Notes for Lecture 9

The Variational Principle

Suppose there is a principle that is very useful practically and is very easy to prove. Suppose also that, despite these pleasant features, the answer that it gives is never quite exact, unless you get really really lucky with your intuition, or, more realistically, the luck finds you since you trained your intuition so hard. How should we think about such a principle? Should we complain about it since the solution is almost never exact? No. We value such a principle very much! It is one of the cherished tools to attack unsolvable problems of the real world!

9.1 The Variational Principle

As remarked above, the variational principle is very easy to state and prove.

$$\left\langle \psi \left| \hat{H} \right| \psi \right\rangle \geq E_0 \qquad\qquad \text{for any state } |\psi\rangle \tag{9.1}$$

where E_0 is the ground state energy, i.e. the lowest energy eigenvalue of $\hat{H}$.

To prove, we use the resolution of identity for the energy eigenstates. Here, we will assume that the energy spectrum is discrete, while the proof for the continuum case or the mixed case is left for your exercise. Assuming that $|E_n\rangle$'s are energy eigenstates, with n being just a serial number for eigenvalues[1]: $E_0 \leq E_1 \leq \ldots$, we get

$$\begin{aligned}
\left\langle \psi \left| \hat{H} \right| \psi \right\rangle &= \langle \psi | \hat{H} \sum_n |E_n\rangle \langle E_n | \psi \rangle && \text{Resolution of identity} \quad (9.2) \\
&= \sum_n E_n \langle \psi | E_n \rangle \langle E_n | \psi \rangle && \hat{H} |E\rangle_n = E_n |E\rangle_n \quad (9.3) \\
&= \sum_n E_n |C_n|^2 && C_n \equiv \langle \psi | E_n \rangle,\ C_n^* = C_n^\dagger = \langle \psi | E_n \rangle^\dagger = \langle E_n | \psi \rangle \quad (9.4) \\
&\geq E_0 \sum_n |C_n|^2 && \text{because } E_0 \leq E_n \quad (9.5) \\
&= E_0 && \text{Probability sum rule: } \sum_n |C_n|^2 = 1 \quad (9.6)
\end{aligned}$$

Essentially, this proof amounts to the following observation. For an arbitrary state not exactly in the ground state there is a probability that it is in excited states—so its energy expectation value is naturally higher than the ground state energy.

9.2 Some remarks

The principle is almost trivial, but why give it a grand name like "the variational principle"? Also, as stated above, it concerns only with the ground state energy, i.e. only one of infinitely many energy values that one could obtain if one were to solve the problem exactly!

The answers to these questions is that (1) the ground state is singularly important to learn the physics of any system, (2) there are infinite real world problems for which, even when the problem is simplified to bare essence, the exact solution is unavailable, and (3) there are problems for which the perturbation approach simply fails. Big problems that correspond to (2) and (3) are numerous, and they tend to make a list of problems for which Nobel prizes have been given for progress in theory (BCS superconductivity, Kondo effect, quantum Hall effect, asymptotic freedom in QCD).

The variational principle can be *really useful* for these real and really tough problems, while it goes without saying that it pales in comparison to exact solutions, if available.

Of course, there are other approximate methods than the variational principle. The mean field theory (e.g. Hartree theory or Hartree-Fock theory) is an example. Generally, the answer provided by the variable principle is better, assuming that the variational wave function is chosen well.

---

[1] So, please do not confuse this with n in the Hydrogen-like atom problem. Here, we are using n as a collective group index for all quantum numbers that characterize an energy eigenstate.

In practice, what happens is that one writes a variational wave function, by making an intelligent guess, which contains one or more variational parameters. The energy expectation value of such **a trial wave function** is easy to compute. And the minimization of the energy expectation value with respect to variational parameters leads to an estimate of the ground state energy. Due to the above principle, that estimate is always an upper bound.

In addition, note that this principle can be extended to an excited state, if one knows the symmetry property. For instance, a variational wave function for a parity symmetric potential in one dimension can be written as an even function to give the ground state energy, but it can be written as an odd function to give the first excited energy.

9.3 Helium atom problem

For an unsolvable problem for which the variational principle can be really helpful, let us look at the problem of the Helium atom.

Please read other more elementary examples in the textbook. In particular, the Hydrogen molecule ion in Section T7.3 should be of interest—I leave it to your reading. You will see that the structure of this problem is much like Homework 3.1 or any 2×2 matrix problem that we dealt with so far in connection with the degenerate perturbation (Lecture 5).

Now, the Helium atom problem.

$$\begin{aligned} \hat{H} &= \hat{H}_1 + \hat{H}_2 + \hat{V}_{ee} && (9.7) \\ \hat{H}_1 &= \frac{\hat{p}_1^2}{2m_e} - \frac{2e^2}{4\pi\epsilon_0 \hat{r}_1} && (9.8) \\ \hat{H}_2 &= \frac{\hat{p}_2^2}{2m_e} - \frac{2e^2}{4\pi\epsilon_0 \hat{r}_2} && (9.9) \\ \hat{V}_{ee} &= \frac{e^2}{4\pi\epsilon_0 \left|\hat{\vec{r}}_1 - \hat{\vec{r}}_2\right|} && (9.10) \end{aligned}$$

So, the Hamiltonian, $\hat{H}_1$ is Eq. 6.12 with $Z = 2$, the Hamiltonian for *one* electron in the presence of a Helium nucleus. The Hamiltonian $\hat{H}_2$ is the same, except that it is written in terms of $\hat{r}_2$ and $\hat{p}_2$. This is, of course, due to the fact that our Hilbert space is doubled due to there being two electrons in the problem, instead of one.

This problem is unsolvable due to $\hat{V}_{ee}$. What might we do about this problem?

Let us start by recalling what the $1s$ wave function looks like for the Hydrogen atom.

$$\psi_{100}(\vec{r}) = \sqrt{\frac{1}{\pi a_B^3}} \exp\left(-\frac{r}{a_B}\right) \qquad \text{for Hydrogen } (Z = 1) \tag{9.11}$$

Here, the spatial wave function is written with the quantum numbers, nlm_l, as the subscript. This wave function is the product of the radial wave function $R_{10}(r)$ and the s-wave spherical Harmonic function, $Y_{00}(\theta, \phi) = \sqrt{1/(4\pi)}$.

Now, what happens to this wave function when Z in Eq. 6.12 is taken into account? As can be guessed from the form of various r integrals (e.g., Eqs. 7.6–7.9), the answer is very simple: $a_B \to a_B/Z$. That is, the length scale of the problem gets reduced by the factor $1/Z$. So, we can write

$$\psi_{100}(\vec{r}) = \sqrt{\frac{Z^3}{\pi a_B^3}} \exp\left(-\frac{Zr}{a_B}\right) \qquad \text{Hydrogenic atom} \tag{9.12}$$

$$\psi_{100}(\vec{r}) = \sqrt{\frac{8}{\pi a_B^3}} \exp\left(-\frac{2r}{a_B}\right) \qquad \text{Helium ion, He}^{1+} \tag{9.13}$$

The problem that we have now is that we do not have a Helium ion, but we have a Helium atom with two electrons. What would happen when there are two electrons, instead of one electron, in the presence of the Coulomb potential due to the Helium nucleus? Here are some useful observations.

1. Due to the fact that electrons repel each other, each electron will not get as close to the nucleus as it would in He^{1+}. So, clearly the mean distance between each electron and the nucleus will become longer.
2. Due to the fact that the other electron *screens* the positive charge of the nucleus (not completely but partially), each electron will experience a smaller positive net charge of the nucleus than $+2e$.

These two views are completely complimentary, and point to a qualitative picture in which each electron is farther away from the nucleus than in He^{1+} but closer to the nucleus than the Bohr radius (the value for the Hydrogen atom).

Given this intuition, it can be argued that a very sensible idea is to try the following variational wave function.

$$\psi_{100,\xi}(\vec{r}) = \sqrt{\frac{\xi^3}{\pi a_B^3}} \exp\left(-\frac{\xi r}{a_B}\right) \qquad \text{trial wave function; } \xi \equiv Z_{eff} \tag{9.14}$$

Note that $\xi\, e$ represents the effective charge $Z_{eff}\, e$ of the nucleus + the other electron. As such, we should expect that

$$1 < \xi < 2 \tag{9.15}$$

If our calculation gives a final ξ value outside this range, that will ring a loud alarm bell.

Now, all we have to do is to calculate the energy expectation value for this trial wave function and then minimize it with respect to ξ, to get an estimate of the ground state energy.

But, hold on ... we do not yet have a full wave function! That is, we need to realize that this is a two electron problem, and the wave function must be written as a two electron wave function.

What we are conjecturing here is that in the Helium atom, the $1s$ orbital looks like the above trial wave function for the spatial part. The two body state is then described as

$$\begin{aligned} \left|1s^2\right\rangle &= \left|(100,\xi)^2, S_{tot}=0\right\rangle & (9.16)\\ &= \left|(100,\xi),(100,\xi)\right\rangle \frac{\left|\uparrow\downarrow\right\rangle - \left|\downarrow\uparrow\right\rangle}{\sqrt{2}} & (9.17) \end{aligned}$$

Here, the first $(100,\xi)$ is for the Hilbert space corresponding to the coordinate $\vec{r}_1$, and the second $(100,\xi)$ is for the Hilbert space corresponding to the coordinate $\vec{r}_2$. Likewise, the first arrow in the spin ket corresponds to the Hilbert space for spin 1 (S_1) and the second arrow in the same ket corresponds to the Hilbert space for spin 2 (S_2). The spatial ket $|(100,\xi),(100,\xi)\rangle$ and the spin ket, e.g. $|\uparrow\downarrow\rangle$, are "multiplied" in the sense of the direct product[2], since they belong in different spaces. The fact that the spin part is anti-symmetric with respect to the exchange symmetry, i.e. it acquires a negative sign if we exchange the first spin value and the second spin value in all spin kets, ensures the overall odd symmetry under particle exchange, required for two electrons since electrons are Fermions. This is due to the fact that the (ordinary) spatial part *must be* be symmetric for $|1s^2\rangle$ under exchange of the first part and the second part, since two parts are identical (= $(100,\xi)$).

Now, let us represent the above state.

$$\left|1s^2\right\rangle \doteq \psi_{100,\xi}(\vec{r}_1)\psi_{100,\xi}(\vec{r}_2)\frac{\begin{pmatrix}1\\0\end{pmatrix}_1 \otimes \begin{pmatrix}0\\1\end{pmatrix}_2 - \begin{pmatrix}0\\1\end{pmatrix}_1 \otimes \begin{pmatrix}1\\0\end{pmatrix}_2}{\sqrt{2}} \tag{9.18}$$

---

[2] Or the direct sum, whichever terminology you prefer.

Here, the spin wave function is written explicitly as a direct product, to avoid any confusion, and the subscripts 1 and 2 in the spin wave function refer to the spin Hilbert spaces 1 and 2 respectively[3].

Because the Hamiltonian, Eq. 9.7, is spin independent, and because what we like to calculate is just $\left\langle 1s^2 \left| \hat{H} \right| 1s^2 \right\rangle$, the spin wave function is irrelevant for the current problem. All we need is

$$\left|1s^2\right\rangle \doteq \psi_{100,\xi}(\vec{r}_1)\psi_{100,\xi}(\vec{r}_2) \qquad \text{(ordinary) spatial part only} \tag{9.19}$$

With this, and keeping in mind that an inner product involves a *double* integral $\int d\vec{r}_1 d\vec{r}_2$, we are ready to solve this problem.

Before we calculate the integral, a significant amount of work is saved if we write

$$\hat{H}_1 = \hat{H}_B(\hat{p}_1, \hat{r}_1, \xi) - \frac{(2-\xi)e^2}{4\pi\epsilon_0 \hat{r}_1} = \hat{H}_B(\hat{p}_1, \hat{r}_1, \xi) + \left(\frac{2-\xi}{\xi}\right)\hat{V}_B(\hat{r}_1, \xi) \tag{9.20}$$

$$\hat{H}_2 = \hat{H}_B(\hat{p}_2, \hat{r}_2, \xi) - \frac{(2-\xi)e^2}{4\pi\epsilon_0 \hat{r}_2} = \hat{H}_B(\hat{p}_2, \hat{r}_2, \xi) + \left(\frac{2-\xi}{\xi}\right)\hat{V}_B(\hat{r}_2, \xi) \tag{9.21}$$

where $\hat{H}_B(\hat{p}, \hat{r}, \xi)$ (B for Bohr) is the Hamiltonian for an Hydrogen-like atom, Eq. 6.12, but with $Z = \xi$, and $\hat{V}_B(\hat{r}, \xi) = -\xi e^2/(4\pi\epsilon_0 \hat{r})$ is the potential energy part of it.

As $\psi_{100,\xi}(\vec{r}_1)$ is an eigenfunction of $\hat{H}_B(\hat{p}_1, \hat{r}_1, \xi)$, and since the expectation value of $\hat{V}_B(\hat{r}_1, \xi)$ is readily available if one uses the virial theorem, it is rather straightforward to calculate $\left\langle \hat{H}_1 \right\rangle$ on $|1s^2\rangle$. Regarding the double integral, $\psi_{100,\xi}(\vec{r}_2)$ simply integrates out due to the normalization.

$$\begin{aligned}
\left\langle 1s^2 \left| \hat{H}_B(\hat{p}_1, \hat{r}_1, \xi) \right| 1s^2 \right\rangle &= \int d\vec{r}_1\, \psi^*_{100,\xi}(\vec{r}_1) H_B(p_1, r_1, \xi)\psi_{100,\xi}(\vec{r}_1) \\
&\quad \times \int d\vec{r}_2\, \psi^*_{100,\xi}(\vec{r}_2)\psi_{100,\xi}(\vec{r}_2) \\
&= \int d\vec{r}_1\, \psi^*_{100,\xi}(\vec{r}_1) H_B(p_1, r_1, \xi)\psi_{100,\xi}(\vec{r}_1) && \text{wave function normalization} \\
&= -\xi^2 R \qquad (\because H_B(\xi)\,\psi_{100,\xi} = -\xi^2 R \psi_{100,\xi}) && \text{Eq. 7.10 with } Z = \xi, n = 1 \\
\left\langle 1s^2 \left| \hat{V}_B(\hat{r}_1, \xi) \right| 1s^2 \right\rangle &= 2\left\langle 1s^2 \left| \hat{H}_B(\hat{p}_1, \hat{r}_1, \xi) \right| 1s^2 \right\rangle && \text{virial theorem} \\
&= -2\xi^2 R && R = 13.61 \text{ eV}
\end{aligned}$$

Here, the virial theorem for the Hydrogen-like problem is used to deduce that $\left\langle \hat{H}_B \right\rangle = \left\langle \hat{T} \right\rangle + \left\langle \hat{V}_B \right\rangle = -\frac{1}{2}\left\langle \hat{V}_B \right\rangle + \left\langle \hat{V}_B \right\rangle = \left\langle \hat{V}_B \right\rangle/2$.

---

[3] The direct product can be seen as essentially nothing more than combining the two Hilbert spaces. The spin wave function $\begin{pmatrix} a \\ b \end{pmatrix}_1 \otimes \begin{pmatrix} c \\ d \end{pmatrix}_2$ is equivalent to $\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}_{1\otimes 2}$.

So, we get

$$\left\langle \hat{H}_1 \right\rangle = \left(\xi^2 - 4\xi\right) R$$

It should not come as a surprise that $\left\langle \hat{H}_2 \right\rangle$ is identical to this.

$$\left\langle \hat{H}_1 + \hat{H}_2 \right\rangle = \left(2\xi^2 - 8\xi\right) R \tag{9.22}$$

For the expectation value of $\hat{V}_{ee}$ we do need to evaluate the full double integral since the potential depends on both coordinates.

$$\left\langle \hat{V}_{ee} \right\rangle = \int d\vec{r}_1 d\vec{r}_2\, \psi^*_{100,\xi}(\vec{r}_1)\psi^*_{100,\xi}(\vec{r}_2)\, \frac{e^2}{4\pi\epsilon_0\,|\vec{r}_1 - \vec{r}_2|}\, \psi_{100,\xi}(\vec{r}_1)\psi_{100,\xi}(\vec{r}_2)$$

Plugging in Eq. 9.14, and using $e^2/(4\pi\epsilon_0) = 2Ra_B$, we get

$$\begin{aligned}\left\langle \hat{V}_{ee} \right\rangle &= \left(\frac{\xi^3}{\pi a_B^3}\right)^2 2Ra_B \int d\vec{r}_1 d\vec{r}_2 \exp\left(-\frac{2\xi(r_1+r_2)}{a_B}\right) \frac{1}{|\vec{r}_1 - \vec{r}_2|} \\ &= \left(\frac{\xi^3}{\pi a_B^3}\right)^2 2Ra_B \int d\vec{r}_1\, dr_2\, d\cos\theta_2\, d\phi_2\, r_2^2 \exp\left(-\frac{2\xi(r_1+r_2)}{a_B}\right) \frac{1}{|\vec{r}_1 - \vec{r}_2|}\end{aligned}$$

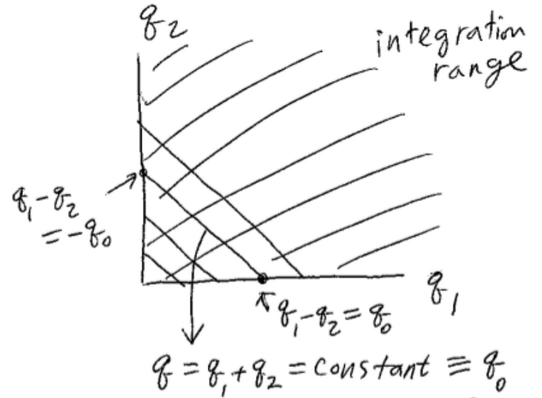
The integral over $d\Omega_2 = d\cos\theta_2\, d\phi_2$ can be easily done, if for any arbitrary but fixed $\vec{r}_1$, we set up the coordinate system for $\vec{r}_2$ so that its z axis is along $\vec{r}_1$. Then, $|\vec{r}_1 - \vec{r}_2| = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos\theta_2}$.

$$\begin{aligned}\int d\cos\theta_2\, d\phi_2\, \frac{1}{|\vec{r}_1 - \vec{r}_2|} &= 2\pi \int_{-1}^{1} d\eta\, (r_1^2 + r_2^2 - 2r_1 r_2 \eta)^{-1/2} \\ &= \frac{2\pi}{-r_1 r_2} (r_1^2 + r_2^2 - 2r_1 r_2 \eta)^{1/2}\Big|_{-1}^{1} \\ &= \frac{2\pi}{r_1 r_2}\left(r_1 + r_2 - |r_1 - r_2|\right)\end{aligned}$$

Clearly the integral over $d\Omega_1 = d\cos\theta_1\, d\phi_1$ is now trivial, and we get

$$\begin{aligned}\left\langle \hat{V}_{ee} \right\rangle &= \frac{16\xi^6 R}{a_B^5} \int dr_1 dr_2 \exp\left(-\frac{2\xi(r_1+r_2)}{a_B}\right) r_1 r_2 \left(r_1 + r_2 - |r_1 - r_2|\right) \\ &= \frac{16\xi^6 R}{a_B^5} \left(\frac{a_B}{2\xi}\right)^5 \int_0^\infty dq_1 \int_0^\infty dq_2 \exp\left(-(q_1+q_2)\right) q_1 q_2 \left(q_1 + q_2 - |q_1 - q_2|\right)\end{aligned}$$

where the substitution $q_i \equiv 2\xi r_i/a_B$ has been made. Now, let us make the change of variables $q = q_1 + q_2$ and $p = q_1 - q_2$. The integration range is easy to figure out (see diagram), and $dq_1 dq_2 = dq\, dp/2$ considering the Jacobian matrix (see diagram).



$p \equiv q_1 - q_2$

$$\therefore \text{ Integral range } = \int_0^\infty dq \int_{-q}^{q} dp$$

$$p = q_1 - q_2$$
$$q = q_1 + q_2$$

$$\overleftrightarrow{J} \equiv \frac{\partial(p,q)}{\partial(q_1,q_2)} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \qquad \left|\overleftrightarrow{J}\right| = 2$$

$$dq_1\, dq_2 = \frac{dp\, dq}{\left|\frac{\partial(p,q)}{\partial(q_1,q_2)}\right|} = \frac{1}{2}\, dp\, dq$$

$$\begin{aligned}
\left\langle \hat{V}_{ee} \right\rangle &= \frac{\xi R}{4} \int_0^\infty dq \int_{-q}^{q} dp\, e^{-q}\, \frac{q^2 - p^2}{4}\, (q - |p|) \\
&= \frac{\xi R}{8} \int_0^\infty dq\, e^{-q} \int_0^q dp\, (q^3 - pq^2 - qp^2 + p^3) \\
&= \frac{\xi R}{8} \int_0^\infty dq\, e^{-q} q^4 \left(1 - \frac{1}{2} - \frac{1}{3} + \frac{1}{4}\right) \\
&= \frac{5\xi R}{8 \cdot 12} \times 4! \\
&= \frac{5\xi R}{4}
\end{aligned}$$

Therefore, we get, by adding this to Eq. 9.22,

$$\left\langle \hat{H} \right\rangle = \left(2\xi^2 - \frac{27}{4}\xi\right) R \tag{9.23}$$

This expression attains the minimum value when

$$\xi = Z_{eff} = \frac{27}{16} = 1.69 \tag{9.24}$$

Indeed, this number is greater than 1 and less than 2, as we expected. And the minimum energy value is given by

$$\left\langle \hat{H} \right\rangle_{\xi=\xi_{min}} = -\frac{3^6}{2^7} R = -77.489 \text{ eV} \tag{9.25}$$

The experimental value is -78.975 eV, which is very close to this value!

Note that had we been very naive and ignored the electron-electron interaction $\hat{V}_{ee}$ completely, then we would have found that the state $|1s\rangle^2$ with $\xi = 2$ *is* the eigenstate of the Hamiltonian $\hat{H}_1 + \hat{H}_2$, with its eigenvalue given by Eq. 9.22 with $\xi = 2$. That value is $-8R = -109$ eV, which is quite a poor solution, compared to the above value obtained by the variational principle. Had we been less naive, but not so enlightened, as to use Eq. 9.23 with $\xi = 2$ to estimate the ground state energy, then we would have obtained $-5.5R = -74.9$ eV, which is great improvement over -109 eV, but still a bit short compared with the above variational value.