Lecture 5: Perturbation Theory I

April 1, 2020

To remind you of the current task at hand, we are trying to find good quality approximations to electronic wavefunctions in order to obtain the electronic energy of a molecular system. That is, we seek to solve the Schrödinger equation in the best way possible.

$$H \ket{\Psi} = E \ket{\Psi}$$

We already learned that HF theory is a decent and cheap approximation, and full-CI is exact but very expensive. Perturbation theory is one of many approaches to finding wavefunctions and energies which are of reasonable cost and reasonable accuracy. Recall, however, that the HF wavefunction Φ gives us a very good starting point, so we are really just interested in solving for the correlation energy, which is obtained by shifting our Hamiltonian by the HF energy.

$$H_{\mathrm{c}} \ket{\Psi} = E_{\mathrm{corr}} \ket{\Psi}$$

where

$$H_{\mathrm{c}} = H - \braket{\Phi | H | \Phi}$$

$$E_{\mathrm{corr}} = E_{\mathrm{exact}} - E_{\mathrm{HF}}$$

We will start by discussing Rayleigh-Schrödinger perturbation theory (RSPT), which applies the ideas of PT to the time-independent Schrödinger equation. Next, we will apply the results of RSPT to derive Møller-Plesset methods (MPn) for approximating the electronic correlation energy of molecular systems.

1 Rayleigh-Schrödinger Perturbation Theory

Perturbation theory (PT), in the most general sense, is applicable in nearly every domain of physics. If a problem can be separated into some easily solvable piece A and not-so-easily solvable piece B ("the perturbation"), one can approach the exact solution of the problem in terms of the solution to A plus a series of corrections to account for the influence of perturbation. Rayleigh-Schrödinger Perturbation Theory (RSPT) applies the ideas of PT to the time-independent Schrödinger equation. The Hamiltonian is decomposed into a sum of a simple (easy to solve) problem H_0 and a perturbation V:

$$H = H_0 + V$$

If the perturbation is "small" (i.e., does not wildly shift the solutions of the problem) the energy levels/eigenstates associated with the *perturbed system* described by H can be expressed as sum of the energy levels/eigenstates of the *simple system* (described by H_0) and a series of corrections obtained by considering the influence of the perturbation V on the energy levels/eigenstates of the simple system. Each of these corrections contain eigenvalues of H_0 and matrix elements of the perturbation between the eigenfunctions of H_0, as we will see. In summary, RSPT is applying age-old ideas of general perturbation theory to the time independent Schrödinger equation.

The eigenvalue problem we wish to solve is the following:

$$H \ket{\Psi_i} = (H_0 + V) \ket{\Psi_i} = E_i \ket{\Psi_i}$$

We are assuming H_0 is a simple, easy-to-solve system, for which we have access to all the eigenstates and eigenvalues.

$$H_0 \, |\Phi_i^{(0)}\rangle = E_i^{(0)} \, |\Phi_i^{(0)}\rangle$$

These eigenstates $|\Phi_i^{(0)}\rangle$ and eigenvalues E_i for the simple system are hopefully very close to the exact results $|\Psi_i\rangle$ and E_i, that is, the perturbation V is small. Now, we do something a bit weird. We introduce a parameter λ in front of V, which allows us to toggle the influence of the perturbation on the energy levels and eigenstates. Before saying much else about λ, let's write it down and see what it allows us to do.

$$(H_0 + \lambda V) \, |\Psi_i\rangle = E_i \, |\Psi_i\rangle$$

When $\lambda = 0$, our Hamiltonian is just H_0, and our eigenstates and energy levels become that of the simple system $|\Phi_i^{(0)}\rangle$ and $E_i^{(0)}$. When $\lambda = 1$, our Hamiltonian is the full Hamiltonian H, and our eigenstates and energy levels are the perturbed system $|\Psi_i^{(0)}\rangle$ and E_i. By doing this, *the energy levels and eigenstates of the perturbed system are a function of λ*. To put it another way, our energy levels and eigenstates depend on lambda; we might express this mathematically as $E_i(\lambda)$ and $\Psi_i(\lambda)$. Since these quantities are *functions of λ*, we can write down a Taylor series expansion of the energy levels and eigenstates of H in terms of λ:

$$E_i = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \cdots$$

$$|\Psi_i\rangle = \Phi_i^{(0)} + \lambda \Phi_i^{(1)} + \lambda^2 \Phi_i^{(2)} + \cdots$$

where each term is,

$$E_i^{(k)} = \frac{1}{k!} \frac{d^k E_i}{d\lambda^k} \bigg|_{\lambda=0}$$

$$\Phi_i^{(k)} = \frac{1}{k!} \frac{d^k \, |\Psi_i\rangle}{d\lambda^k} \bigg|_{\lambda=0}$$

We call $E_i^{(0)}$ the 0th order energy, which is the energy of the simple system. Every additional term in the energy expansion is the nth order energy correction, and these are what we must solve for. Likewise terminology is used for the eigenstate corrections $|\Phi_i^{(k)}\rangle$

Many textbooks and other resources simply say λ is a "fictitious dummy parameter" or "book-keeping device" that is used to keep track of the "orders" of the terms when working with the perturbation expansion. This is a bit of an oversimplification. The parameter λ has the utility of turning the perturbation on/off, and motivates the use of a Taylor expansion in the first place. Without λ, you have no Taylor expansion, and no perturbation theory. While it does serve as a useful label for the orders of each term, it's not just a "dummy" parameter. It's the key to all of perturbation theory.

We have a form for the wavefunction and the energy. Plugging these into the time-independent Schrödinger equation, we get

$$(H_0 + \lambda V)(\Phi_i^{(0)} + \lambda \Phi_i^{(1)} + \lambda^2 \Phi_i^{(2)} + \cdots) = (E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \cdots)(\Phi_i^{(0)} + \lambda \Phi_i^{(1)} + \lambda^2 \Phi_i^{(2)} + \cdots)$$

Subtracting from each side so that the left is 0 and grouping like powers of λ gives

$$\left(H_0 \, |\Phi_i^{(0)}\rangle - E_i^{(0)} \, |\Phi_i^{(0)}\rangle \right)$$
$$+ \lambda \left(H_0 \, |\Phi_i^{(1)}\rangle + V \, |\Phi_i^{(0)}\rangle - E_i^{(0)} \, |\Phi_i^{(1)}\rangle - E_i^{(1)} \, |\Phi_i^{(0)}\rangle \right)$$
$$+ \lambda^2 \left(H_0 \, |\Phi_i^{(2)}\rangle + V \, |\Phi_i^{(1)}\rangle - E_i^{(0)} \, |\Phi_i^{(2)}\rangle - E_i^{(1)} \, |\Phi_i^{(1)}\rangle - E_i^{(2)} \, |\Phi_i^{(0)}\rangle \right)$$
$$+ \cdots$$
$$= 0$$

The only way the above equation can be satisfied for any arbitrary value of λ is if the coefficient in parentheses for each power of λ is equal to 0. Therefore, since each term for each power of λ is equal to 0,

this equation is fully separable into an infinite series of simultaneous equations

$$H_0 \ket{\Phi_i^{(0)}} = E_i^{(0)} \ket{\Phi_i^{(0)}}$$
$$H_0 \ket{\Phi_i^{(1)}} + V \ket{\Phi_i^{(0)}} = E_i^{(0)} \ket{\Phi_i^{(1)}} + E_i^{(1)} \ket{\Phi_i^{(0)}}$$
$$H_0 \ket{\Phi_i^{(2)}} + V \ket{\Phi_i^{(1)}} = E_i^{(0)} \ket{\Phi_i^{(2)}} + E_i^{(1)} \ket{\Phi_i^{(1)}} + E_i^{(2)} \ket{\Phi_i^{(0)}}$$
$$\cdots$$

Notice we have now dropped the λ's since each of the above equations hold whether or not λ^n is applied to each term. You could also interpret this as us setting $\lambda = 1$, and fully "turning on" the perturbation.

We will assume henceforth that our reference state is normalized $\braket{\Phi_i^{(0)} | \Phi_i^{(0)}} = 1$ and our exact wavefunction is intermediately normalized $\braket{\Phi_i^{(0)} | \Psi_i} = 1$. This yields,

$$\braket{\Phi_i^{(0)} | \Psi_i} = 1 = \braket{\Phi_i^{(0)} | \Phi_i^{(0)}} + \lambda \braket{\Phi_i^{(0)} | \Phi_i^{(1)}} + \lambda^2 \braket{\Phi_i^{(0)} | \Phi_i^{(2)}} + \cdots$$

In order for the above expression to be true for arbitrary λ, it must be the case that $\braket{\Phi_i^{(0)} | \Phi_i^{(k)}} = 0$ for $k = 1, 2, 3, \ldots$

We can now project each equation in our system of equations on the left by $\bra{\Phi_i^{(0)}}$. We can simplify quite a bit by noting $\braket{\Phi_i^{(0)} | \Phi_i^{(k)}} = 0$ for $k = 1, 2, 3, \ldots$ and that since H_0 is Hermitian, we can act it act on the left to give $\braket{\Phi_i^{(0)} | H_0 | \Phi_n^{(k)}} = E_i^{(0)} \braket{\Phi_i^{(0)} | \Phi_i^{(k)}}$.

The final result is:

$$E_i^{(0)} = \braket{\Phi_i^{(0)} | H_0 | \Phi_i^{(0)}}$$
$$E_i^{(1)} = \braket{\Phi_i^{(0)} | V | \Phi_i^{(0)}}$$
$$E_i^{(2)} = \braket{\Phi_i^{(0)} | V | \Phi_i^{(1)}}$$

Notice that $E_i^{(1)}$ depends on $\Phi_i^{(0)}$, and $E_i^{(2)}$ depends on $\Phi_i^{(1)}$. This trend continues on for higher order energy corrections: the nth order correction to the wavefunction is required for finding the $(n+1)^{th}$ order energy correction.

1.1 The first-order wavefunction and second-order energy corrections

The third equation above, the second-order energy $(E_i^{(2)})$ correction, requires you to know the 1st order correction to the wavefunction $\Phi_i^{(1)}$. To obtain it, we take the equation containing our 1st-order terms and rearrange:

$$H_0 \ket{\Phi_i^{(1)}} + V \ket{\Phi_i^{(0)}} = E_i^{(0)} \ket{\Phi_i^{(1)}} + E_i^{(1)} \ket{\Phi_i^{(0)}}$$

Projecting by $\bra{\Phi_n^{(0)}}$ for some $n \neq i$ gives

$$\braket{\Phi_n^{(0)} | H_0 | \Phi_i^{(1)}} + \braket{\Phi_n^{(0)} | V | \Phi_i^{(0)}} = E_i^{(0)} \braket{\Phi_n^{(0)} | \Phi_i^{(1)}} + E_i^{(1)} \braket{\Phi_n^{(0)} | \Phi_i^{(0)}}$$

To simplify, note that $\bra{\Phi_n^{(0)}} H_0 = E_n^{(0)} \bra{\Phi_n^{(0)}}$, so the first term on the left becomes $E_n^{(0)} \braket{\Phi_n^{(0)} | \Phi_i^{(1)}}$ The second term on the right goes to 0 since $\braket{\Phi_n^{(0)} | \Phi_i^{(0)}} = 0$ when $n \neq i$. The reason is these are both eigenfunctions of H_0, which are orthonormal. We are left with

$$E_n^{(0)} \braket{\Phi_n^{(0)} | \Phi_i^{(1)}} + \braket{\Phi_n^{(0)} | V | \Phi_i^{(0)}} = E_i^{(0)} \braket{\Phi_n^{(0)} | \Phi_i^{(1)}}$$

Rearranging we get

$$\braket{\Phi_n^{(0)} | \Phi_i^{(1)}} = -\frac{\braket{\Phi_n^{(0)} | V | \Phi_i^{(0)}}}{E_n^{(0)} - E_i^{(0)}}$$

We want to get $|\Phi_i^{(1)}\rangle$ by itself on the left side. The left side above is the expansion coefficients of $\Phi_i^{(1)}$ in the basis of $\Phi_n^{(0)}$. To see this, project each side by $|\Phi_n^{(0)}\rangle$ and sum over all $n \neq i$

$$\sum_{n \neq i} |\Phi_n^{(0)}\rangle \langle \Phi_n^{(0)}|\Phi_i^{(1)}\rangle = \sum_{n \neq i} -\frac{\langle \Phi_n^{(0)}|V|\Phi_i^{(0)}\rangle}{E_n^{(0)} - E_i^{(0)}} |\Phi_n^{(0)}\rangle$$

The left side above would be equal to $|\Phi_i^{(1)}\rangle$ if the sum was over all n (resolution of the identity), but if we sum over all n, the energy denominator on the right will be 0 for $n = i$. Fortunately, the expansion coefficient $\langle \Phi_n^{(0)}|\Phi_i^{(1)}\rangle$ for $n = i$ is necessarily 0. Why? Because of intermediate normalization.

$$\langle \Phi_i^{(0)}|\Psi_i\rangle = 1 = \langle \Phi_i^{(0)}|\Phi_i^{(0)}\rangle + \lambda \langle \Phi_i^{(0)}|\Phi_i^{(1)}\rangle + \lambda^2 \langle \Phi_i^{(0)}|\Phi_i^{(2)}\rangle + \cdots$$

Since our reference wavefunction is normalized $\langle \Phi_i^{(0)}|\Phi_i^{(0)}\rangle = 1$, in order for this expression to be true for arbitrary λ, it must be the case that $\langle \Phi_i^{(0)}|\Phi_i^{(k)} = 0 \text{ for } k = 1, 2, 3, ...\rangle$ This implies that when $n = i$, the expansion coefficient is 0. Therefore, we can sum over all n (rather than $n \neq i$) on the left side of our expression for $|\Phi_i^{(1)}\rangle$ without penalty; it is the same as adding zero to both sides:

$$\sum_{n} |\Phi_n^{(0)}\rangle \langle \Phi_n^{(0)}|\Phi_i^{(1)}\rangle = \sum_{n \neq i} -\frac{\langle \Phi_n^{(0)}|V|\Phi_i^{(0)}\rangle}{E_n^{(0)} - E_i^{(0)}} |\Phi_n^{(0)}\rangle$$

Noting the left side contains the resolution of the identity, we can just remove it:

$$|\Phi_i^{(1)}\rangle = \sum_{n \neq i} -\frac{\langle \Phi_n^{(0)}|V|\Phi_i^{(0)}\rangle}{E_n^{(0)} - E_i^{(0)}} |\Phi_n^{(0)}\rangle$$

This now enables us to derive an evaluable expression for the second-order energy

$$E_i^{(2)} = \langle \Phi_i^{(0)}|V|\Phi_i^{(1)}\rangle$$

Plugging in the definition for $\Phi_i^{(1)}$ above we obtain

$$E_i^{(2)} = \sum_{n \neq i} -\frac{\langle \Phi_n^{(0)}|V|\Phi_i^{(0)}\rangle \langle \Phi_i^{(0)}|V|\Phi_n^{(0)}\rangle}{E_n^{(0)} - E_i^{(0)}} |\Phi_n^{(0)}\rangle$$

$$E_i^{(2)} = \sum_{n \neq i} \frac{|\langle \Phi_i^{(0)}|V|\Phi_n^{(0)}\rangle|^2}{E_i^{(0)} - E_n^{(0)}}$$

One could keep working through the tedious algebra and derive the expressions for the second order wavefunction, third order wavefunction, etc, and in turn obtain energy corrections $E_i^{(3)}$, $E_i^{(4)}$, etc, but we will stop here. The above results are sufficient for deriving second-order Møller–Plesset perturbation theory.

There does exist a more advanced scheme which generalizes the results of all orders of perturbation theory using a mysterious new thing called "resolvents". We will touch on this at the end of these notes and continue down that path after learning diagrams.

RSPT, as described above, is a completely general scheme, applicable to any application of the time-independent Schrodinger equation. We will use these ideas to come up with a way to obtain electronic energies with RSPT.

2   Møller–Plesset Perturbation Theory

Møller–Plesset Perturbation Theory (MPn) is the direct application of time-independent RSPT to the problem of finding electronic wavefunctions and energies. In this section, we will derive MP2 by applying the

RSPT equations derived above for a specific partition of our electronic Hamiltonian into H_0 and V. The MP2 energy which is correct up to second-order.

We first must divide our Hamiltonian into two parts, one which has a simple solution H_0 and the perturbation V.

$$H = H_0 + V$$

In MPn methods, the H_0 part of the Hamiltonian is taken to be the Hartree-Fock Hamiltonian,

$$H_0 = F = (h_p^q + \bar{g}_{pi}^{qi})a_q^p = f_p^q a_q^p$$

Note here we are **not** using the Φ-normal ordered Fock operator $f_p^q \tilde{a}_q^p$. The perturbation V will just be everything not included in H_0.

$$V = H - F$$

$$V = \frac{1}{4}\bar{g}_{pq}^{rs}a_{rs}^{pq} - \bar{g}_{pi}^{qi}a_q^p$$

One could interpret V, the perturbation, as being all of the two-electron interactions, except for the mean-field picture already given by the HF Hamiltonian.

Suppose we seek the energy of the ground state of a system (which we will denote by a subscript 0). The ground state 0th order wavefunction $\Phi_0^{(0)}$ is the Hartree-Fock Slater determinant wavefunction. The ground state 0th order energy is given by $E_0^{(0)} = \langle \Phi_0^{(0)} | H_0 | \Phi_0^{(0)} \rangle$

$$\begin{aligned} E_0^{(0)} &= \langle \Phi_0^{(0)} | H_0 | \Phi_0^{(0)} \rangle \\ &= \sum_i h_{ii} + \sum_{ij} \langle ij || ij \rangle \end{aligned}$$

We note the above expression, from Hartree-Fock theory, is equal to the sum of the occupied orbital energies $\sum_i \varepsilon_i$

$$E_0^{(0)} = \sum_i h_{ii} + \sum_{ij} \langle ij || ij \rangle = \sum_i \varepsilon_i$$

The 1st order energy correction is given by $E_0^{(1)} = \langle \Phi_0^{(0)} | V | \Phi_0^{(0)} \rangle$

$$\begin{aligned} E_0^{(1)} &= \langle \Phi_0^{(0)} | V | \Phi_0^{(0)} \rangle \\ &= -\frac{1}{2} \sum_{ij} \langle ij || ij \rangle \end{aligned}$$

We notice immediately that the sum of the 0th and 1st order energies is the Hartree-Fock energy.

$$E_0^{(0)} + E_0^{(1)} = \sum_i h_{ii} + \frac{1}{2} \sum_{ij} \langle ij || ij \rangle$$

Thus, the first two terms of our energy perturbation expansion is the HF energy.

$$E_0 = E_0^{(0)} + E_0^{(1)} + E_0^{(2)} + \cdots$$

$$E_0 = E_{\text{HF}} + E_0^{(2)} + \cdots$$

It must be that all higher-order terms constitute *correlation energies*, since in principle the expansion should approach the exact energy so long as the perturbation is small.

The second-order energy correction to the ground state energy is given by

$$E_0^{(2)} = \sum_{n \neq 0} \frac{|\langle \Phi_0^{(0)} | V | \Phi_n^{(0)} \rangle|^2}{E_0^{(0)} - E_n^{(0)}}$$

What sort of states are $\Phi_n^{(0)}$? They can be any excitation from our ground state reference HF determinant Φ_0 (any determinant in our CI expansion). However, we find that single excitations do not survive:

$$\langle \Phi_0 | V | \Phi_i^a \rangle = \langle \Phi_0 | H - H_0 | \Phi_i^a \rangle = \langle \Phi_0 | H | \Phi_i^a \rangle - \langle \Phi_0 | H_0 | \Phi_i^a \rangle = 0$$

Triple and higher excitations do not survive because you cannot form complete contractions, so $\Phi_n^{(0)}$ can only be the set of unique doubly excited determinants. What about the $E_0^{(0)}$ and $E_n^{(0)}$? Recall $E_0^{(0)}$ is the eigenvalue of H_0 acting on our Hartree-Fock determinant,

$$H_0 | \Phi_0^{(0)} \rangle = E_0^{(0)} | \Phi_0^{(0)} \rangle$$

which is just the sum of the occupied orbital energies $E_0^{(0)} = \sum_k \varepsilon_k$. The same is true for $E_n^{(0)}$, but it has a sum over occupied orbital energies which differs by two spin orbital occupations. Thus,

$$E_0^{(0)} - E_n^{(0)} = \sum_k \varepsilon_k - \left(\sum_k \varepsilon_k - (\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b) \right)$$

$$E_0^{(0)} - E_n^{(0)} = \varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b$$

So we finally obtain for the second order energy expression,

$$E_0^{(2)} = \sum_{n \neq 0} \frac{| \langle \Phi_0^{(0)} | V | \Phi_n^{(0)} \rangle |^2}{E_0^{(0)} - E_n^{(0)}}$$

$$E_0^{(2)} = \sum_{\substack{i < j \\ a < b}} \frac{| \langle \Phi_0 | V | \Phi_{ij}^{ab} \rangle |^2}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}$$

Noting that only the first term of V yields a nonzero result,

$$E_0^{(2)} = \sum_{\substack{i < j \\ a < b}} \frac{| \langle \Phi_0 | \frac{1}{4} \bar{g}_{pq}^{rs} a_{rs}^{pq} | \Phi_{ij}^{ab} \rangle |^2}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}$$

$$E_0^{(2)} = \frac{1}{16} \sum_{\substack{i < j \\ a < b}} \frac{| \langle ij || ab \rangle |^2}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}$$

$$E_0^{(2)} = \frac{1}{4} \sum_{ijab} \frac{| \langle ij || ab \rangle |^2}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}$$