7 Approximation Methods

In our logical progression from simple systems to more complex systems of chemical interest, the next system to study after the hydrogen atom is the helium atom.

The Schrödinger equation for the helium atom is:

$$\left(-\frac{\hbar^2}{2M}\nabla^2-\frac{\hbar^2}{2m_e}\nabla_1^2-\frac{\hbar^2}{2m_e}\nabla_2^2\right)\psi(\mathbf{R},\mathbf{r}_1,\mathbf{r}_2)+\left(-\frac{2e^2}{4\pi\varepsilon_0|\mathbf{R}-\mathbf{r}_1|}-\frac{2e^2}{4\pi\varepsilon_0|\mathbf{R}-\mathbf{r}_2|}+\frac{e^2}{4\pi\varepsilon_0|\mathbf{r}_1-\mathbf{r}_2|}\right)\psi(\mathbf{R},\mathbf{r}_1,\mathbf{r}_2)=E\psi(\mathbf{R},\mathbf{r}_1,\mathbf{r}_2)$$

R is the position of the helium nucleus, $\mathbf{r}_1$ and $\mathbf{r}_2$ are the positions of the two electrons, M is the mass of the nucleus and m_e is the mass of the electrons. ∇^2 is the Laplacian operator with respect to the position of the nucleus (*i.e.* the derivatives are with respect to the nucleus position). ∇_1^2 and ∇_2^2 are the Laplacian operators with respect to the position of the electrons.

Because this is a three-body problem, the analytical separation between relative coordinates and center of mass coordinates is more complicated. However, because the nucleus of helium (with two protons and two neutrons) is 7300 times more massive than the electron, it is a good approximation to assume that the center of mass is at the nucleus. The center of mass motion is then the motion of this point through space. For the internal problem of the electrons in the center of mass coordinate system, this amounts to ignoring the kinetic energy of the nucleus, since its position is fixed at the origin. This is essentially equivalent to using m_e instead of μ in the hydrogen atom problem, but in this case it is even a better approximation because the helium nucleus is 4 times the mass of a proton.

We can therefore write (to a very high degree of approximation):

$$-\frac{\hbar^2}{2m_e}\left(\nabla_1^2+\nabla_2^2\right)\psi(\mathbf{r}_1,\mathbf{r}_2)-\frac{2e^2}{4\pi\varepsilon_0}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)\psi(\mathbf{r}_1,\mathbf{r}_2)+\frac{e^2}{4\pi\varepsilon_0|\mathbf{r}_1-\mathbf{r}_2|}\psi(\mathbf{r}_1,\mathbf{r}_2)=E\psi(\mathbf{r}_1,\mathbf{r}_2)$$

Even though this is simplified by ignoring the nuclear kinetic energy, this Schrödinger equation cannot be solved exactly due to the term:

$$\frac{e^2}{4\pi\varepsilon_0|\mathbf{r}_1-\mathbf{r}_2|}$$

This term depends upon the relative separation of the two electrons and represents the potential energy due to inter-electronic repulsion.

If this term were gone, the Hamiltonian would just be a sum of two hydrogen atom Hamiltonians. In this case, the wave functions would simply be a product of H atom wave functions and the energy would be a sum of H atom energies. The presence of the inter-electronic repulsion term makes such a separation impossible, however. We must therefore turn to approximation techniques to deal with this problem. I will discuss approximation techniques in general, using several different systems as examples, and afterwards I will return specifically to the problem of the helium atom.

You can see that in any three-(or more)-particle system, we will always run into the problem of the Hamiltonian not being separable if the particles interact. Thus, the Schrödinger equation for many-electron atoms or molecules cannot be solved exactly.

However, there are two different approximation techniques which can yield extremely good results and are widely used in quantum mechanics.

The first of these approaches is called perturbation theory.

7.1 Perturbation Theory

Let me rewrite the Hamiltonian for the Helium atom in scalar form using the coordinate r_{12} to denote $|\mathbf{r}_1\text{-}\mathbf{r}_2|$.

$$\hat{H} = -\frac{\hbar^2}{2m_e}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{2e^2}{4\pi\varepsilon_0}\left(\frac{1}{r_1} + \frac{1}{r_2}\right) + \frac{e^2}{4\pi\varepsilon_0 r_{12}}$$

We can see that this Hamiltonian has the form

$$\hat{H} = \hat{H}_{H1} + \hat{H}_{H2} + \frac{e^2}{4\pi\varepsilon_0}\frac{1}{r_{12}}$$

where

$$\hat{H}_{Hi} = -\frac{\hbar^2}{2m_e}\nabla_i^2 - \frac{2e^2}{4\pi\varepsilon_0}\frac{1}{r_i} \qquad i = 1,2$$

$\hat{H}_{Hi}$ is just the Hamiltonian for a hydrogen-like atom (a one-electron atom) in which the nuclear charge $Z = 2$. So the Hamiltonian for the Helium atom looks like the sum of two hydrogen atom Hamiltonians plus a term which represents the inter-electronic repulsion.

If the inter-electronic repulsion term,

$$\frac{e^2}{4\pi\varepsilon_0}\frac{1}{r_{12}}$$

weren't there, we could separate the Hamiltonian into two Hamiltonians which depend on coordinates of different electrons. We would then know how to solve that problem. The wave functions would be the product of one-electron wave functions and the energies would be the sum of energies of each electron. However, the inter-electronic repulsion term prevents us from doing this.

Let's think back to another problem that we could not solve exactly – the anharmonic oscillator problem. Recall that the vibrational motions of real molecules are not exactly harmonic, but for small amplitude vibrations a harmonic approximation works pretty well. We demonstrated this by writing the potential energy as a Taylor series expansion about the bottom of the well:

$$U(x) = \frac{1}{2}kx^2 + \frac{1}{6}\gamma x^3 + \frac{1}{24}bx^4 + \ldots$$

The Hamiltonian is then

$$\hat{H} = -\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2} + \frac{1}{2}kx^2 + \frac{1}{6}\gamma x^3 + \frac{1}{24}bx^4 + \ldots$$

We then pointed out that for small amplitude vibrations, the latter two terms would be small and we could neglect them. The Hamiltonian for the *anharmonic* oscillator then reduces to that for the harmonic oscillator.

The problem of the Helium atom and that of the anharmonic oscillator have an important property in common -- they both can be written in the form

$$\hat{H} = \hat{H}^{(0)} + \hat{H}'$$

where $\hat{H}^{(0)}$ is the Hamiltonian for a problem that we know how to solve exactly, and $\hat{H}'$ is a small additional term that prevents an exact solution.

$\hat{H}^{(0)}$ is called the unperturbed Hamiltonian and $\hat{H}'$ is called the perturbation.

In each case we know how to solve the Schrödinger equation for $\hat{H}^{(0)}$, that is we know the solution of

$$\hat{H}^{(0)}\psi^{(0)} = E^{(0)}\psi^{(0)}$$

The basic approach of *perturbation theory* is to write the Hamiltonian in the form

$$\hat{H} = \hat{H}^{(0)} + \hat{H}'$$

If the perturbation to the Hamiltonian is sufficiently small, the energies and the wave functions will not be very different from the unperturbed problem. Perturbation theory then formulates the solution to the perturbed problem in the form of corrections to the energy and wave functions of the unperturbed problem.

I will call the unperturbed wave functions $\psi^{(0)}$ the *zeroth-order wave functions* and the energies $E^{(0)}$ the *zeroth-order energies*.

In the case of the Helium atom, the zeroth-order energy is the sum of two hydrogen atom energies and the zeroth-order wavefunctions the product of two hydrogen atom wave functions.

In the case of the anharmonic oscillator, the zeroth-order energies and wave functions are simply those of the harmonic oscillator problem.

We will solve these problems after we develop the theory.

7.1.1 Derivation of the Perturbation Theory

The basic problem is that we need to find solutions to the Schrödinger equation

$$\hat{H}\psi = E\psi$$

but many times we cannot solve it exactly. This means that the Hamiltonian is such that we cannot find analytic solutions for the energies and wave functions. However, we can often write $\hat{H}$ in the form:

$$\hat{H} = \hat{H}^{(0)} + \hat{H}'$$

where $\hat{H}^{(0)}$ is a Hamiltonian for which we can solve the Schrödinger Equation exactly and $\hat{H}'$ is a small *perturbation* to that Hamiltonian.

For Perturbation Theory to yield reasonable results, the perturbation $\hat{H}'$ must be small compared to $\hat{H}^{(0)}$. To emphasize this we will write the perturbation as $\lambda\hat{H}'$ where λ is some small number <<1.

So we will in general write:

$$\hat{H} = \hat{H}^{(0)} + \lambda\hat{H}'$$

We never need to know what the parameter λ is besides to know that it is a small number. It is a useful construct that will drop out.

The problem we are trying to solve is then

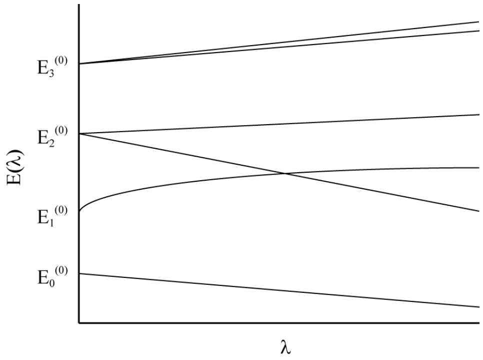
$$\left(\hat{H}^{(0)} + \lambda\hat{H}'\right)\psi_n = E_n\psi_n$$

Since $\hat{H}$ depends on the parameter λ, both E_n and ψ_n will depend on λ, that is

$$\psi_n = \psi_n(\lambda, q) \qquad \text{and} \qquad E_n = E_n(\lambda)$$

where q are the coordinates.

We can look at the problem in the following way:



The eigenvalues will depend upon the value of the parameter λ, that is, how large the perturbation is. As $\lambda \to 0$, they will approach the eigenvalues of the unperturbed system.

There is at least one wave function for each energy, and for some there is more than one (degenerate states). As we turn on the perturbation by making λ different from zero, both the energies and eigenfunctions will change. One can actually do this with magnetic or electric fields. In some cases the perturbation will split the degeneracy ($E_3^{(0)}$), in others it may not ($E_4^{(0)}$). The perturbation treatment of degenerate and non-degenerate states differs. I will deal only with non-degenerate states.

Since $\hat{H}$ depends on λ, E_n and ψ_n should as well. We will therefore assume that we can expand the wave function ψ and the energy E_n in powers of λ,

$$\psi_n = \psi_n^{(0)} + \lambda\psi_n^{(1)} + \lambda^2\psi_n^{(2)} + \ldots + \lambda^k\psi_n^{(k)} + \ldots$$

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots + \lambda^k E_n^{(k)} + \ldots$$

where we call $E_n^{(k)}$ and $\psi_n^{(k)}$ the k^{th} order corrections to the energy and wave function.

For small perturbations this series should converge, and the energy and wave function should be fairly well approximated by carrying just a few of the correction terms. Putting these expressions in to the Schrödinger equation:

$$\left(\hat{H}^{(0)} + \lambda\hat{H}'\right)\left(\psi_n^{(0)} + \lambda\psi_n^{(1)} + \lambda^2\psi_n^{(2)} + \ldots + \lambda^k\psi_n^{(k)} + \ldots\right) =$$

$$\left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots + \lambda^k E_n^{(k)} + \ldots\right)\left(\psi_n^{(0)} + \lambda\psi_n^{(1)} + \lambda^2\psi_n^{(2)} + \ldots + \lambda^k\psi_n^{(k)} + \ldots\right)$$

If we multiply out all the terms and group similar powers of λ we get:

$$H^{(0)}\psi_n^{(0)} + \lambda\left(\hat{H}^{(0)}\psi_n^{(1)} + \hat{H}'\psi_n^{(0)}\right) + \lambda^2\left(\hat{H}^{(0)}\psi_n^{(2)} + \hat{H}'\psi_n^{(1)}\right) + \ldots$$

$$= E_n^{(0)}\psi_n^{(0)} + \lambda\left(E_n^{(1)}\psi_n^{(0)} + E_n^{(0)}\psi_n^{(1)}\right) + \lambda^2\left(E_n^{(2)}\psi_n^{(0)} + E_n^{(1)}\psi_n^{(1)} + E_n^{(0)}\psi_n^{(2)}\right) + \ldots$$

For this equality to hold, the coefficient of each power of λ must be equal.

For the λ^0 terms, we get:

$$\hat{H}^{(0)}\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)}$$

There is no new information here, since this is simply the Schrödinger equation for the unperturbed system, the solution to which we already know.

For the λ^1 terms we get:

$$\hat{H}^{(0)}\psi_n^{(1)} + \hat{H}'\psi_n^{(0)} = E_n^{(1)}\psi_n^{(0)} + E_n^{(0)}\psi_n^{(1)}$$

$$\left(\hat{H}^{(0)} - E_n^{(0)}\right)\psi_n^{(1)} = \left(E_n^{(1)} - \hat{H}'\right)\psi_n^{(0)}$$

We need to find $E_n^{(1)}$ and $\psi_n^{(1)}$. Since $H^{(0)}$ is Hermitian, its eigenfunctions form a complete orthonormal set. We can then expand $\psi_n^{(1)}$ in terms of these functions:

$$\psi_n^{(1)} = \sum_j a_j \psi_j^{(0)}$$

Our equation then becomes:

$$\sum_j a_j \left(\hat{H}^{(0)} - E_n^{(0)}\right)\psi_j^{(0)} = \left(E_n^{(1)} - \hat{H}'\right)\psi_n^{(0)}$$

but we know that

$$\hat{H}^{(0)}\psi_j^{(0)} = E_j^{(0)}\psi_j^{(0)}$$

so we can write

$$\sum_j a_j \left(E_j^{(0)} - E_n^{(0)}\right)\psi_j^{(0)} = \left(E_n^{(1)} - H'\right)\psi_n^{(0)}$$

If we now multiply both sides by $\psi_n^{(0)*}$ and integrate, one has:

$$\int \psi_n^{(0)*} \sum_j a_j \left(E_j^{(0)} - E_n^{(0)}\right)\psi_j^{(0)} d\tau = \int \psi_n^{(0)*} \left(E_n^{(1)} - \hat{H}'\right)\psi_n^{(0)} d\tau$$

We can rewrite the left side of this equation

$$\sum_j a_j \left(E_j^{(0)} - E_n^{(0)}\right)\int \psi_n^{(0)*}\psi_j^{(0)} d\tau = \sum_j a_j \left(E_j^{(0)} - E_n^{(0)}\right)\delta_{nj}$$

You can see that the summation yields zero since when we do it, the only term that survives is when *j*=*n*, and

$$\left(E_n^{(0)} - E_n^{(0)}\right) = 0$$

Thus the right side of our equation must also equal zero

$$\int \psi_n^{(0)*} \left(E_n^{(1)} - \hat{H}'\right)\psi_n^{(0)} d\tau = 0$$

Solving for $E_n^{(1)}$

$$E_n^{(1)} = \int \psi_n^{(0)*} \hat{H}' \psi_n^{(0)} d\tau$$

Remember that the 1st order correction term to the energy was given by $\lambda E_n^{(1)}$

$$\lambda E_n^{(1)} = \lambda \int \psi_n^{(0)*} \hat{H}' \psi_n^{(0)} d\tau = \int \psi_n^{(0)*} \lambda \hat{H}' \psi_n^{(0)} d\tau$$

We define H'_{nn} as the integral of $\lambda H'$

$$H'_{nn} = \int \psi_n^{(0)*} \lambda \hat{H}' \psi_n^{(0)} d\tau$$

Recall that $\lambda H'$ is just the perturbation part of the Hamiltonian. Thus, the first order correction to the energy, $\lambda H'$, represents the *average value of the perturbation*.

We can now do away with the parameter λ. What we called $\lambda H'$ before we will now call H' – the perturbation part of the Hamiltonian. We can also absorb λ into the definition of the first order correction $E_n^{(1)}$ as well.

We can then write to first order:

$$E_n = E_n^{(0)} + E_n^{(1)} = E_n^{(0)} + H'_{nn}$$

where the first order correction to the energy, H'_{nn} , represents the average value of the perturbation.

We now need to find the first order correction to the wave function, $\psi_n^{(1)}$. Remember we had expanded $\psi_n^{(1)}$ in basis functions of the eigenfunctions of the unperturbed Hamiltonian:

$$\psi_n^{(1)} = \sum_j a_j \psi_j^{(0)}$$

Finding the correction to the wave function consists of finding the coefficients a_j in this expansion.

If we go back to our equation

$$\sum_j a_j \left(E_j^{(0)} - E_n^{(0)} \right) \psi_j^{(0)} = \left(E_n^{(1)} - \hat{H}' \right) \psi_n^{(0)}$$

and multiply by $\psi_m^{(0)*}$ and integrate (where $m \neq n$)

$$\int \psi_m^{(0)*} \sum_j a_j \left(E_j^{(0)} - E_n^{(0)} \right) \psi_j^{(0)} d\tau = \int \psi_m^{(0)*} \left(E_n^{(1)} - \hat{H}' \right) \psi_n^{(0)} d\tau$$

$$\sum_j a_j \left(E_j^{(0)} - E_n^{(0)} \right) \int \psi_m^{(0)*} \psi_j^{(0)} d\tau = E_n^{(1)} \int \psi_m^{(0)*} \psi_n^{(0)} d\tau - \int \psi_m^{(0)*} \hat{H}' \psi_n^{(0)} d\tau$$

$$\sum_j a_j \left(E_j^{(0)} - E_n^{(0)} \right) \delta_{mj} = E_n^{(1)} \delta_{mn} - H'_{mn}$$

$\delta_{mn} = 0$ since we are restricting ourselves to $m \neq n$. Also, we can evaluate the sum over *j* since only the term in which $j = m$ will survive:

$$a_m \left(E_m^{(0)} - E_n^{(0)} \right) = -H'_{mn}$$

Since $m \neq n$, we can divide both sides by $\left(E_m^{(0)} - E_n^{(0)} \right)$ (we can do this because states *m* and *n* are non-degenerate)

$$a_m = \frac{H'_{mn}}{\left(E_n^{(0)} - E_m^{(0)}\right)}$$ (Note the change of order and the loss of the minus sign)

So our first order correction to the wave function is

$$\psi_n^{(1)} = \sum_{m \neq n} \frac{H'_{mn}}{\left(E_n^{(0)} - E_m^{(0)}\right)} \psi_m^{(0)}$$

where the sum is for *m≠n*.

Note that although there is no $\psi_n^{(0)}$ in the sum, the zeroth-order wave function does appear in the expression for the corrected ψ_n.

The wave function ψ_n is then

$$\psi_n = \psi_n^{(0)} + \sum_{m \neq n} \frac{H'_{mn}}{\left(E_n^{(0)} - E_m^{(0)}\right)} \psi_m^{(0)}$$

Note that once again, we have absorbed λ into our definition of H'_{mn}.

This quantity H'_{mn} is also an integral involving the perturbation part of the Hamiltonian, but now the integral involves two different functions. We call both H'_{nn} and H'_{mn} *matrix elements*, in that we can form a matrix of all the integrals of the perturbation with all of the different wave functions. The first order correction to the energy, H'_{nn} would lie along the diagonal, and the H'_{mn} would be the various off-diagonal elements.

Looking back at the expression for the first order correction to the wave function, you can see that the effect of a perturbation is to "mix" the unperturbed (or "zeroth-order") state with contributions from other states. In general, those contributions are larger the closer they are to the level under consideration because of the denominator of this expression.

Our treatment thus far has been concerned with finding only the first order correction to the wave function and energy. If we go back to our original Schrödinger equation in which we had substituted expansions in powers of λ for *E* and ψ and had grouped powers of λ, the terms involving λ^2 will give us the second order corrections. I will simply quote the result:

$$E_n^{(2)} = \sum_{m \neq n} \frac{\left|H'_{mn}\right|^2}{\left(E_n^{(0)} - E_m^{(0)}\right)}$$

The energy to second order can therefore be written

$$E_n = E_n^{(0)} + H'_{nn} + \sum_{m \neq n} \frac{\left|H'_{mn}\right|^2}{\left(E_n^{(0)} - E_m^{(0)}\right)}$$

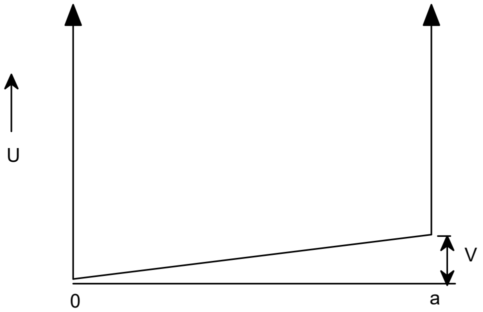
To get the first order correction to the energy, only a single integral needs to be evaluated, whereas for the second order correction to the energy and the first order correction to the wave function, we must evaluate all the integrals in the infinite sum. It is often difficult to evaluate these sums. For higher corrections it is more difficult still.

However, the most important terms come from states that are closest in energy to the one being calculated because of the $\frac{1}{\left(E_n^{(0)} - E_m^{(0)}\right)}$ term.

You should know what all the terms in the expression for the 1st order correction to the wave function and second order correction to the energy mean. Often, many of the matrix elements that you need to evaluate will be zero on the basis of symmetry. You should be aware of this and look for it.

7.1.2 Examples of Using Perturbation Theory

Example 1: Particle in a box with slanted bottom.



We need to formulate the problem as an zeroth-order part plus a perturbation. The zeroth-order part must be one that we know how to solve.

If we let H$^{(0)}$ be the particle in the box Hamiltonian, then we can write

$$\hat{H} = \hat{H}^{(0)} + \hat{H}'$$

where $$\hat{H}' = \frac{Vx}{a}$$

We know the zeroth-order energy is just the particle in a box energy. Let us then find the first order correction to the energies, $E_n^{(1)}$.

$$E_n^{(1)} = H'_{nn} = \int_0^a \psi_n^{(0)*}(x)\left(\frac{Vx}{a}\right)\psi_n^{(0)}(x)dx$$

$$= \frac{2}{a}\frac{V}{a}\int_0^a x\sin^2\left(\frac{n\pi}{a}x\right)dx$$

$$= \frac{2}{a}\frac{V}{a}\frac{a^2}{4}$$

So $$E_n^{(1)} = \frac{V}{2}$$

Note that this is independent of *n*. In general, the effect of the perturbation may depend upon the quantum numbers. In this case, however, each level is shifted by the same amount.

Thus, $$E_n = E_n^{(0)} + E_n^{(1)} = \frac{n^2h^2}{8ma^2} + \frac{V}{2}$$

Example 2: Anharmonic oscillator.

Recall that in discussing the harmonic oscillator problem I indicated that real molecules are somewhat anharmonic (that is, the potential contains terms other than $\frac{1}{2}kx^2$). Consider an anharmonic oscillator whose potential is given by

$$U(x) = \frac{1}{2}kx^2 + \frac{1}{6}\gamma x^3 + \frac{1}{24}bx^4$$

Let us determine the first-order correction to the ground state energy of an anharmonic oscillator with a potential as given above. Comparing $U(x)$ above to the harmonic oscillator potential, we can see that

$$\hat{H}' = \frac{1}{6}\gamma x^3 + \frac{1}{24}bx^4$$

Recall that the ground state wave function for the harmonic oscillator is

$$\psi_0(x) = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\alpha}{2}x^2}$$

where $\alpha = \sqrt{\dfrac{k\mu}{\hbar^2}}$

The first order correction to the ground state energy is given by

$$E_0^{(1)} = H'_{00} = \int_{-\infty}^{\infty} \psi_0^{(0)*}(x)\hat{H}'\psi_0^{(0)}(x)dx$$

$$= \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \int_{-\infty}^{\infty} \left(\frac{1}{6}\gamma x^3 + \frac{1}{24}bx^4\right) e^{-\alpha x^2} dx$$

$$= \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \left[\int_{-\infty}^{\infty} \left(\frac{1}{6}\gamma x^3\right) e^{-\alpha x^2} dx + \int_{-\infty}^{\infty} \left(\frac{1}{24}bx^4\right) e^{-\alpha x^2} dx\right]$$

The first of these integrals equals zero since the function under the integral is odd. The second term gives

$$E_0^{(1)} = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \frac{2}{24} b \int_0^{\infty} x^4 e^{-\alpha x^2} dx = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \frac{b}{12} \frac{3\sqrt{\pi}}{8\alpha^{\frac{5}{2}}}$$

Thus, $E_0^{(1)} = \dfrac{b}{32\alpha^2} = \dfrac{\hbar^2 b}{32k\mu}$

The total ground state energy is then

$$E_0 = \frac{1}{2}h\nu + \frac{\hbar^2 b}{32k\mu} + \ldots$$

7.2 The Variational Principle

I would now like to introduce a second approximation method for solving the Schrödinger equation – the *variational principle*. This method does not require that a similar problem has been solved exactly as in the case of perturbation theory. It is widely used for solving the *electronic* Schrödinger equation (that is, the Schrödinger equation for electronic motion in molecules).

Consider the ground (lowest energy) state of some arbitrary system. The wave function ψ_0 , and energy E_0, satisfy the relation

$$\hat{H}\psi_0 = E_0\psi_0$$

If we multiply this equation by ψ_0^* and integrate over all space we get

$$E_0 = \frac{\int \psi_0^* \hat{H}\psi_0 \, d\tau}{\int \psi_0^* \psi_0 \, d\tau}$$

where $d\tau$ is the volume element.

Note that ψ_0 is the true wave function for the lowest energy state, even though we don't know what it is.

We could set the denominator equal to one if we choose normalized functions, but if we leave it we can allow for the possibility that the functions are not normalized.

If we substitute any other function for ψ into this expression (any function that we might guess to be the true wave function), the variation principle says that energy we would calculate from this expression will be greater than or equal to the true ground state energy.

That is, if we calculate

$$E_\varphi = \frac{\int \Phi^* \hat{H}\Phi \, d\tau}{\int \Phi^* \Phi \, d\tau}$$

then

$$E_\varphi \geq E_0$$

The expression for E_φ is called the *variational function*, and the wave function Φ is called a *trial function*.

The only requirements on the function Φ are that it satisfies the proper boundary conditions. We can determine this without solving the problem.

7.2.1 Proof of the Variational Principle

Let us expand our function Φ as a linear combination of eigenfunctions of the Hamiltonian for the system of interest. That is:

$$\Phi = \sum_n c_n \psi_n$$

where $\hat{H}\psi_n = E_n\psi_n$

We can see that

$$\hat{H}\Phi = \sum_n \hat{H} c_n \psi_n = \sum_n c_n E_n \psi_n$$

Substituting these expressions back into the variational function we get

$$E_\varphi = \frac{\int \left(\sum_m c_m^* \psi_m^* \right) \left(\sum_n c_n E_n \psi_n \right) d\tau}{\int \left(\sum_m c_m^* \psi_m^* \right) \left(\sum_n c_n \psi_n \right) d\tau}$$

We can rearrange this to get

$$E_\varphi = \frac{\sum_m \sum_n c_m^* c_n E_n \int \psi_m^* \psi_n d\tau}{\sum_m \sum_n c_m^* c_n \int \psi_m^* \psi_n d\tau}$$

Because the eigenfunctions form a complete orthonormal set we can write:

$$E_\varphi = \frac{\sum_n c_n^* c_n E_n}{\sum_n c_n^* c_n}$$

If we now subtract E_0 from the left side and the quantity below (which is equal to E_0) from the right side

$$\frac{\sum_n c_n^* c_n E_0}{\sum_n c_n^* c_n} = E_0$$

we get

$$E_\varphi - E_0 = \frac{\sum_n c_n^* c_n \left(E_n - E_0 \right)}{\sum_n c_n^* c_n}$$

We know that $E_n \geq E_0$ for all n since E_0 is the lowest energy (the ground state).

Also, $c_n^* c_n \geq 0$. Thus, the right hand side of the equation is positive.

Thus we have

$$E_\varphi - E_0 \geq 0$$

or

$$E_\varphi \geq E_0$$

Basically, *any trial function will always yield an energy greater than the exact solution.* We can then choose the function to contain "variational" parameters, and we can minimize E_φ as a function of the parameters. The value will be the best value for the adjustable parameter.

7.2.2 Examples of Variational Calculations

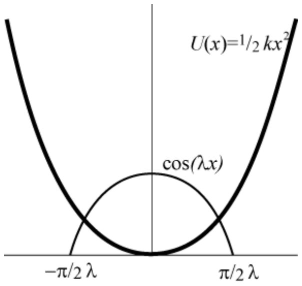
Example 1:

Consider the ground state of a harmonic oscillator. Assume for the sake of this example that we did not know how to solve it by exact methods. We might expect, however, that the ground state wave function would be symmetric about x=0 (by analogy to problems like the particle in a box).

Let us try the function

$$\Phi = \cos(\lambda x) \qquad -\frac{\pi}{2\lambda} \le x \le \frac{\pi}{2\lambda}$$

where λ is a variable parameter. We can vary the parameter λ to get the best estimate of the ground state energy.



We will only consider the function between $-\pi/2\lambda$ and $\pi/2\lambda$ since we know the function must damp out at $x = \pm\infty$ and not oscillate as this does. If we consider it just to the point where it goes to zero it should be ok.

Remember the variational function has the form

$$E_\varphi = \frac{\int \Phi^* \hat{H} \Phi \, d\tau}{\int \Phi^* \Phi \, d\tau}$$

Recall that the Hamiltonian for the harmonic oscillator is

$$\hat{H} = -\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2} + \frac{1}{2}kx^2$$

Substituting for Φ and $\hat{H}$

$$E_\varphi = \frac{\int_{-\frac{\pi}{2\lambda}}^{\frac{\pi}{2\lambda}} \cos(\lambda x)\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2} + \frac{1}{2}kx^2\right]\cos(\lambda x)\,dx}{\int_{-\frac{\pi}{2\lambda}}^{\frac{\pi}{2\lambda}} \cos(\lambda x)\cos(\lambda x)\,dx}$$

Let's do the denominator first:

$$\int_{-\frac{\pi}{2\lambda}}^{\frac{\pi}{2\lambda}} \cos^2(\lambda x)\,dx = \frac{\pi}{2\lambda}$$

Next, the numerator:

$$\int_{-\frac{\pi}{2\lambda}}^{\frac{\pi}{2\lambda}} \cos(\lambda x)\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2}+\frac{1}{2}kx^2\right]\cos(\lambda x)dx = \frac{\hbar^2\lambda^2}{2\mu}\int_{-\frac{\pi}{2\lambda}}^{\frac{\pi}{2\lambda}}\cos^2(\lambda x)dx + \frac{k}{2}\int_{-\frac{\pi}{2\lambda}}^{\frac{\pi}{2\lambda}} x^2\cos^2(\lambda x)dx$$

$$= \frac{\hbar^2\lambda^2}{2\mu}\left(\frac{\pi}{2\lambda}\right)+\frac{k}{\lambda^3}\left(\frac{\pi^3}{48}-\frac{\pi}{8}\right)$$

The function E_φ is then

$$E_\varphi = \frac{\frac{\hbar^2\lambda^2}{2\mu}\left(\frac{\pi}{2\lambda}\right)+\frac{k}{\lambda^3}\left(\frac{\pi^3}{48}-\frac{\pi}{8}\right)}{\frac{\pi}{2\lambda}}$$

$$= \frac{\hbar^2\lambda^2}{2\mu}+\frac{k}{\lambda^2}\left(\frac{\pi^2}{24}-\frac{1}{4}\right)$$

We now want to minimize this function with respect to the parameter λ. So we need to take the first derivative with respect to λ and set it equal to zero (recall that this is the prescription for finding a minimum or maximum of a function).

$$\frac{dE_\varphi}{d\lambda} = \frac{2\hbar^2\lambda}{2\mu}-\frac{2k}{\lambda^3}\left(\frac{\pi^2}{24}-\frac{1}{4}\right)=0$$

Multiplying both sides by λ^3 we get:

$$\frac{\hbar^2}{\mu}\lambda^4-2k\left(\frac{\pi^2}{24}-\frac{1}{4}\right)=0$$

The minimum is then at

$$\lambda^2_{\min} = +\frac{\sqrt{2k\mu\left(\frac{\pi^2}{24}-\frac{1}{4}\right)}}{\hbar}$$

Note that we took the positive root only as λ^2 must be positive.

Substituting λ^2 into E_φ:

$$E_{\varphi\min} = \frac{\hbar^2}{2\mu}\frac{\sqrt{2k\mu\left(\frac{\pi^2}{24}-\frac{1}{4}\right)}}{\hbar}+\left(\frac{\pi^2}{24}-\frac{1}{4}\right)\frac{k\hbar}{\sqrt{2k\mu\left(\frac{\pi^2}{24}-\frac{1}{4}\right)}}$$

This gives

$$E_{\varphi\min} = \sqrt{2}\sqrt{\frac{\pi^2}{24}-\frac{1}{4}}\,\hbar\sqrt{\frac{k}{\mu}} = 0.57\hbar\omega = 0.57h\nu$$

Compare this to the exact solution:

$$E_0 = 0.5\ h\nu$$

Note that $E_\varphi > E_0$. Also note that even with this crude approximation we get excellent results.

Example 2:

Now let us repeat the variational calculation using the trial function

$$\Phi = e^{-\frac{\alpha}{2}x^2}$$

Then

$$E_\varphi = \frac{\int_{-\infty}^{\infty} e^{-\frac{\alpha}{2}x^2}\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2} + \frac{1}{2}kx^2\right]e^{-\frac{\alpha}{2}x^2}\,dx}{\int_{-\infty}^{\infty} e^{-\frac{\alpha}{2}x^2}e^{-\frac{\alpha}{2}x^2}\,dx}$$

Numerator:

$$\int_{-\infty}^{\infty} e^{-\frac{\alpha}{2}x^2}\left[-\frac{\hbar^2}{2\mu}\left(\alpha^2x^2e^{-\frac{\alpha}{2}x^2} - \alpha e^{-\frac{\alpha}{2}x^2}\right) + \frac{1}{2}kx^2e^{-\frac{\alpha}{2}x^2}\right]dx = \int_{-\infty}^{\infty}\frac{\hbar^2\alpha}{2\mu}e^{-\alpha x^2}dx + \int_{-\infty}^{\infty}\left(\frac{1}{2}k - \frac{\hbar^2\alpha^2}{2\mu}\right)x^2e^{-\alpha x^2}dx$$

$$= \frac{\hbar^2\alpha}{2\mu}\sqrt{\frac{\pi}{\alpha}} + \left(\frac{1}{2}k - \frac{\hbar^2\alpha^2}{2\mu}\right)\frac{1}{2\alpha}\sqrt{\frac{\pi}{\alpha}}$$

Denominator:

$$\int_{-\infty}^{\infty} e^{-\alpha x^2}dx = \sqrt{\frac{\pi}{\alpha}}$$

$$E_\varphi = \frac{\hbar^2\alpha}{2\mu} + \left(\frac{1}{2}k - \frac{\hbar^2\alpha^2}{2\mu}\right)\frac{1}{2\alpha}$$

So,

$$= \frac{\hbar^2\alpha}{2\mu} + \frac{1}{4\alpha}k - \frac{\hbar^2\alpha}{4\mu}$$

$$= \frac{\hbar^2\alpha}{4\mu} + \frac{1}{4\alpha}k$$

We now have to take the derivative and set it equal to zero to find the best value of α.

$$\frac{dE_\varphi}{d\alpha} = \frac{\hbar^2}{4\mu} - \frac{1}{4\alpha^2}k = 0$$

yielding:

$$\alpha^2 = \frac{k\mu}{\hbar^2}$$

Now substitute this back into the expression for E_φ to get the minimum value.

$$E_{\varphi\min} = \frac{\hbar^2}{4\mu}\sqrt{\frac{k\mu}{\hbar^2}} + \frac{k}{4}\sqrt{\frac{\hbar^2}{k\mu}}$$

$$= \frac{\hbar}{4}\sqrt{\frac{k}{\mu}} + \frac{\hbar}{4}\sqrt{\frac{k}{\mu}}$$

$$= \frac{\hbar}{2}\sqrt{\frac{k}{\mu}} = \frac{1}{2}\hbar\omega = \frac{1}{2}h\nu$$

Note: this is the same as the exact result!!

The better agreement (exact in this case) using this trial function stems from the fact that the function is closer to (in this case identical to) the true wave function. It came out exact because we chose a function that was in the exact same form as the ground state solution. The α we found was simply the definition for α we had used previously.

However, the general principle holds true: The closer the wave function is to the true one, the better the estimate of the energy. The general approach is to include many variable parameters in the trial function and vary them in a systematic manner to obtain the lowest energy possible. The wave function becomes closer to the exact solution.

In principle one can use this approach to find not only the ground state energy but also the ground state wave function, however the estimate of the energy approaches the true energy faster than the wave function approaches the true wave function. You therefore need an energy very close to the true energy before the wave function is close to the true one.

Up to now, we have only used a single variable parameter in our trial function. You can imagine that if we choose a trial function with several variable parameters, calculate the variational function, and then vary each of the parameters to minimize the energy, one could get an even better estimate of the energy.

7.2.3 Variational Principle with more than one variable parameter

A systematic way to handle a trial function with more than one variational parameter is to write it as a linear combination of functions

$$\Phi = \sum_{n=1}^{N} c_n f_n$$

where the c_n are the parameters which we are going to vary, and the f_n are arbitrary known functions that are not necessarily orthonormal.

The more terms we take in the sum, the better our estimate of the ground state energy will be. This accuracy comes at the cost of computational complexity, not in the sense of conceptual difficulty, but in the sense of the raw computer power required.

Consider the simplest case where we have a linear combination of only two functions

$$\Phi = c_1 f_1 + c_2 f_2$$

We will choose the c_n and f_n to be real.

Let us plug this into the variational function E_φ.

First evaluate the numerator:

$$\int \Phi^* \hat{H} \Phi \, d\tau = \int \left(c_1 f_1 + c_2 f_2 \right)^* \hat{H} \left(c_1 f_1 + c_2 f_2 \right) d\tau$$

$$= c_1^2 \int f_1 \hat{H} f_1 \, d\tau + c_1 c_2 \int f_1 \hat{H} f_2 \, d\tau + c_2 c_1 \int f_2 \hat{H} f_1 \, d\tau + c_2^2 \int f_2 \hat{H} f_2 \, d\tau$$

(I dropped the * since the *c*'s and *f*'s are real)

Let us define a matrix element H_{ij} as

$$H_{ij} = \int f_i \hat{H} f_j \, d\tau$$

Note that since $\hat{H}$ is Hermitian $H_{ij} = H_{ji}$

We can therefore write

$$\int \Phi^* \hat{H} \Phi \, d\tau = c_1^2 H_{11} + 2 c_1 c_2 H_{12} + c_2^2 H_{22}$$

Now the denominator:

$$\int \Phi^* \Phi \, d\tau = \int \left(c_1 f_1 + c_2 f_2 \right)^* \left(c_1 f_1 + c_2 f_2 \right) d\tau$$

$$= c_1^2 \int f_1 f_1 \, d\tau + 2 c_1 c_2 \int f_1 f_2 \, d\tau + c_2^2 \int f_2 f_2 \, d\tau$$

Let's define an overlap integral S_{ij}

$$S_{ij} = \int f_i f_j \, d\tau$$

Since the functions f_i and f_j are not by definition eigenfunctions of an operator, they are not necessarily orthonormal and thus we can write:

$$\int \Phi^* \Phi \, d\tau = c_1^2 S_{11} + 2 c_1 c_2 S_{12} + c_2^2 S_{22}$$

This gives us

$$E_\varphi(c_1, c_2) = \frac{c_1^2 H_{11} + 2 c_1 c_2 H_{12} + c_2^2 H_{22}}{c_1^2 S_{11} + 2 c_1 c_2 S_{12} + c_2^2 S_{22}}$$

We now want to differentiate this with respect to c_1 and c_2. Before we do this let's write this expression in the following form

$$E_\varphi(c_1, c_2) \left(c_1^2 S_{11} + 2 c_1 c_2 S_{12} + c_2^2 S_{22} \right) = c_1^2 H_{11} + 2 c_1 c_2 H_{12} + c_2^2 H_{22}$$

Now let us differentiate it with respect to c_1. We need to use the chain rule on the left side:

$$\frac{\partial E_\varphi(c_1, c_2)}{\partial c_1} \left(c_1^2 S_{11} + 2 c_1 c_2 S_{12} + c_2^2 S_{22} \right) + E_\varphi(c_1, c_2) \left(2 c_1 S_{11} + 2 c_2 S_{12} \right) = 2 c_1 H_{11} + 2 c_2 H_{12}$$

To find the minimum we set the derivative to zero:

$$\frac{\partial E_\varphi(c_1,c_2)}{\partial c_1}=0$$

This gives

$$E_\varphi(c_1,c_2)\left(2c_1S_{11}+2c_2S_{12}\right)=2c_1H_{11}+2c_2H_{12}$$

and consequently we get:

$$c_1\left(H_{11}-E_\varphi(c_1,c_2)S_{11}\right)+c_2\left(H_{12}-E_\varphi(c_1,c_2)S_{12}\right)=0$$

Similarly, if we differentiate *E* with respect to c_2 and set the derivative equal to zero we get

$$c_1\left(H_{12}-E_\varphi(c_1,c_2)S_{12}\right)+c_2\left(H_{22}-E_\varphi(c_1,c_2)S_{22}\right)=0$$

These two equations constitute a pair of linear algebraic equations for the variables c_1 and c_2.

You should know from linear algebra that this set of equations will only have a non-trivial solution if

$$\begin{vmatrix} H_{11}-E_\varphi S_{11} & H_{12}-E_\varphi S_{12} \\ H_{12}-E_\varphi S_{12} & H_{22}-E_\varphi S_{22} \end{vmatrix}=0$$

This determinant is called a ***Secular Determinant.***

When we multiply out the determinant, we will get a quadratic equation in *E* which is called a ***Secular Equation***. Solving the secular equation will give us two values of *E*. The lowest one will be the approximation to the ground state energy. The next highest root will be an upper limit to the next energy level.

If one takes the value of *E* obtained from this procedure and substitutes it back into the set of linear equations, one can then solve for the values of c_1 and c_2 which gives the best estimate of Φ.

This whole procedure can be generalized to the case in which we use *N* functions rather than just 2 functions. In that case the secular determinant is

$$\begin{vmatrix} H_{11}-E_\varphi S_{11} & H_{12}-E_\varphi S_{12} & \dots & H_{1N}-E_\varphi S_{1N} \\ H_{12}-E_\varphi S_{12} & H_{22}-E_\varphi S_{22} & & \vdots \\ \vdots & & & \vdots \\ H_{1N}-E_\varphi S_{1N} & \cdots & \dots & H_{NN}-E_\varphi S_{NN} \end{vmatrix}=0$$

In this case, when we multiply out the determinant, we get an N^{th} order secular equation with *N* roots for the energy *E*. The lowest root is an upper limit to the ground state energy. The estimate of the energy gets better as you add more functions, but you can see that the needed computational power also increases.

If we choose our functions f_n to be an orthonormal set, then all the off-diagonal overlap integrals will equal zero and all the diagonal ones will equal one.

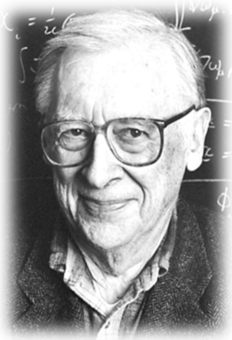
That is: $S_{ij}=\delta_{ij}$

The secular determinant then becomes

$$\begin{vmatrix} H_{11}-E_{\varphi} & H_{12} & \cdots & H_{1N} \\ H_{12} & H_{22}-E_{\varphi} & & \vdots \\ \vdots & & & \vdots \\ H_{1N} & \cdots & \cdots & H_{NN}-E_{\varphi} \end{vmatrix} = 0$$

This approach is one of the two major techniques used in computational chemistry for solving the electronic Schrödinger equation; we will discuss this later. The key to making it work well is to choose the right trial functions.

John Pople and Walter Kohn won the 1998 Nobel Prize in chemistry for their work developing and using these two different techniques.



John Pople

Walter Kohn