3\psi_2 - \psi_1$. We recognise these as precisely (up to normalisation) the spherical harmonics $Y_0^0$ and $Y_2^0$.

While we have ended up landing on exact eigenstates/eigenvalues, we didn't get the first two, but rather the first and fifth lowest eigenstates of $L^2$—we missed the three $\ell = 1$ states. This is, of course, compatible with our eigenvalues being upper bounds for the first two eigenvalues. However, in our trial basis we explicitly chose only $\varphi$-independent functions, so we have effecively enfoced by hand that we are only studying states with $m = 0$. In the $m = 0$ sector, we have ended up with the first and *third* eigenstates. This is frequently a useful trick in applying Rayleigh–Ritz; you can specialise to a subspace of the Hilbert space with some definite behaviour with respect to other symmetries and then avoid worrying about many lower-energy states that are orthogonal to that subspace.