

Advanced Quantum Physics

Week 7

Michel Ferrero
Thomas Schäfer

Fall 2019

Perturbation theory

We have seen many examples where we were able to derive formal analytical solutions for the stationary Schrödinger equation. However, in the vast majority of practical cases, it is impossible to obtain such exact solutions. For example, finding the solution of a problem as simple as the helium atom is already a formidable task. The description of a silicon atom requires a Hilbert space for 14 electrons in a three-dimensional space, namely $\mathcal{E}_H = \mathcal{L}^2(\mathbb{R}^{3 \times 14}) = \mathcal{L}^2(\mathbb{R}) \otimes \mathcal{L}^2(\mathbb{R}) \otimes \cdots \otimes \mathcal{L}^2(\mathbb{R})$. Even if we only keep 10 states (e.g. Hermite polynomials) for every $\mathcal{L}^2(\mathbb{R})$ subspace, the Hamiltonian would be described by a $10^{42} \times 10^{42}$ matrix! Needless to say that manipulating such a matrix is completely out of reach, even with modern computational tools. There is hence a need for tractable methods that capture the essence of a problem even if they are approximate. This is something that Paul Dirac realized very early:

The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed which can lead to an explanation of the main features of complex atomic systems without too much computation.

The principle of perturbation theory

Perturbation theory relies on the idea that the Hamiltonian of the system can be decomposed into

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{W},$$

where $\hat{\mathcal{H}}_0$ is a Hamiltonian whose solution we know and where \hat{W} is a small perturbation. When exactly \hat{W} can be considered a small perturbation will be discussed more later. For now, we suppose that we can write $\hat{W} = \lambda \hat{\mathcal{H}}_1$ where λ is a small parameter and $\hat{\mathcal{H}}_1$ is "of the same order" as $\hat{\mathcal{H}}_0$. We write the eigenvalues and eigenstates of $\hat{\mathcal{H}}_0$ as

$$\hat{\mathcal{H}}_0 |n, r\rangle = E_n |n, r\rangle$$

where n labels the energy eigenvalues and $r \in \{1, \dots, g_n\}$ labels different degenerate eigenstates. We are looking for the eigenstates and eigenenergies of the full Hamiltonian

$$\hat{\mathcal{H}}|\Psi(\lambda)\rangle = E(\lambda)|\Psi(\lambda)\rangle$$

The idea of perturbation theory is to expand the eigenstates $|\Psi(\lambda)\rangle$ and eigenenergies $E(\lambda)$ in powers of λ :

$$\begin{aligned} |\Psi(\lambda)\rangle &= |\Psi^{(0)}\rangle + \lambda|\Psi^{(1)}\rangle + \lambda^2|\Psi^{(2)}\rangle + \dots \\ E(\lambda) &= E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots \end{aligned}$$

Using these expressions in

$$(\hat{\mathcal{H}}_0 + \lambda\hat{\mathcal{H}}_1)|\Psi(\lambda)\rangle = E(\lambda)|\Psi(\lambda)\rangle$$

and collecting terms that correspond to identical powers of λ , one obtains

$$\begin{aligned} \hat{\mathcal{H}}_0|\Psi^{(0)}\rangle &= E^{(0)}|\Psi^{(0)}\rangle \\ \hat{\mathcal{H}}_0|\Psi^{(1)}\rangle + \hat{\mathcal{H}}_1|\Psi^{(0)}\rangle &= E^{(0)}|\Psi^{(1)}\rangle + E^{(1)}|\Psi^{(0)}\rangle \\ \hat{\mathcal{H}}_0|\Psi^{(2)}\rangle + \hat{\mathcal{H}}_1|\Psi^{(1)}\rangle &= E^{(0)}|\Psi^{(2)}\rangle + E^{(1)}|\Psi^{(1)}\rangle + E^{(2)}|\Psi^{(0)}\rangle \end{aligned}$$

If one looks at the first of these equations, one recognizes that $|\Psi^{(0)}\rangle$ is one of the eigenstates associated to $E^{(0)} = E_n$. Starting from this eigenstate, the next terms, that can be computed successively, are corrections in powers of λ introduced by the perturbation \hat{W} . This method for computing eigenstates and eigenvalues in the presence of a perturbation is called perturbation theory. It assumes that the perturbed states are only a small modification of the original unperturbed states, an assumption called *adiabatic continuity*. It turns out that perturbation theory is applicable to a large body of systems. It is probably the most important method for solving problems in quantum mechanics and is widely used in atomic physics, condensed matter and particle physics.

Non-degenerate perturbation theory

Let us start again from

$$\hat{\mathcal{H}}_0|\Psi^{(0)}\rangle = E^{(0)}|\Psi^{(0)}\rangle$$

and suppose that $E^{(0)} = E_n$ for a given n and that the associated eigenvalue is *non-degenerate*. In that case we know that $|\Psi^{(0)}\rangle = |n\rangle$ and we drop the index r which is irrelevant when there is no degeneracy. What we are going to compute is the correction to the n^{th} eigenstate and eigenvalue brought by \hat{W} .

First-order perturbation theory

We first consider the equation at order 1 in λ , namely

$$\hat{\mathcal{H}}_0|\Psi^{(1)}\rangle + \hat{\mathcal{H}}_1|n\rangle = E_n|\Psi^{(1)}\rangle + E^{(1)}|n\rangle$$

By multiplying on both sides by $\langle m|$, one obtains

$$\langle m|\hat{\mathcal{H}}_0|\Psi^{(1)}\rangle + \langle m|\hat{\mathcal{H}}_1|n\rangle = E_n\langle m|\Psi^{(1)}\rangle + E^{(1)}\langle m|n\rangle$$

Because $|m\rangle$ is an eigenstate of $\hat{\mathcal{H}}_0$ with energy E_m this can be reorganized into

$$\langle m|\hat{\mathcal{H}}_1|n\rangle = (E_n - E_m)\langle m|\Psi^{(1)}\rangle + E^{(1)}\langle m|n\rangle$$

A first important result is obtained when $m = n$. Indeed, then we just have

$$E^{(1)} = \langle n|\hat{\mathcal{H}}_1|n\rangle$$

In other words, the first-order correction to the n^{th} energy is the average value of $\hat{\mathcal{H}}_1$ in the *unperturbed* state $|n\rangle$.

When $m \neq n$ we have that

$$\langle m|\Psi^{(1)}\rangle = \frac{\langle m|\hat{\mathcal{H}}_1|n\rangle}{E_n - E_m} \quad (m \neq n)$$

We have an equation for the projection of $|\Psi^{(1)}\rangle$ on the eigenstates $|m\rangle$ with $m \neq n$. This does not tell us what the projection of $|\Psi^{(1)}\rangle$ on $|n\rangle$ is. So the general expression for $|\Psi^{(1)}\rangle$ is

$$|\Psi^{(1)}\rangle = \alpha|n\rangle + \sum_{m \neq n} \frac{\langle m|\hat{\mathcal{H}}_1|n\rangle}{E_n - E_m} |m\rangle$$

One can shown that asking $\langle \Psi^{(1)}|\Psi^{(1)}\rangle = 1$ imposes that α is a purely imaginary number that basically acts as a modification of the overall phase of $|\Psi^{(1)}\rangle$. This global phase is usually chosen such that $\alpha = 0$ and we have

$$|\Psi^{(1)}\rangle = \sum_{m \neq n} \frac{\langle m|\hat{\mathcal{H}}_1|n\rangle}{E_n - E_m} |m\rangle \quad \text{and} \quad \langle n|\Psi^{(1)}\rangle = 0$$

Second-order perturbation theory

There are many practical cases where the first-order correction to the energy vanishes because the matrix element $\langle n|\hat{\mathcal{H}}_1|n\rangle = 0$. One then has to go to the equations at second-order in λ

$$\hat{\mathcal{H}}_0|\Psi^{(2)}\rangle + \hat{\mathcal{H}}_1|\Psi^{(1)}\rangle = E_n|\Psi^{(2)}\rangle + E^{(1)}|\Psi^{(1)}\rangle + E^{(2)}|n\rangle$$

Multiplying by $\langle n|$ on both sides yields

$$E_n\langle n|\Psi^{(2)}\rangle + \langle n|\hat{\mathcal{H}}_1|\Psi^{(1)}\rangle = E_n\langle n|\Psi^{(2)}\rangle + E^{(1)}\langle n|\Psi^{(1)}\rangle + E^{(2)}$$

We then find the second-order correction to the energy

$$E^{(2)} = \sum_{m \neq n} \frac{|\langle m|\hat{\mathcal{H}}_1|n\rangle|^2}{E_n - E_m}$$

Note that this result is clearly independent from the global phase chosen for the state $|\Psi^{(1)}\rangle$.

We have obtained the following result. If we start from the *non-degenerate* unperturbed eigenstate $|n\rangle$ with eigenenergy E_n and introduce a perturbation \hat{W} , the modification of the state and its energy are given by

$$E = E_n + \langle n|\hat{W}|n\rangle + \sum_{m \neq n} \frac{|\langle m|\hat{W}|n\rangle|^2}{E_n - E_m} + \dots$$

$$|\Psi\rangle = |n\rangle + \sum_{m \neq n} \frac{\langle m|\hat{W}|n\rangle}{E_n - E_m} |m\rangle + \dots$$

Several observations can be made here

- Perturbation theory is valid when the corrections to the original state are small, namely when

$$|\langle m|\hat{W}|n\rangle| \ll |E_n - E_m|$$

This is giving a more precise definition of what one means when saying that \hat{W} is a small perturbation of $\hat{\mathcal{H}}_0$.

- If we consider a non-degenerate ground state $|0\rangle$ such that $\langle 0|\hat{W}|0\rangle = 0$, then the modified ground-state energy is always lower.

$$\sum_{m>0} \frac{|\langle m|\hat{W}|0\rangle|^2}{E_0 - E_m} < 0$$

- In a perturbation theory, the energy of the n^{th} state is mainly modified by those states that are close in energy, if they are connected via a matrix element of \hat{W} .

Degenerate perturbation theory

In the section above, we have considered the modification induced by \hat{W} on a non-degenerate eigenstate of the Hamiltonian $\hat{\mathcal{H}}_0$. Here, we will inspect the case when the state belongs to a degenerate eigensubspace of the non-interacting Hamiltonian. We start again from the equation at order 0

$$\hat{\mathcal{H}}_0|\Psi^{(0)}\rangle = E^{(0)}|\Psi^{(0)}\rangle$$

Now we consider a degenerate energy $E^{(0)} = E_n$, so that $|\Psi^{(0)}\rangle \in \mathcal{E}_n$, where \mathcal{E}_n is the eigensubspace associated to the energy E_n . We can write

$$|\Psi^{(0)}\rangle = \sum_{r=1}^{g_n} c_r |n, r\rangle$$

At order λ we have

$$\hat{\mathcal{H}}_0|\Psi^{(1)}\rangle + \hat{\mathcal{H}}_1|\Psi^{(0)}\rangle = E_n|\Psi^{(1)}\rangle + E^{(1)}|\Psi^{(0)}\rangle$$

Multiplying on both sides by $\langle n, r|$ we obtain

$$E_n\langle n, r|\Psi^{(1)}\rangle + \langle n, r|\hat{\mathcal{H}}_1|\Psi^{(0)}\rangle = E_n\langle n, r|\Psi^{(1)}\rangle + E^{(1)}\langle n, r|\Psi^{(0)}\rangle$$

Using the expression for $|\Psi^{(0)}\rangle$ above, this becomes

$$\sum_{r'=1}^{g_n} \langle n, r|\hat{\mathcal{H}}_1|n, r'\rangle c_{r'} = E^{(1)}c_r$$

$$\sum_{r'=1}^{g_n} \langle n, r|\hat{W}|n, r'\rangle c_{r'} = \lambda E^{(1)}c_r$$

We can see that the equation at order λ imposes some constraint on the values of the coefficients c_r . Namely, the equation introduces a correction at order 0 of the perturbed states. The equation can be seen as an eigenvalue equation, written in matrix form

$$\begin{pmatrix} \langle n, 1|\hat{W}|n, 1\rangle & \dots & \langle n, 1|\hat{W}|n, g_n\rangle \\ \langle n, 2|\hat{W}|n, 1\rangle & \dots & \langle n, 2|\hat{W}|n, g_n\rangle \\ \vdots & & \vdots \\ \langle n, g_n|\hat{W}|n, 1\rangle & \dots & \langle n, g_n|\hat{W}|n, g_n\rangle \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{g_n} \end{pmatrix} = \delta E^{(1)} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{g_n} \end{pmatrix},$$

where $\delta E^{(1)} = \lambda E^{(1)}$ is the first-order correction to the energy. We see that in order to find the perturbed states at order 0 and the corresponding perturbed energies at order 1, one has to diagonalize the restriction of \hat{W} in the eigensubspace \mathcal{E}_n associated to the unperturbed energy E_n . In doing so, one will in general lift the degeneracy of the unperturbed states. Another way to think about degenerate perturbation theory is to imagine we write the full Hamiltonian as a matrix

$$\hat{H} = \hat{H}_0 + \hat{W} =$$

At first order in λ , one neglects all matrix elements $\langle m, r_m|\hat{W}|n, r_n\rangle = 0$ that connect blocks associated to different unperturbed subspaces $m \neq n$. It is as if the matrix was treated as a block-diagonal matrix. This approximation will be justified as long as

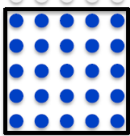
$$|\langle m, r_m|\hat{W}|n, r_n\rangle| \ll |E_m - E_n|$$

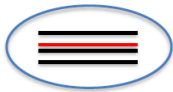
Quasi-degenerate perturbation theory

Sometimes, the unperturbed energy spectrum does not display perfect degeneracies but is still organized in groups of eigenenergies that are well separated. In this situation, it is common to use the same strategy as in degenerate perturbation theory and only consider the restriction of the Hamiltonian in the subspace spanned by a set of unperturbed states lying close in energy, as illustrated below

$$\hat{H} = \hat{H}_0 + \hat{W} = \left(\begin{array}{c} \text{matrix of grey dots} \\ \text{with a central } 4 \times 4 \text{ block of blue dots} \end{array} \right) \quad \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array}$$


$n \in \mathcal{N}$




 $n \in \mathcal{N}$

$\updownarrow \sim$

$|\langle n | \hat{W} | m \rangle|$



Such an approach will again be justified only if the matrix elements between the states that belong to \mathcal{N} and those outside of \mathcal{N} are much smaller than the typical energy distance between these states.

Variational approach

One of the assumption behind the use of perturbation theory is that the original unperturbed states provide a suitable platform to construct the perturbed states. This assumption of adiabatic continuity is not always realized in quantum systems. In strongly-interacting systems, correlations can induce very strong effects, such as phase transitions to new states of matter. In that case, the perturbed states are not adiabatically connected to the non-interacting ones. To address this kind of problems, it can be useful to use a variational approach where one tries to find a suitable guess for the wavefunction and then optimizes it.

To be more specific, consider a set of trial wavefunctions $\{|\varphi(\alpha)\rangle\}$ characterized by a variational parameter α . Here we use a single variational parameter α but one can think of it as being a set of several variational parameters. Let us then write the average energy in the state $|\varphi(\alpha)\rangle$:

$$E(\alpha) = \frac{\langle \varphi(\alpha) | \hat{\mathcal{H}} | \varphi(\alpha) \rangle}{\langle \varphi(\alpha) | \varphi(\alpha) \rangle}$$

The idea of the variational approach is to find the parameter α_{\min} that minimizes the energy $E(\alpha)$. The state $|\varphi(\alpha_{\min})\rangle$ and the energy $E(\alpha_{\min})$ are then used as an approximation for the ground state and ground-state energy. The variational approximation has the following properties

- The variational energy provides an *upper bound* to the exact ground-state energy

$$E(\alpha) \geq E_0 \quad \forall \alpha$$

Indeed, let us consider any normalized state $|\psi\rangle = \sum_n c_n |n\rangle$ expressed in the Hamiltonian eigenstates basis. We have $\langle E \rangle = \sum_n |c_n|^2 E_n \geq \sum_n |c_n|^2 E_0 = E_0$.

- The state $|\varphi(\alpha_{\min})\rangle$ will not, in general, be an eigenstate of the Hamiltonian.
- The variational approach is usually much better at finding an approximation to the ground-state energy than at finding a good approximation for the wavefunction. There are many examples where the variational energy is only 1% away from the exact answer, but the corresponding wavefunction is very different from the true answer.

Generalization to excited states

In the discussion above, we only designed a scheme that aims at finding an approximation to the ground-state energy. The variational principle can however be generalized to excited states by observing the following important property (the Rayleigh-Ritz variational principle)

The functional $E(|\varphi\rangle)$ acting on elements of the Hilbert space and yielding

$$E(|\varphi\rangle) = \frac{\langle \varphi | \hat{\mathcal{H}} | \varphi \rangle}{\langle \varphi | \varphi \rangle}$$

is *stationary* in $|\varphi\rangle$ if and only if $|\varphi\rangle$ is an eigenstate of $\hat{\mathcal{H}}$

$$\hat{\mathcal{H}}|\varphi\rangle = E|\varphi\rangle \quad \Leftrightarrow \quad \frac{\delta E}{\delta|\varphi\rangle} = 0$$

This is easy to prove by differentiating $\langle\varphi|\varphi\rangle E = \langle\varphi|\hat{\mathcal{H}}|\varphi\rangle$

$$\langle\delta\varphi|\varphi\rangle E + \langle\varphi|\delta\varphi\rangle E + \langle\varphi|\varphi\rangle \delta E = \langle\delta\varphi|\hat{\mathcal{H}}|\varphi\rangle + \langle\varphi|\hat{\mathcal{H}}|\delta\varphi\rangle$$

which yields

$$\langle\varphi|\varphi\rangle \delta E = \langle\delta\varphi|(\hat{\mathcal{H}} - E)|\varphi\rangle + \text{c.c.}$$

Now we see that

- If $|\varphi\rangle$ is an eigenstate of the Hamiltonian

$$\hat{\mathcal{H}}|\varphi\rangle = E|\varphi\rangle \quad \Rightarrow \quad \delta E = 0$$

and $E(|\varphi\rangle)$ is stationary with respect to $|\varphi\rangle$.

- If $\delta E = 0 \quad \forall |\varphi\rangle$ then it must also be true for the special choice $|\delta\varphi\rangle = \eta(\hat{\mathcal{H}} - E)|\varphi\rangle$, where $\eta \rightarrow 0$. Therefore

$$\eta \langle\varphi|(\hat{\mathcal{H}} - E)^2|\varphi\rangle + \text{c.c.} = 0 \quad \Rightarrow \quad \|(\hat{\mathcal{H}} - E)|\varphi\rangle\|^2 = 0 \quad \Rightarrow \quad \hat{\mathcal{H}}|\varphi\rangle = E|\varphi\rangle$$

and $|\varphi\rangle$ is an eigenstate of the Hamiltonian.

This important property of stationarity allows to generalize the variational approach described above. If one considers a set of trial wavefunctions $\mathcal{E}_{\text{trial}} \subset \mathcal{E}_H$ within the full Hilbert space, one obtains an approximation of the eigenstates of $\hat{\mathcal{H}}$ by finding those state $|\varphi\rangle \in \mathcal{E}_{\text{trial}}$ for which the functional $E(|\varphi\rangle)$ is stationary. The corresponding values of $E(|\varphi\rangle)$ are approximations for the eigenenergies.

The linear variational method (Ritz method)

When the set of trial wavefunctions $\mathcal{E}_{\text{trial}}$ is a vector subspace of the Hilbert space, any trial wavefunction can be written as a linear combination of states that span the subspace

$$|\varphi\rangle = \sum_n c_n |\phi_n\rangle$$

In this case, finding the stationary states is equivalent to diagonalizing the restriction of $\hat{\mathcal{H}}$ in the subspace spanned by the $\{|\phi_n\rangle\}$. One can show that in this linear variational method, the eigenvalues are *upper bounds* for the exact energies. It is interesting to note that the degenerate perturbation theory discussed above is a special case of a linear variational method.

Example: the helium atom

Let us illustrate the variational method in the calculation of the helium atom ground-state energy. The Hamiltonian for a helium atom (neglecting the fine-structure) is

$$\hat{\mathcal{H}} = -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) - \frac{e^2}{4\pi\epsilon_0} \left(\frac{Z}{r_1} + \frac{Z}{r_2} - \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right),$$

where $Z = 2$ is the atomic number of helium. In the absence of the electronic repulsion term, this Hamiltonian would just be the sum of two hydrogen atoms with $+2e$ atomic charge. The ground state would be the product of two $1s$ orbitals

$$\Psi(\vec{r}_1, \vec{r}_2) = \Psi_{1s}(\vec{r}_1)\Psi_{1s}(\vec{r}_2) = \frac{Z^3}{\pi a_0^3} e^{-Z(r_1+r_2)/a_0}$$

The energy in the absence of the interaction term would be twice the hydrogen atom energy E_1 multiplied by Z^2 , namely $2Z^2 E_1 \simeq -109eV$.

First-order perturbation theory

We can use first-order perturbation theory to get a first estimate of the effect of the interaction between the electrons. We then have to correct the energy by

$$\Delta E = \langle \Psi | \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\vec{r}_1 - \vec{r}_2|} | \Psi \rangle$$

We therefore have to compute the integral

$$\Delta E = \frac{e^2}{4\pi\epsilon_0} \left(\frac{Z^3}{\pi a_0^3} \right)^2 \int d\vec{r}_1 d\vec{r}_2 \frac{1}{|\vec{r}_1 - \vec{r}_2|} e^{-2Z(r_1+r_2)/a_0} = \frac{5}{4} \frac{e^2}{4\pi\epsilon_0} \frac{Z}{2a_0} = -\frac{5}{4} Z E_1 \simeq 34eV$$

Therefore, from first-order perturbation theory we would estimate that the ground-state energy is $E = -109 + 34 = -75eV$. This is not such a bad estimate as it is only about 5% higher than the experimental value $\simeq -79eV$.

Variational approach

But one can go one step further using the variational method. From a physical point of view, one may expect that the presence of another electron partially screens the charge of the nucleus. One can therefore construct a variational ansatz with a wavefunction

$$\Psi_\alpha(\vec{r}_1, \vec{r}_2) = \frac{\alpha^3}{\pi a_0^3} e^{-\alpha(r_1+r_2)/a_0},$$

where we replaced the atomic charge Z by a variational parameter α . We must now compute the energy in this new variational state

$$E(\alpha) = \langle \Psi_\alpha | \hat{\mathcal{H}} | \Psi_\alpha \rangle$$

The average of one of the kinetic terms is

$$\begin{aligned}
-\frac{\hbar^2}{2m} \frac{\alpha^3}{\pi a_0^3} \int d\vec{r} e^{-\alpha r/a_0} \nabla^2 e^{-\alpha r/a_0} &= -\frac{\hbar^2}{2m} \frac{\alpha^3}{\pi a_0^3} \int d\vec{r} e^{-\alpha r/a_0} \left(\frac{\partial^2}{\partial^2 r} + \frac{2}{r} \frac{\partial}{\partial r} \right) e^{-\alpha r/a_0} \\
&= -\frac{\hbar^2}{2m} \frac{\alpha^3}{\pi a_0^3} 4\pi \int dr r^2 e^{-\alpha r/a_0} \left(\frac{\alpha^2}{a_0^2} - \frac{2}{r} \frac{\alpha}{a_0} \right) e^{-\alpha r/a_0} \\
&= -\frac{\hbar^2}{2m} \frac{\alpha^3}{\pi a_0^3} 4\pi \left(\frac{\alpha^2}{a_0^2} \frac{a_0^3}{4\alpha^3} - \frac{\alpha}{a_0} \frac{a_0^2}{2\alpha^2} \right) \\
&= -\frac{\hbar^2}{2m} \frac{\alpha^3}{\pi a_0^3} 4\pi \left(\frac{a_0}{4\alpha} - \frac{a_0}{2\alpha} \right) = \frac{\hbar^2}{2m} \frac{\alpha^2}{a_0^2} = -\alpha^2 E_1
\end{aligned}$$

For two electrons, the total kinetic energy is therefore given by $-2\alpha^2 E_1$. The potential energy is instead given by

$$\begin{aligned}
-\frac{e^2}{4\pi\epsilon_0} \frac{\alpha^3}{\pi a_0^3} \int d\vec{r} \frac{Z}{r} e^{-2\alpha r/a_0} &= -\frac{e^2}{4\pi\epsilon_0} \frac{\alpha^3}{\pi a_0^3} 4\pi \int dr Z r e^{-2\alpha r/a_0} \\
&= -\frac{e^2}{4\pi\epsilon_0} \frac{\alpha^3}{\pi a_0^3} 4\pi \frac{a_0^2 Z}{4\alpha^2} = 2\alpha Z E_1
\end{aligned}$$

The total potential energy is $4\alpha Z E_1$. Finally, the interaction energy is obtained from the same calculation as the one we did just above and yields

$$\langle \Psi_\alpha | \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\vec{r}_1 - \vec{r}_2|} | \Psi_\alpha \rangle = -\frac{5}{4} \alpha E_1$$

If we put all these terms together we find the variational energy

$$E(\alpha) = \left(-2\alpha^2 + 4\alpha Z - \frac{5}{4}\alpha \right) E_1$$

This energy is minimized for

$$\alpha_{\min} = Z - \frac{5}{16} \quad E(\alpha_{\min}) \simeq -77.5 eV$$

This is only 2% greater than the experimental value $\simeq 79 eV$. It also shows that the presence of another electron leads to a shielding of the nuclear charge by an amount of about $(5/16)e$.