Exact Results

Alston J. Misquitta

Centre for Condensed Matter and Materials Physics
Queen Mary, University of London

February 20, 2019

Outline of the lecture I

We are going to be dealing with many-body (i.e., many electron) wavefunctions. So we will begin by looking at what we know about these wavefunctions and what tools we have to analyse them:

- Properties of Ψ: All the exact properties the wavefunction (and density) are expected to satisfy. Afterall, any approximate wavefunction should satisfy as many of these as is practically possible.
- Methods for approximating the wavefunction. We cannot solve much beyond the 1-electron, hydrogen atom. For anything more complex, we must solve the Schrödinger equation approximately.
- Mathematical methods: We will need some advanced mathematical methods in this course.

Atomic Units I

Atomic units will be used throughout: effectively $\hbar = 1$, $m_{\mathrm{e}} = 1$, $e = 1$.

- Mass: Has units of free-electron mass m. SI value is 9.10938×10^{-31} kg.
- Charge: Has units of absolute value of free-electron charge e. SI value is 1.60218×10^{-19} C.
- Angular momentum: Units of reduced Planck's constant $\hbar$. SI value is 1.05457×10^{-34} J s.
- Length: Units of Bohr radius of the H atom, $a_0 = 4\pi\epsilon_0\hbar^2/me^2$. SI value 5.29177×10^{-11} m.

In these units $c = \alpha^{-1} = 1/137.036$, where the fine structure constant is defined as $\alpha = e^2/4\pi\epsilon_0\hbar c$.

Atomic Units II

If you are not already familiar with atomic units then please read and work through Szabo & Ostlund §2.1.1. You should be able to start from the standard form of the non-relativistic N-electron Hamiltonian and show that if we measure energies in Hartree and distances in Bohr then the electronic Hamiltonian reduces to

$$\mathcal{H}_{\mathrm{e}} = -\sum_{i}^{N}\frac{1}{2}\nabla_i^2 - \sum_{i}^{N}\sum_{\alpha}^{M}\frac{Z_\alpha}{r_{i\alpha}} + \sum_{i}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}}$$

To get to this form you will need to use the concept of coordinate scaling. This is a simple but powerful way of learning about the fundamental length and energy scales of the system.

Properties I

Properties II

Notation:

- $|\Psi\rangle$: exact many-body wavefunction
- $|0\rangle$: approximate wavefunction

Our goal here is to list properties satisfied by $|\Psi\rangle$ that we'd also like $|0\rangle$ to satisfy. Some will be obvious, others not so obvious...

Properties III

$|\Psi\rangle$ *contains N electrons*:
We definitely want $|0\rangle$ to describe the same number of electons and so expect that, if $\hat{N}$ is a number operator (more on these when we cover second-quantization), then

$$\hat{N}|0\rangle = N|0\rangle$$

i.e., $|0\rangle$ is an eigenstate of the number operator with eigenvalue N. This is not normally an issue and is made explicit when we define the wavefunction in second-quantized form (more on this later).

Properties IV

Antisymmetry

$$\hat{P}_{ij}|\Psi\rangle = -|\Psi\rangle$$

We expect $|0\rangle$ to be antisymmetric.

- In second-quantization (which we will use), the Pauli principle is built into the anticommutating relations of the creation/annihilation operators.
- In first-quantization, we need to make the wavefunction antisymmetric by expressing it as a linear combination of Slater determinants.

Properties V

Square-integrability of the bound states.

$$\langle \Psi | \Psi \rangle = 1$$

To ensure that $|0\rangle$ will always satisfy this we expand it in a basis of normalized orbitals. These orbitals will be, in turn, expanded in a basis of, say, Gaussian-type orbitals (GTOs). These orbitals are, by definition, square-integrable. Other basis sets are possible and some are not strictly integrable (plane-waves) but conditions may be imposed to ensure integrability (plane-waves can be integrated in a finite box).

Variational stability I

Variational stability
If Ψ is the exact g.s. wavefunction and δ is a small variation in Ψ giving an approximate wavefunction $\tilde{\Psi} = \Psi + \delta$, then the variational principle states that the error in the energy $E[\tilde{\Psi}]$ is only second-order in the error in variation δ, i.e.,

$$E[\tilde{\Psi}] = E[\Psi] + \mathcal{O}(\delta^2). \tag{1}$$

We will define $|0\rangle$ s.t. this principle is preserved. But not all methods gaurantee variational stability.

Variational stability II

What does the notation $\mathcal{O}(x^k)$ mean?

When we expand a quantity, $E(x)$, in terms of a small parameter x we often truncate the expansion at some power, say n. If the series was convergent then the largest of the missing terms would be the term $c_{n+1}x^{n+1}$, where c_{n+1} is a (usually unknown) coefficient. A compact way of denoting that we have skipped this and other terms (an infinite number) is to write the series as

$$E(x) = \sum_{i=0}^{n} c_i x^i + \mathcal{O}(x^{n+1}).$$

If we use the order notation then the above expression is written as an identity, not an approximation. This is because the last term is simply a short form for the missing terms.

Asymptotic form I

Asymptotic form of $|\Psi\rangle$:

$$\Psi \to \Psi_{N-1} e^{-\sqrt{2E_I}\, r},$$

where $E_I = E_{N-1} - E_N$ is the vertical first ionisation energy of the system.

This result holds for any bound state. It is quite difficult to satisfy as the slow decay of the exponential function can be only approximately modelled by GTOs.

Cusp conditions I

The electron–nuclear cusp condition:
The Hamiltonian has a singularity in the electon–nuclear potential, however the energy of a bound state is finite. So this singularity must be cancelled by another singularity of the opposite sign. This is possible if the wavefunction has a cusp at the nuclei:

$$\left.\frac{\partial \Psi}{\partial r_{i\alpha}}\right|_{r_{i\alpha}=0} = -Z_\alpha \Psi(r_{i\alpha} = 0)$$

Cusp conditions II

The electron–electron cusp condition:
For similar reasons, there must be cusps in the wavefunction when two electrons are at the same location. In this case, the cusp condition depends on the electronic spins. For a singlet system (electrons with opposite spins) we have

$$\left.\frac{\partial \Psi}{\partial r_{ij}}\right|_{r_{ij}=0} = +\frac{1}{2}\Psi(r_{ij}=0)$$

While for a triplet, we get a $+1/4$ on the R.H.S.
This condition is responsible for correlation effects at short-range.

Size-extensivity I

Size-Extensivity
For a system containing non-interacting subsystems, the total energy is equal to the sum of energies of the individual subsystems. Consider the total Hamiltonian:

$$H_T = \sum_i^M H_i,$$

where the H_i are the Hamiltonians of non-interacting systems. If we have $H_T \Psi_T = E_T \Psi_T$, then we must have

$$E_T = \sum_i^M E_i,$$

where $H_i \Psi_i = E_i \Psi_i$.

Size-extensivity II

For a system of N non-interacting *identical* components, a size-extensive method will result in an energy of

$$E(N) = NE_0,$$

where E_0 is the energy of an individual part. So the *energy per component* is a constant: $E(N)/N = E_0$. If this is not the case, that is, if the energy per component depends on N, then the method is not size-extensive. In this case as you add (or remove) components, even though they do not interact, the energy per component will change.

Spin I

Spin

In non-relativistic theory, the exact stationary eigenstates of $\hat{S}^2$ and $\hat{S}_z$:

$$\hat{S}^2\Psi = S(S+1)\Psi$$
$$\hat{S}_z\Psi = M\Psi$$

Reminder: Atomic units used throughtout!

Born–Oppenheimer approximation I

The Born–Oppenheimer approximation

We have electrons and nuclei in our general Hamiltonian. The nuclei complicate matters. But we can simplify life by arguing that since the nuclei are nearly 2000 times heavier than the electrons, they can be considered fixed while we solve the electronic Hamiltonian. That is we solve

$$\mathcal{H}_{\mathrm{e}}\Psi_{\mathrm{e}} = E_{\mathrm{e}}\Psi_{\mathrm{e}} \tag{2}$$

where

$$\mathcal{H}_{\mathrm{e}} = -\sum_i \frac{1}{2}\nabla_i^2 - \sum_i \sum_\alpha \frac{Z_\alpha}{r_{i\alpha}} + \sum_i \sum_{j>i} \frac{1}{r_{ij}} \tag{3}$$

Born–Oppenheimer approximation II

to get the wavefunction and energy that will be parametrically dependent on the positions of the nuclei:

$$\Psi_{\mathrm{e}} = \Psi_{\mathrm{e}}(\{r_i\};\{R_\alpha\})$$
$$E_{\mathrm{e}} = E_{\mathrm{e}}(\{R_\alpha\})$$

From the latter we get our notion of an energy landscape on which the nuclei move (often assumed to be Classically using Newtons Laws - i.e, Molecular Dynamics).

Born–Oppenheimer approximation III

There are two cases then the BO approximation is invalid:

- Fast nuclei: The BO condition should not really be stated in terms of the mass ratios of the electrons and nuclei, but rather in terms of the kinetic energy ratios. If nuclei are very fast, they can have kinetic energies comparable with those of the electrons. Example: Radiation damage.
- Level crossing: This is a subtle one. If two electronic energy levels cross (often happens with excited states) then if there is a vibrational mode of appropriate symmetry, the BO approximation breaks down. This is the Jahn–Teller effect where we must consider a coupling of the electronic and nuclear motions.

We will henceforth always use the electronic Hamiltonian and wavefunction.

Asymptotic form of $|\Psi\rangle$ I

Proof of the asymptotic form of the density/wavefunction:

$$\Psi \rightarrow \Psi_{N-1} e^{-\sqrt{2E_I} r},$$

or

$$\rho(r) \rightarrow e^{-2\sqrt{2E_I} r},$$

Asymptotic form of $|\Psi\rangle$ II

We begin this proof by first proving a simpler result for a 1-electron system. Consider the Hamiltonian

$$
\begin{aligned}
\hat{H} &= -\frac{1}{2}\nabla^2 + V(\mathbf{r}) && (4) \\
&= -\frac{1}{2}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{\hat{l}^2}{r^2}\right) + V(\mathbf{r}). && (5)
\end{aligned}
$$

Asymptotic form of $|\Psi\rangle$ III

Now, as $r \to \infty$, assuming $V(r) \to 0$ (valid for all reasonable potentials, apart for constant shifts which can be absorbed into the energy), only the first term in the above Hamiltonian survives and the eigenvalue problem reduces to

$$-\frac{1}{2}\frac{\partial^2}{\partial r^2}\psi(r) = E\psi(r). \tag{6}$$

The solution of this equation is the asymptotic 1-electron wavefunction, thus

$$\psi(r) \to c e^{-\sqrt{-2E}r}. \tag{7}$$

For the H atom, $E_0 = -\frac{1}{2}$ giving $\psi_0 \to e^{-r}$.

Asymptotic form of $|\Psi\rangle$ IV

The *N*-electron case:

$$\hat{H}\Psi_N(1,2,...,N) = E_N\Psi_N(1,2,...,N) \tag{8}$$

As we pull one electron out, Ψ_N collapses into the state $\Psi_{N-1}\phi(r)$ — that is, the product on an $N-1$-electron wavefunction and a 1-electron wavefunction $\phi(r)$. It is the latter that determines the asymptotic properties of Ψ_N.

We will make the assumption that the interaction terms in $\hat{H}$ that act between Ψ_{N-1} and $\phi(r)$ can be neglected. This is valid if the electronic state Ψ_{N-1} is sufficiently compact. This results in the separable Hamiltonian:

$$\hat{H} = \hat{H}_{N-1} + \hat{H}_1. \tag{9}$$

Asymptotic form of $|\Psi\rangle$ V

Therefore

$$E_N \Psi_N = \hat{H}\Psi_N = (\hat{H}_{N-1} + \hat{H}_1)\Psi_{N-1}\phi(r) \tag{10}$$

$$= E_{N-1}\Psi_{N-1}\phi(r) + \Psi_{N-1}(\hat{H}_1\phi(r)). \tag{11}$$

And so

$$\Psi_{N-1}(\hat{H}_1\phi(r)) = (E_N - E_{N-1})\Psi_{N-1}\phi(r). \tag{12}$$

On integrating out Ψ_{N-1} we get the 1-electron eigenvalue problem:

$$\hat{H}_1\phi(r) = (E_N - E_{N-1})\phi(r). \tag{13}$$

Asymptotic form of $|\Psi\rangle$ VI

Using the result for the 1-electron case (Eq. (7)) we get

$$\phi(r) \rightarrow e^{-\sqrt{-2(E_N - E_{N-1})}r} \tag{14}$$

$$\rightarrow e^{-\sqrt{2E_I}r}, \tag{15}$$

where $E_I = E_{N-1} - E_N$ is the vertical ionization potential. This will prove to be a useful result in the analysis of density functionals.

The more general result is:

Q:

$$\Psi \rightarrow r^{\beta} e^{-\sqrt{2E_I}r}.$$

Prove it and show that $\beta = -1 + 1/\alpha$ where $\alpha = \sqrt{2E_I}$.

Electron–Nuclear cusp condition I

$$\boxed{\left.\frac{\partial \Psi}{\partial r_{i\alpha}}\right|_{r_{i\alpha}=0} = -Z_\alpha \Psi(r_{i\alpha}=0)}$$

Electron–Nuclear cusp condition II

Consider a 1-electron system (that's all we need for this cusp condition) with a nucleus of charge Z at the origin:

$$\begin{aligned} \hat{H} &= -\frac{1}{2}\nabla^2 + V(\mathbf{r}) & (16) \\ &= -\frac{1}{2}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{\hat{l}^2}{r^2}\right) - \frac{Z}{r}. & (17) \end{aligned}$$

If $\hat{H}\Psi = E\Psi$, where E is a finite, bound-state energy, then we can define the local energy function:

$$\begin{aligned} E\Psi &= \hat{H}\Psi \\ &= -\frac{1}{2}\frac{\partial^2\Psi}{\partial r^2} \\ &\quad - \frac{1}{r}\frac{\partial\Psi}{\partial r} - \frac{\hat{l}^2\Psi}{r^2} - \frac{Z}{r}\Psi. \end{aligned}$$

Electron–Nuclear cusp condition III

Now, since the energy E is finite, the L.H.S. is finite everywhere, and therefore the R.H.S. must also remain finite for all r. In fact, it must be a constant for all r! But the R.H.S. contains three terms that diverge as $r \rightarrow 0$, i.e., as the electron approaches the nucleus. The only way for the R.H.S. to remain finite is for these three terms to cancel as $r \rightarrow 0$.

The $\hat{l}^2\Psi$ term vanishes for all spherical states, and can be made to vanish for more general states by taking the spherical average about the nuclear position.

Q: How do we see this?
Hint: $\hat{l}^2\Psi$ can be expanded in terms of spherical harmonics and radial functions. What is the spherical average of Y_{lm} for $l > 0$?

Electron–Nuclear cusp condition IV

The other two terms must cancel giving:

$$0 = -\frac{2}{r}\frac{\partial \Psi}{\partial r} - \frac{Z}{r}\Psi$$

and since $\Psi(0) \neq 0$, we have

$$\left.\frac{\partial \Psi}{\partial r}\right|_{r=0} = -Z\Psi(0).$$

Or, more generally,

$$\left.\left\langle \frac{\partial \Psi}{\partial r} \right\rangle_{\mathrm{sph}}\right|_{r=0} = -Z\langle \Psi(0) \rangle_{\mathrm{sph}}.$$

Electron–Electron cusp condition I

The e-e cusp condition proofs are more involved. Here we will look at a simplified ‘proof’ of the condition.

Consider two electrons *i* and *j* approaching each other. We will work in the limit $r \equiv r_{ij} = |\mathrm{r_i} - \mathrm{r_j}| \to 0$. Here, the only interactions that matter are those involving these two electrons. Hence the effective 2-particle Schrödinger equation can be written as follows:

$$E\Psi = \left(-\frac{1}{2\mu}\nabla^2 + \frac{1}{r} \right) \Psi,$$

where the reduced mass is $\mu = 1/2$ (Why?). Now let’s use the same ideas we used for the electron-nuclear cusp. For the system in a spherically symmetric state (this requires that the electron

Electron–Electron cusp condition II

spins are opposite, i.e. we have a singlet state. A triplet state would have a node at $r_{ij} = 0$) we can write

$$\begin{aligned} E\Psi &= \hat{H}\Psi \\ &= -\frac{\partial^2\Psi}{\partial r^2} \\ &\quad -\frac{2}{r}\frac{\partial\Psi}{\partial r} - \frac{\hat{l}^2\Psi}{2r^2} + \frac{1}{r}\Psi. \end{aligned}$$

Since the L.H.S. is finite, the divergent terms on the R.H.S. need to cancel. For a spherically symmetric state, $\hat{l}^2\Psi = 0$, so we must have, at electron coalscence

$$-\frac{2}{r}\frac{\partial\Psi}{\partial r} + \frac{1}{r}\Psi = 0$$

Electron–Electron cusp condition III

or, remembering that r is short for r_{ij},

$$\left.\frac{\partial \Psi}{\partial r_{ij}}\right|_{r_{ij}=0} = +\frac{1}{2}\Psi(r_{ij}=0).$$

This holds for the singlet state only! For the triplet state instead of a half, we get a $+1/4$ (no proof). There are cusp conditions involving more than two electrons. Or indeed, the conditions could involve two electrons and a nucleus, and another other number and permutation of particles.

Variational Principle I

Variational stability

If Ψ is the exact g.s. wavefunction and δ is a small variation in Ψ giving an approximate wavefunction $\tilde{\Psi} = \Psi + \delta$, then the variational principle states that the error in the energy $E[\tilde{\Psi}]$ is only second-order in the error in variation δ, i.e.,

$$E[\tilde{\Psi}] = E[\Psi] + \mathcal{O}(\delta^2). \tag{18}$$

Variational Principle II

Let the exact (ground state) eigenvalue equation be

$$\hat{H}|0\rangle = E_0|0\rangle.$$

Let $|\tilde{0}\rangle$ be an approximation to $|0\rangle$ and let

$$|\tilde{0}\rangle = |0\rangle + |\delta\rangle$$

where $|\delta\rangle$ is an allowed variation.

Variational Principle III

$$
\begin{aligned}
E[0+\delta] &= \frac{\langle 0+\delta|H|0+\delta\rangle}{\langle 0+\delta|0+\delta\rangle} \\
&= \frac{\langle 0|H|0\rangle + \langle 0|H|\delta\rangle + \langle \delta|H|0\rangle + \langle \delta|H|\delta\rangle}{\langle 0|0\rangle + \langle 0|\delta\rangle + \langle \delta|0\rangle + \langle \delta|\delta\rangle} \\
&= \left(E_0 + \langle 0|H|\delta\rangle + \langle \delta|H|0\rangle + \mathcal{O}(\delta^2)\right) \\
&\quad \times \left(1 - (\langle 0|\delta\rangle + \langle \delta|0\rangle) + \mathcal{O}(\delta^2)\right) \\
&= E_0 + \langle 0|H - E_0|\delta\rangle + \langle \delta|H - E_0|0\rangle + \mathcal{O}(\delta^2) \\
&= E_0 + \delta E + \mathcal{O}(\delta^2)
\end{aligned}
$$

Now $\delta E = 0$ whenever $\hat{H}|0\rangle = E_0|0\rangle$, therefore the state $|0\rangle$ is a stationary state point in the energy functional $E[\tilde{0}]$.

Variational Principle IV

Conversely, we will now show that all stationary states of $E[\tilde{0}]$ are eigenstates of $\hat{H}$. Notice that here we are treating $E[\tilde{0}]$ as a *functional* of $\tilde{0}$.

Let $|0\rangle$ be a stationary point (state) of $E[\tilde{0}]$. By definition, for the variation $|\delta\rangle$ in $|0\rangle$, all terms in $E[0+\delta]$ that linear in $|\delta\rangle$ must vanish. From the previous page his means that $\delta E = 0$, or

$$\delta E = 0 = \langle 0|H - E[0]|\delta\rangle + \langle \delta|H - E[0]|0\rangle.$$

Now consider the variation $i|\delta\rangle$. This gives us the condition

$$\delta E = 0 = \langle 0|H - E[0]|\delta\rangle - \langle \delta|H - E[0]|0\rangle.$$

Therefore, adding them up we get

$$0 = \langle \delta|H - E[0]|0\rangle \quad \forall \;\; |\delta\rangle.$$

Variational Principle V

Since this must hold for all variations $|\delta\rangle$, we must have

$$\hat{H}|0\rangle = E[0]|0\rangle \equiv E_0|0\rangle,$$

i.e., $|0\rangle$ is an eigenstate of $\hat{H}$.

Hence the *variational principle* states that the solution of $\hat{H}|0\rangle = E_0|0\rangle$ is equivalent to a variational optimization of $E[\tilde{0}]$.

The Hellmann–Feynman Theorem I

For a stationary state, the first-order change in the energy due to a perturbation may be calculated as the expectation value of the perturbation operator:

$$\left.\frac{\partial E(\alpha)}{\partial \alpha}\right|_{\alpha=0} = \langle \Psi | V | \Psi \rangle .$$

Proof:
Let $H(\alpha) = H + \alpha V$ where V is the perturbation. Further, let $|\Psi_\alpha\rangle$ be a stationary state of $H(\alpha)$, i.e.,

$$H(\alpha)|\Psi_\alpha\rangle = E(\alpha)|\Psi_\alpha\rangle .$$

The Hellmann–Feynman Theorem II

This implies

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

Therefore, using the definition $|\Psi\rangle = |\Psi_\alpha\rangle|_{\alpha=0}$, and $\langle \Psi_\alpha | \Psi_\alpha \rangle = 1$,

$$\begin{aligned} \left.\frac{\partial E(\alpha)}{\partial \alpha}\right|_{\alpha=0} &= \frac{\partial}{\partial \alpha} \left.\frac{\langle \Psi_\alpha | H(\alpha) | \Psi_\alpha \rangle}{\langle \Psi_\alpha | \Psi_\alpha \rangle}\right|_{\alpha=0} \\ &= 2\Re \langle \left.\frac{\partial \Psi_\alpha}{\partial \alpha}\right|_{\alpha=0} | H - E(0) | \Psi \rangle + \langle \Psi | V | \Psi \rangle \end{aligned}$$

and since $H(0)|\Psi\rangle = E(0)|\Psi\rangle$, we get

$$\left.\frac{\partial E(\alpha)}{\partial \alpha}\right|_{\alpha=0} = \langle \Psi | V | \Psi \rangle.$$

The Hellmann–Feynman Theorem III

Properties

This theorem is useful for evaluating properties and in geometry optimizations. For example, in the presence of an electric field F along x, the Hamiltonian is $H(F) = H - Fx$. From the Hellmann–Feynman theorem we can evaluate the dipole moment μ_x as follows:

$$\left.\frac{\partial E}{\partial F}\right|_{F=0} = \langle \Psi | x | \Psi \rangle = \mu_x.$$

So, if we know the derivative of the energy w.r.t. the field F at zero field strength, $\left.\frac{\partial E}{\partial F}\right|_{F=0}$, then we will also know the dipole moment μ_x. This can be evaluated using *finite differences*:

$$\left.\frac{\partial E}{\partial F}\right|_{F=0} = \lim_{\epsilon=0} \frac{E(F=+\epsilon) - E(F=-\epsilon)}{2\epsilon}.$$

The Hellmann–Feynman Theorem IV

In practice, we must choose ϵ small enough that the derivative converges, but yet large enough that numerical noise is low. This technique can be used to calculate a number of molecular properties, but there usualy will be ways to do the same analytically. For example, a better way to calculatet the dipole moment μ_x would be to evaluate the integral:

$$\mu_x = \int \rho(\mathrm{r}) \mathrm{x} \mathrm{d} \mathrm{r}.$$

But the point of using the Hellmann–Feynman theorem is that all you need are the energies!

The Hellmann–Feynman Theorem V

Forces
One of the many important applications of this theorem is the calculation of forces. We will often want to perform a geometry optimization. One way of doing this is to perform a number of energy calculations as a function of molecular geometry, and the lowest energy structure will then be our best guess to the optimized structure. This may work in a couple of dimensions, but it will not when there are too many internal degrees of freedom. Remember that these grow as $3N - 6$, so a water molecule has 3 internal parameters to optimize, but a small lipid of 130 atoms will have $3 \times 130 - 6 = 384$ degrees of freedom!

The Hellmann–Feynman Theorem VI

Instead, the Hellmann–Feynman theorem allows us to use a variational solution to calculate the *forces* on each nucleus. The forces allow us to move the nuclei so as to (generally) reduce the energy. This gives us a new geometry at which we repeat the calculation till the forces are as close to zero as we need them to be. In practice, in addition to the forces, we also calculate the Hessians: these involve the second derivatives and give us even more information and generally allow the process to converge faster.

To see how it can be used to determine the force on a nucleus consider the Hamiltonian with nuclei at R_I displaced by α_I: $H(R_I + \alpha_I)$. Let $|\Psi\rangle$ be the eigenfunction of $H(R_I)$, i.e., at $\alpha_I = 0$.

The Hellmann–Feynman Theorem VII

We can write

$$\begin{aligned}
H(R_I+\alpha_I) &= -\frac{1}{2}\sum_i \nabla_i^2 + \sum_{j>i}\frac{1}{|r_i-r_j|} \\
&\quad -\sum_{iI}\frac{Z_I}{|r_i-R_I-\alpha_I|} \\
&\quad +\sum_{J>I}\frac{Z_I Z_J}{|R_I+\alpha_I-R_J-\alpha_J|} \\
&= H(R_I)-\sum_{iI}\frac{Z_I(r_i-R_I)\cdot\alpha_I}{|r_i-R_I|^3} \\
&\quad -\sum_{J>I}\frac{Z_I Z_J(R_I-R_J)\cdot\alpha_I}{|R_I-R_J|^3}+\mathcal{O}(\alpha^2)
\end{aligned}$$

The Hellmann–Feynman Theorem VIII

We have used the expansion in the last step as we know we need to differentiate w.r.t. α_I and then set $\alpha_I = 0$. I.e., we have written the Hamiltonian in the form $H(R_I + \alpha_I) = H(R_I) + \alpha_I \cdot V$, with

$$V = -\sum_{iI} \frac{Z_I(r_i - R_I)}{|r_i - R_I|^3} - \sum_{J>I} \frac{Z_I Z_J(R_I - R_J)}{|R_I - R_J|^3}$$

Hence, from the H-F theorem,

$$\left.\frac{\partial E}{\partial \alpha_I}\right|_{\alpha_I=0} = -\langle\Psi|\sum_{iI} \frac{Z_I(r_i - R_I)}{|r_i - R_I|^3}|\Psi\rangle - \sum_{J>I} \frac{Z_I Z_J(R_I - R_J)}{|R_I - R_J|^3}$$

This makes it easy to evaluate first derivatives of the energy w.r.t. nuclear coordinates. So we can do geometry optimizations.

The Hellmann–Feynman Theorem IX

Q: This theorem holds not only for the exact wavefunction but also for variationally optimized wavefunctions such as $|\mathbf{c}\rangle$. Prove this.

Pulay Forces: We usually obtain variational wavefunctions using nuclear-centered basis sets. In this case, the basis space alters with changes in the nuclear positions. So there are terms in the gradient w.r.t. the basis functions. These terms are called Pulay forces after Peter Pulay.

The Hellmann–Feynman Theorem X

Outline of proof for the validity of the Hellmann–Feynman theorem for variational wavefunctions.

Consider a variational wavefunction of the form $\tilde{\Psi} = \tilde{\Psi}(\{c_i\})$ with energy

$$\tilde{\mathcal{E}} = \frac{\langle \tilde{\Psi} | H(\alpha) | \tilde{\Psi} \rangle}{\langle \tilde{\Psi} | \tilde{\Psi} \rangle}.$$

At the extremum, the energy $\tilde{\mathcal{E}}$ depends explicitely on α through the Hamiltonian, but also contains an *implicit* dependence through the variational wavefunction $\tilde{\Psi}$ as the optimum values of the parameters $\{c_i\}$ are dependent on the Hamiltonian, and so they are implicitely dependent on α. That is, at the extremum, we have

$$\tilde{\mathcal{E}} = \tilde{\mathcal{E}}[\tilde{\Psi}_\alpha, \alpha] \equiv \tilde{\mathcal{E}}[\{c_i^\alpha\}, \alpha],$$

The Hellmann–Feynman Theorem XI

and we have the usual variational condition:

$$\left.\frac{\partial \tilde{\mathcal{E}}}{\partial c_i}\right|_{c_i=c_i^\alpha} = 0.$$

Here we have used α in the sub/super-script to denote the *implicit* dependence on α.
Now consider the total differential of $\tilde{\mathcal{E}}$ w.r.t. α evaluated at the extremum:

$$\begin{aligned}\frac{d\tilde{\mathcal{E}}}{d\alpha} &= \frac{\partial \tilde{\mathcal{E}}[\{c_i^\alpha\},\alpha]}{\partial \alpha} + \sum_i \left.\frac{\partial \tilde{\mathcal{E}}}{\partial c_i}\right|_{c_i=c_i^\alpha} \left.\frac{\partial c_i}{\partial \alpha}\right|_{c_i=c_i^\alpha} \\ &= \frac{\partial \tilde{\mathcal{E}}[\{c_i^\alpha\},\alpha]}{\partial \alpha}.\end{aligned}$$

The Hellmann–Feynman Theorem XII

But the explicit differential can only act on the explicit α dependence in the Hamiltonian. That is

$$\frac{\partial \tilde{\mathcal{E}}[\{c_i^\alpha\},\alpha]}{\partial\alpha} = \frac{\partial}{\partial\alpha}\left(\frac{\langle\tilde{\Psi}_\alpha|H(\alpha)|\tilde{\Psi}_\alpha\rangle}{\langle\tilde{\Psi}_\alpha|\tilde{\Psi}_\alpha\rangle}\right) = \frac{\langle\tilde{\Psi}_\alpha|\frac{dH(\alpha)}{d\alpha}|\tilde{\Psi}_\alpha\rangle}{\langle\tilde{\Psi}_\alpha|\tilde{\Psi}_\alpha\rangle},$$

and hence we get the result

$$\frac{d\tilde{\mathcal{E}}}{d\alpha} = \frac{\langle\tilde{\Psi}_\alpha|\frac{dH(\alpha)}{d\alpha}|\tilde{\Psi}_\alpha\rangle}{\langle\tilde{\Psi}_\alpha|\tilde{\Psi}_\alpha\rangle}.$$

Q.E.D.

The Virial Theorem I

You will have come across the Virial theorem in the form:

$$\langle V \rangle = -2\langle T \rangle,$$

that is, the expectation value of a Coulomb potential $\hat{V}$ is twice the expectation value of the kinetic energy operator $\hat{T}$. The *molecular virial theorem* takes the form

$$2\langle T \rangle + \langle V(R) \rangle + \langle \left. \frac{\partial V(\alpha R)}{\partial \alpha} \right|_{\alpha=1} \rangle = 0.$$

This may seem rather strange, but, as we shall see, the additional term is a consequence of the Born–Oppenhiemer approximation.

The Virial Theorem II

To prove the molecular virial theory we will use the idea of uniform scaling (we will also use this idea a lot when we take up density functional theory):

Theorem

Under uniform scaling of the electronic coordinates:

$$r_i \to \alpha r_i,$$

$$\psi(r_i) \to \psi_\alpha(r_i) = \alpha^{3N/2}\psi(\alpha r_i).$$

This form of the scaling is needed to ensure normalization.

Q: Prove that as defined above, ψ_α is normalized.

The Virial Theorem III

We need two more identities:

- $\langle\psi_\alpha|T|\psi_\alpha\rangle = \alpha^2\langle\psi|T|\psi\rangle$
- $\langle\psi_\alpha|V(R)|\psi_\alpha\rangle = \alpha\langle\psi|V(\alpha R)|\psi\rangle$

The Virial Theorem IV

Proof of the first result:

$$\begin{aligned}
\langle\psi_\alpha|T|\psi_\alpha\rangle &= \alpha^{3N}\int \psi^*(\alpha r_i)\left(-\frac{1}{2}\sum_i \nabla_i^2\right)\psi(\alpha r_i)\prod_i \mathrm{d}r_i \\
&= \int \psi^*(\alpha r_i)\left(-\frac{1}{2}\alpha^2\sum_i \nabla^{\alpha}{}_i^2\right)\psi(\alpha r_i)\prod_i \mathrm{d}(\alpha r_i) \\
&= \alpha^2\int \psi^*(r_i')\left(-\frac{1}{2}\sum_i \nabla'^2_i\right)\psi(r_i')\prod_i \mathrm{d}(r_i') \\
&= \alpha^2\langle\psi|T|\psi\rangle
\end{aligned}$$

The Virial Theorem V

Q: Show that the matrix element of the potential energy operator satisfies the following relation upon uniform scaling of the electronic corrdinates:

$$\langle\psi_\alpha|V(R)|\psi_\alpha\rangle = \alpha\langle\psi|V(\alpha R)|\psi\rangle$$

Hint: If you get stuck have a look at the Exercise 4.3 in the book by Helgaker *et al.*.

The Virial Theorem VI

Now consider

$$\begin{aligned}
\left.\frac{\partial E(\alpha)}{\partial\alpha}\right|_{\alpha=1} &= \left.\frac{\partial}{\partial\alpha}\langle\psi_\alpha|H(R)|\psi_\alpha\rangle\right|_{\alpha=1} \\
&= \left.\frac{\partial}{\partial\alpha}\left(\alpha^2\langle\psi_\alpha|T|\psi_\alpha\rangle + \alpha\langle\psi_\alpha|V(\alpha R)|\psi_\alpha\rangle\right)\right|_{\alpha=1} \\
&= \left.2\alpha\langle\psi|T|\psi\rangle + \langle\psi|V(\alpha R)|\psi\rangle + \alpha\langle\psi|\frac{\partial V(\alpha R)}{\partial\alpha}|\psi\rangle\right|_{\alpha=1}
\end{aligned}$$

Therefore we get the molecular virial theorem:

$$\boxed{2\langle\psi|T|\psi\rangle + \langle\psi|V(R)|\psi\rangle = -\langle\psi|\left.\frac{\partial V(\alpha R)}{\partial\alpha}\right|_{\alpha=1}|\psi\rangle}$$

The Virial Theorem VII

For exact or variational states, using the Hellmann–Feynman theorem we know that

$$\left.\frac{\partial E(\alpha R)}{\partial \alpha}\right|_{\alpha=1} = \langle\psi| \left.\frac{\partial V(\alpha R)}{\partial \alpha}\right|_{\alpha=1} |\psi\rangle$$

so, the virial theorem becomes

$$\boxed{2\langle\psi|T|\psi\rangle + \langle\psi|V(R)|\psi\rangle = -\left.\frac{\partial E(\alpha R)}{\partial \alpha}\right|_{\alpha=1}}$$

The Virial Theorem VIII

The R.H.S. is just the classical force w.r.t. uniform scaling of the nuclear framework.
This force vanishes at molecular equilibrium, at which point the molecular virial theorem becomes the more famaliar:

$$\langle \psi | T | \psi \rangle = -\frac{1}{2} \langle \psi | V(R_e) | \psi \rangle$$

with all quantities evaluated at $R = R_e$.