Section 9
Variational Method
Lecture 27:
The Variational Method

Date Given: 2008/12/03
Date Revised: 2008/12/03
Motivation

It is not always possible to find an analytic solution to the Schrödinger Equation. One can always solve the equation numerically, but this is not necessarily the best way to go; one may not be interested in the detailed eigenfunctions, but rather only in the energy levels and the qualitative features of the eigenfunctions. And numerical solutions are usually less intuitively understandable.

Fortunately, one can show that the values of the energy levels are only mildly sensitive to the deviation of the wavefunction from its true form, and so the expectation value of the energy for an approximate wavefunction can be a very good estimate of the corresponding energy eigenvalue. By using an approximate wavefunction that depends on some small set of parameters and minimizing its energy with respect to the parameters, one makes such energy estimates. The technique is called the variational method because of this minimization process.

This technique is most effective when trying to determine ground state energies, so it serves as a nice complement to the WKB approximation, which works best when one is interested in relatively highly excited states, ones whose deBroglie wavelength is short compared to the distance scale on which the wavelength changes.
The Hamiltonian's Eigenstates and Eigenvalues are Stationary Points of the Energy Functional

For any wavefunction $\psi$, we may calculate the expectation value of the energy,

$$E[\psi] = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\int dx \, \psi^*(x) H \psi(x)}{\int dx \, \psi^*(x) \psi(x)}$$  \hspace{1cm} (9.1)

(We will explain below why we need to explicitly include the normalizing denominator.) We call $E$ a functional of the wavefunction $\psi$ because the above expression maps a wavefunction $\psi$ to a number $E$. We aim to show that we can obtain the energy eigenvalues by requiring that $E[\psi]$ be stationary with respect to $\psi$. By this, we mean that, if there is a function $\psi$ such that, for small variations $\delta \psi$ away from $\psi$, the corresponding variation $\delta E$ in $E[\psi]$ vanishes, then $\psi$ is an eigenstate of the Hamiltonian with energy $E[\psi]$. This is the same kind of requirement one places on the classical action in Lagrangian mechanics to yield a differential equation for the classical path, so we are essentially just doing the calculus of variations with the $E[\psi(x)]$ functional instead of the $S[x(t)]$ functional.
Let's explicitly insert a variation $\delta \psi$ into our equation for $E[\psi]$ and determine the resulting $\delta E$:

$$E + \delta E = \frac{\int dx \left[ \psi(x) + \delta \psi(x) \right]^* H \left[ \psi(x) + \delta \psi(x) \right]}{\int dx \left[ \psi(x) + \delta \psi(x) \right]^* \left[ \psi(x) + \delta \psi(x) \right]}$$  \hspace{1cm} (9.2)

It is now clear why we had to keep the denominator explicit: since we are varying $\psi$, the normalization of $\psi$ will change. We thus must explicitly include the normalization correction in the denominator to get the correct energy functional. Since having the variations in the denominator on the right side is hard to deal with, but we may multiply through by the denominator to obtain

$$[E + \delta E] \left[ \int dx \left[ \psi(x) + \delta \psi(x) \right]^* \left[ \psi(x) + \delta \psi(x) \right] \right] = \int \left[ \psi(x) + \delta \psi(x) \right]^* H \left[ \psi(x) + \delta \psi(x) \right]$$  \hspace{1cm} (9.3)

$$[E + \delta E] \left[ \int dx \psi^*(x) \psi(x) + \int dx \delta \psi^*(x) \psi(x) + \int dx \psi^*(x) \delta \psi(x) \right] = \int dx \psi^*(x) H \psi(x) + \int dx \delta \psi^*(x) H \psi(x) + \int dx \psi^*(x) H \delta \psi(x)$$  \hspace{1cm} (9.4)

where we have kept only terms to first order in $\delta \psi$. 
Next, we use the unvaried version of the equation to eliminate the terms that include no $\delta\psi$ factors, and we set $\delta E = 0$ to impose the stationarity condition.

$$
E \left[ \int dx \delta\psi^*(x) \psi(x) + \int dx \psi^*(x) \delta\psi(x) \right] = \int \delta\psi^*(x) H \psi(x) + \int \psi^*(x) H \delta\psi(x)
$$

(9.5)

$$
\int dx \delta\psi^*(x) (H - E) \psi(x) + \int dx \psi^*(x) (H - E) \delta\psi(x) = 0
$$

(9.6)

Next, we can show that the two terms must vanish independently by considering two special cases for the variation $\delta\psi$. Suppose $\delta\psi(x) = \chi(x)$ where $\chi(x)$ is purely real. Then we have

$$
\int dx \chi(x) (H - E) \psi(x) = -\int dx \psi^*(x) (H - E) \chi(x)
$$

(9.7)

Next, suppose $\delta\psi$ is completely imaginary, so $\delta\psi(x) = i \chi(x)$ where $\chi(x)$ is again real. This gives

$$
\int dx \chi(x) (H - E) \psi(x) = \int dx \psi^*(x) (H - E) \chi(x)
$$

(9.8)

where we have divided out by $i$. 

Section 9.1
Variational Method: Derivation
Since $\delta \psi$ is completely arbitrary, it is necessary for both equations to hold simultaneously for any real function $\chi$. That is only possible if both terms in the original equation vanish independently. So we have

\[
\int dx \, \delta \psi^* (x) (H - E) \psi(x) = 0 \quad (9.9)
\]

\[
\int dx \, \psi^* (x) (H - E) \delta \psi(x) = 0 \quad (9.10)
\]

Since $\delta \psi$ is arbitrary in each equation, the integrands must vanish. The first equation yields

\[
(H - E) \psi(x) = 0 \quad (9.11)
\]

One must use the Hermiticity of $H$ to transform the second equation so that $H$ acts on $\psi$ and not $\delta \psi$, which will yield the complex conjugate of the above equation. We recover the eigenvector-eigenvalue equation for $H$. 
Thus, we have proven that, if \( \psi(x) \) is an eigenfunction of \( H \), then a small variation \( \delta \psi(x) \) results in no change in \( E \) to first order in \( \delta \psi(x) \). This proof is interesting in its own right – we had no reason to expect that the Hamiltonian’s eigenfunctions would result in an extremum of the energy functional – in much the same way that there is no reason to expect ahead of time that the classical path that solves Newton’s equations would correspond to an extremum of the action functional.

We note that the we could have dealt with the normalization of \( \psi \) differently. We could have imposed a normalization requirement via a Lagrange multiplier by considering instead the alternate functional

\[
E'[\psi] = \int dx \, \psi^*(x) H \psi(x) - \lambda \left[ \int dx \, \psi^*(x) \psi(x) - 1 \right]
\]  

and requiring \( \delta E' = 0 \) under a variation \( \delta \psi \). The two functionals \( E[\psi] \) and \( E'[\psi] \) are the same when one requires that \( \psi \) be normalized, so the requirement \( \delta E' = 0 \) is equivalent to the requirement \( \delta E = 0 \) under that condition. The result would have been the same had we gone this route.
Is the converse of what we have proven true? That is, if $\psi_m$ is an eigenfunction of $H$ with eigenvalue $E_m$, does it hold that the energy functional $E[\psi]$ is stationary with respect to variations away from $\psi_m$? The answer is yes, and this is relatively easy to show. Let us allow for a variation $\delta \psi$, and let's expand $\delta \psi$ in terms of the eigenstates:

$$|\delta \psi\rangle = \sum_n c_n |\psi_n\rangle \quad (9.13)$$

Now, let's calculate the energy of the wavefunction with the variation:

$$E[\psi_m + \delta \psi] = \frac{(\langle \psi_m | + \langle \delta \psi |) H (|\psi_m\rangle + |\delta \psi\rangle)}{(\langle \psi_m | + \langle \delta \psi |)(|\psi_m\rangle + |\delta \psi\rangle)}$$

$$= \frac{E_m + E_m (\langle \psi_m | \delta \psi \rangle + \langle \delta \psi | \psi_m \rangle)}{1 + (\langle \psi_m | \delta \psi \rangle + \langle \delta \psi | \psi_m \rangle)} \quad (9.15)$$

$$= E_m \left( \frac{1 + c_m + c_m^*}{1 + c_m + c_m^*} \right) = E_m \quad (9.16)$$

Hence, $\delta E = E[\psi_m + \delta \psi] - E[\psi_m] = 0$. So, indeed, if $\psi_m$ is an eigenstate of $H$, then $E[\psi]$ is stationary at $\psi = \psi_m$: a variation $\delta \psi$ in $\psi_m$ results in no variation in $E$ to first order in $\delta \psi$. 

Section 9.1

Variational Method: Derivation
For our purposes, the practical implication of the relationship between $\psi$ being an eigenstate and the energy functional being stationary with respect to variations in $\psi$ is that the fractional error in the energy estimate obtained from the trial wavefunction will be much smaller than that fractional error in the wavefunction itself.
Ground State Energy Upper Bounds

It is easy to show that whatever estimate we make for the ground state energy using this technique, it is always an upper bound. Suppose $\psi(x)$ is our trial wavefunction, the energy eigenfunctions are $\{\phi_n(x)\}$ with eigenvalues $E_n$. We may expand $\psi$, $\psi(x) = \sum_n c_n \phi_n(x)$. Let us then calculate the energy:

$$E[\psi] = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2}$$  \hspace{1cm} (9.17)

Subtract off the ground state energy $E_0$

$$E[\psi] - E_0 = \frac{\sum_n |c_n|^2 (E_n - E_0)}{\sum_n |c_n|^2}$$ \hspace{1cm} (9.18)

Because $E_0$ is the ground state energy, $E_n - E_0 > 0$ for all $n$. $|c_n|^2 \geq 0$, so the right side is nonnegative. Therefore

$$E[\psi] \geq E_0$$ \hspace{1cm} (9.19)
Excited State Energy Estimates

If one knows the eigenfunctions for \( n < m \) and one wants to estimate \( E_m \), then one is assured of an upper bound by requiring that the trial wavefunction be orthogonal to the \( \phi_n \) for \( n < m \). Explicitly, we consider a trial wavefunction \( \psi \) and require

\[
\langle \phi_n | \psi \rangle = 0 \text{ for } n < m
\]  

(9.20)

This condition can be met for any trial wavefunction \( \psi(x) \) via Gram-Schmidt orthogonalization (Section 3.3). We then calculate the energy functional

\[
E[\psi] = \frac{\sum_{n=m}^{\infty} |c_n|^2 E_n}{\sum_{n=0}^{\infty} |c_n|^2}
\]  

(9.21)

because the terms with \( n < m \) vanish due to \( c_n = \langle \phi_n | \psi \rangle = 0 \). We then subtract \( E_m \):

\[
E[\psi] - E_m = \frac{\sum_{n=m}^{\infty} |c_n|^2 (E_n - E_m)}{\sum_{n=0}^{\infty} |c_n|^2}
\]  

(9.22)

\( E_n \geq E_m \) for \( n > m \), so we are assured the right side is nonnegative, yielding

\[
E[\psi] \geq E_m \text{ when } \langle \phi_n | \psi \rangle = 0 \text{ for } n < m
\]  

(9.23)
Derivation (cont.)

Of course, the above is somewhat useless because, if one is able to exactly solve the eigenvector-eigenvalue equation for some $n$, then one is usually able to solve it for all $n$ and the above technique is unnecessary. The idea about Gram-Schmidt orthogonalization is good, though: even if one only has approximations to the lower-energy state eigenfunctions, one should still construct trial wavefunctions for the higher states using Gram-Schmidt orthogonalization. One can show, for example, that the error in estimating $E_m$ is related to the mismatch between one’s estimates $\psi_n$ for the actual eigenfunctions $\phi_n$ for $n < m$. In particular, if

$$\delta_0 = 1 - |\langle \psi_0 | \phi_0 \rangle|^2$$  \hspace{1cm} (9.24)

expresses the fractional deviation of the approximate ground state wavefunction $\psi_0$ from the true one $\phi_0$, then one can show that a trial wavefunction $\psi_1$ that has been constructed using Gram-Schmidt orthogonalization with respect to $\psi_0$ (not $\phi_0$!) yields

$$E[\psi_1] \geq E_1 - \delta_0 (E_1 - E_0)$$  \hspace{1cm} (9.25)

That is, $E[\psi_1]$ is no longer an upper bound on $E_1$, but the amount by which it underestimates $E_1$ is proportional to $\delta_0$. Because $\delta_0$ is quadratic in $\langle \psi_0 | \phi_0 \rangle$, the fractional error in estimating $E_1$ is much smaller than the error in estimating the wavefunction, as long as $E_1 - E_0$ is of order $E_1$. 

Section 9.1

Variational Method: Derivation
How it Works in Practice

How do we construct trial wavefunctions for the purpose of making these variational estimates of energy eigenvalues? One cannot usually guess the form for the correct wavefunction exactly. But one usually knows the general features of the wavefunction one wants. So one constructs a trial wavefunction that depends on some small set of parameters, calculates the energy functional for the trial wavefunction as a function of this small set of parameters, and then requires $E$ be stationary with respect to those parameters – *i.e.*, that all the partial derivatives of $E$ with respect to the parameters vanish. In effect, we are explicitly applying the stationarity condition to some subset of all possible wavefunctions under the expectation that we can get a very good approximation to the energy with a good approximation of the correct wavefunction thanks to the stationarity of $E$ with respect to variations in $\psi$.

An important guide is to require that the trial wavefunctions be eigenfunctions of any other Hermitian operators $A$ that commute with $H$. These are usually related to symmetries of the Hamiltonian. For example, a Hamiltonian in which the potential is an even function of position commutes with the parity operator and hence any eigenfunctions of $H$ must have definite parity. The generators of continuous symmetry transformations are also good examples; rotation symmetry, for example, implies that angular momentum commutes with $H$. If one makes these requirements on the trial wavefunctions, one is assured that they are at least members of the same subspace as the true eigenfunctions.
Example 9.1: Particle in a Box

A good first example is to do a variational-method estimate for a problem whose exact solution we already know – the particle in a box. The Hamiltonian guides our choice of trial wavefunction in two ways. First, we know the trial wavefunction should vanish at the box edge and outside. If the wavefunction’s derivative is large at any point, the kinetic energy will be big, so the wavefunction should rise smoothly and slowly away from zero at the box edges. Second, the Hamiltonian commutes with the parity operator, so the trial wavefunction should be even or odd. We try an even function because it can have no zeroes and thus also have the smallest possible derivative and hence kinetic energy. We try

$$\psi(x; c) = \begin{cases} \left( \left( \frac{L}{2} \right)^2 - x^2 \right) \left( 1 + c x^2 \right) & |x| \leq \frac{L}{2} \\ 0 & |x| > \frac{L}{2} \end{cases}$$

(9.26)

This is the simplest polynomial trial function we can use. Evenness requires that we only include even powers of $x$. A simple quadratic could satisfy the requirement of going to zero at the box edges, but admits no free parameters: it must be $\left( (L/2)^2 - x^2 \right)$ (an overall constant multiplier just sets the normalization and cancels out between the numerator and denominator of the energy functional). So the next possibility is a polynomial containing $x^2$ and $x^4$, and the above is just one convenient way to parameterize it (again, neglecting any overall scale factor).
Calculating the integrals for the energy functional is a somewhat nontrivial algebraic exercise that can be done correctly in Mathematica, yielding

\[ E[\psi] \rightarrow E[c] = \frac{3\, \hbar^2}{m\, L^2} \frac{11 \left( \frac{L}{2} \right)^4 \, c^2 + 14 \left( \frac{L}{2} \right)^2 \, c + 35}{\left( \frac{L}{2} \right)^4 \, c^2 + 6 \left( \frac{L}{2} \right)^2 \, c + 21} \]  
\quad (9.27)

We then find the extrema of \( E(c) \) with respect to \( c \), \( dE/dc = 0 \), yielding

\[ 26 \left( \frac{L}{2} \right)^4 \, c^2 + 196 \left( \frac{L}{2} \right)^2 \, c + 42 = 0 \]  
\quad (9.28)

with solutions

\[ c_1 = -\frac{0.221075}{\left( \frac{L}{2} \right)^2} \quad c_2 = -\frac{7.31771}{\left( \frac{L}{2} \right)^2} \]  
\quad (9.29)
We may then calculate $E$ for these values, giving

$$E(c_1) = 4.93488 \frac{\hbar^2}{mL^2} \quad E(c_2) = 51.0652 \frac{\hbar^2}{mL^2} \quad (9.30)$$

The reason that the second solution has so much higher energy is that it has zeros at $x = \pm c_2^{-1/2} \approx 0.185 L$, hence its derivative and its kinetic energy is much larger than the $c_1$ state, which has no zeros. The true ground state energy is

$$E_{n=1} = \frac{\hbar^2 \pi^2}{2mL^2} = 4.93480 \frac{\hbar^2}{mL^2} \quad (9.31)$$

which is shockingly close to $E(c_1)$. Plotting the trial function shows that it is a very good match to the true ground state wavefunction.
Another interesting fact is that the second solution, using $c_2$, is a decent approximation to the second excited state, $n = 3$:

$$E_{n=3} = 9 \ E_{n=1} = 44.4132 \ \frac{\hbar^2}{mL^2} \tag{9.32}$$

There is a general theorem on this point, the Hylleraas-Undheim Theorem, which essentially states that, if the trial wavefunction depends on a set of parameters, the alternate solutions for the parameters giving non-minimal extrema of $E$ yield estimates for the energies of the excited states of the system.
Example 9.2: Other Examples

A few other examples of simple analytically soluble problems that can be approximated quite well or perfectly by the variational method are:

- The simple harmonic oscillator – a Gaussian trial function will recover the ground state energy and wavefunction exactly because the ground state wavefunction is a Gaussian. A Gaussian times a polynomial will cover the excited states exactly, also, via the Gram-Schmidt orthogonalization procedure.

- Bound state of a $\delta$-function potential well. The trial function should be a decaying exponential.