A117

Pre-reads:
A1
A2
A4
A13
A71
A101

How do we derive the corrections in perturbation theory?

How do we split the eigenvalue equation for H into a hierarchy of equations?

In this article, we will derive the expressions for the first-order corrections in time-independent perturbation theory introduced in A71.

We wish to find approximate eigenvalues of the Hamiltonian

$$\hat{H} = \hat{H}_0 + \lambda\hat{V}$$

where the parameter λ is used to keep track of the perturbation in $\hat{V}$. We shall assume that the Hamiltonians $\hat{H}$ and $\hat{H}_0$ are non-degenerate, meaning that every energy eigenstate corresponds to a different energy eigenvalue. It is possible to generalise the following arguments to the degenerate case, where more than one eigenstate has the same energy eigenvalue, but this will not be covered in these articles.

To derive the corrections in perturbation theory, we will first derive a hierarchy of equations by writing the eigenvalue equation for $\hat{H}$,

$$(\hat{H}_0 + \lambda\hat{V})\,|\phi_j\rangle = E_j\,|\phi_j\rangle \tag{1}$$

in terms of power series for E_j and $|\phi_j\rangle$. As introduced in A71, we write these in power series with respect to λ:

$$E_j = \sum_{k=0}^{\infty} \lambda^k E_j^k = E_j^0 + \lambda E_j^1 + \lambda^2 E_j^2 + \cdots$$

$$|\phi_j\rangle = \sum_{k=0}^{\infty} \lambda^k \left|\phi_j^k\right\rangle = \left|\phi_j^0\right\rangle + \lambda\left|\phi_j^1\right\rangle + \lambda^2\left|\phi_j^2\right\rangle + \cdots .$$

If we substitute these power series into Equation 1, we get

$$(\hat{H}_0 + \lambda\hat{V})\sum_{k=0}^{\infty}\lambda^k \left|\phi_j^k\right\rangle = \sum_{l=0}^{\infty}\lambda^l \left|\phi_j^l\right\rangle \sum_{m=0}^{\infty}\lambda^m \left|\phi_j^m\right\rangle$$

which we can rewrite as

$$\hat{H}_0\sum_{k=0}^{\infty}\lambda^k \left|\phi_j^k\right\rangle + \hat{V}\sum_{k=0}^{\infty}\lambda^{k+1} \left|\phi_j^k\right\rangle = \sum_{l,m=0}^{\infty}\lambda^{l+m}E_j^l \left|\phi_j^m\right\rangle \ . \tag{2}$$

The principle of perturbation theory is that we should be able to vary the strength of the perturbation by varying λ, and the eigenvalues and eigenvectors should vary continuously in turn. If λ is a free parameter, then for Equation 2, a hierarchy of equations for each individual power of λ must be independently satisfied.

Therefore, by considering terms proportional to different powers of λ one at a time, we derive a hierarchy of equations.

For example, by considering the terms of Equation 2 proportional to λ^0 (constant terms in λ), we derive the zeroth-order equation

$$\hat{H}_0\left|\phi_j^0\right\rangle = E_j^0\left|\phi_j^0\right\rangle$$

which is the eigenvalue equation for $\hat{H}_0$, reflecting the solutions obtained when λ tends to zero.

The first-order equation is achieved by selecting terms proportional to λ^1:

$$\lambda\hat{H}_0\left|\phi_j^1\right\rangle + \lambda\hat{V}\left|\phi_j^0\right\rangle = \lambda E_j^0\left|\phi_j^1\right\rangle + \lambda E_j^1\left|\phi_j^0\right\rangle$$

which (because $\lambda \neq 0$) we can simplify by dividing by λ to

$$\hat{H}_0\left|\phi_j^1\right\rangle + \hat{V}\left|\phi_j^0\right\rangle = E_j^0\left|\phi_j^1\right\rangle + E_j^1\left|\phi_j^0\right\rangle \tag{3}$$

The first-order corrections can all be derived from this first-order equation.

Similarly, by selecting terms in Equation 2 proportional to higher powers of λ, higher-order equations can be derived that allow us to compute higher-order corrections.

How do we derive the first-order corrections to the eigenvalues?

Our starting point for deriving the first-order eigenvalue correction E_j^1 is first-order Equation 3. If we multiply this equation from the right by $\left\langle\phi_j^0\right|$, we obtain

$$\left\langle\phi_j^0\right|\hat{H}_0\left|\phi_j^1\right\rangle + \left\langle\phi_j^0\right|\hat{V}\left|\phi_j^0\right\rangle = \left\langle\phi_j^0\right|E_j^0\left|\phi_j^1\right\rangle + \left\langle\phi_j^0\right|E_j^1\left|\phi_j^0\right\rangle$$

which, using $\left\langle\phi_j^0\right|\hat{H}_0 = E_j^0\left\langle\phi_j^0\right|$, simplifies to

$$E_j^0 \langle \phi_j^0 | \phi_j^1 \rangle + \langle \phi_j^0 | \hat{V} | \phi_j^0 \rangle = E_j^0 \langle \phi_j^0 |^0 | \phi_j^1 \rangle + E_j^1$$

which, cancelling the identical terms on both sides, simplifies to our final expression:

$$\langle \phi_j^0 | V | \phi_j^0 \rangle = E_j^1 \,.$$

How do we compute the first-order corrections to the eigenvectors?

The states $\left|\phi_j^0\right\rangle$ form an orthonormal basis, which we will use to express the first-order correction to the eigenvector $\left|\phi_j^1\right\rangle$. Our starting point is, again, Equation 3. This time, we multiply it from the left by $\left\langle \phi_k^0 \right|$, where $k \neq j$.

$$\langle \phi_k^0 | \hat{H}_0 | \phi_j^1 \rangle + \langle \phi_k^0 | \hat{V} | \phi_j^0 \rangle = \langle \phi_k^0 | E_j^0 | \phi_j^1 \rangle + \langle \phi_k^0 | E_j^1 | \phi_j^0 \rangle \,.$$

Because $\langle \phi_k^0 \left| \phi_j^0 \right\rangle = 0$, and using $\langle \phi_k^0 | \hat{H}_0 = E_k^0 \left\langle \phi_j^0 \right|$, this simplifies to

$$E_k^0 \langle \phi_k^0 | \phi_j^1 \rangle + \langle \phi_k^0 | \hat{V} | \phi_j^0 \rangle = E_j^0 \langle \phi_k^0 | \phi_j^1 \rangle$$

and hence

$$\langle \phi_k^0 | \phi_j^1 \rangle = \frac{\langle \phi_k^0 | \hat{V} \left| \phi_j^0 \right\rangle}{E_j^0 - E_k^0} \,.$$

This provides almost all the information we need to write $\left|\phi_j^1\right\rangle$ in the basis provided by $|\phi_k^0\rangle$. In the above we assumed that $k \neq j$, so the final coefficient we need to compute is the $k = j$ case, i.e. $\langle \phi_j^0 \left| \phi_j^1 \right\rangle$.

As it turns out, it is possible to show that we have a free choice of the value of $\langle \phi_j^0 \left| \phi_j^1 \right\rangle$. Changing this quantity is equivalent to multiplying the state $|\phi_j\rangle$ by a global phase. Changing the global phase of a state has no measurable consequences, so we conventionally choose the most convenient value for this quantity and set $\langle \phi_j^0 \left| \phi_j^1 \right\rangle = 0$.

Therefore we can express $\left|\phi_j^1\right\rangle$ as follows:

$$|\phi_j^1\rangle = \sum_k |\phi_k^0\rangle \langle \phi_k^0 | \phi_j^1 \rangle = \sum_{k \neq j} \frac{\langle \phi_k^0 | \hat{V} \left| \phi_j^0 \right\rangle}{E_j^0 - E_k^0} |\phi_k^0\rangle \,.$$

Further Reading

- B.H. Bransden and C.J. Joachain, Quantum Mechanics, Prentice Hall (2000), chapter 8.1.
- C. Cohen-Tannoudji, B. Diu and F. Laloe, Quantum Mechanics, Wiley-VCH (1992), Chapter 11.
- D. Bohm, Quantum Theory, Dover (1989), chapter 18.