6 Perturbation theory

Definition 6.1. ***Model Hamiltonian.*** The electronic Hamiltonian[1] can be expressed as the sum of a *zeroth order* or *"model" Hamiltonian* H_0 and a *perturbation* V_c, known as the *fluctuation potential*. For well-behaved electronic systems, a common choice for the model Hamiltonian is the diagonal part of the Fock operator.

$$H_0 \equiv f_p^p \tilde{a}_p^p \qquad\qquad V_c \equiv f_p^q (1 - \delta_p^q) \tilde{a}_q^p + \tfrac{1}{4} \overline{g}_{pq}^{rs} \tilde{a}_{rs}^{pq} \tag{6.1}$$

This choice of H_0 brings the advantage that its eigenbasis is the standard basis of determinants.

$$H_0 \Phi = 0\, \Phi \qquad\qquad H_0 \Phi_{i_1 \cdots i_k}^{a_1 \cdots a_k} = \mathcal{E}_{i_1 \cdots i_k}^{a_1 \cdots a_k} \Phi_{i_1 \cdots i_k}^{a_1 \cdots a_k} \qquad\qquad \mathcal{E}_{q_1 \cdots q_k}^{p_1 \cdots p_k} \equiv \sum_{r=1}^{k} f_{p_r}^{p_r} - \sum_{r=1}^{k} f_{q_r}^{q_r} \tag{6.2}$$

In general the model Hamiltonian is chosen to make the matrix representation of H_c in the model eigenbasis diagonally dominant.[2] Our choice of H_0 is appropriate for *weakly correlated systems*, where the reference determinant can be chosen to satisfy $\langle \Phi | \Psi \rangle \gg \langle \Phi_{i_1 \cdots i_k}^{a_1 \cdots a_k} | \Psi \rangle$ for all substituted determinants. In this context it is convenient to employ intermediate normalization for the wavefunction, which will be assumed from here on out.

Definition 6.2. ***Perturbation theory.*** *Perturbation theory* analyzes the polynomial order with which the wavefunction and its observables depend on the fluctuation potential. For this purpose, we define a continuous series of Hamiltonians $H(\lambda) \equiv H_0 + \lambda V_c$ parametrized by a *strength parameter* λ that smoothly toggles between the model Hamiltonian at $\lambda = 0$ to the exact one at $\lambda = 1$. The mth-*order contribution* to a quantity X is then defined as the mth coefficient in its Taylor series about $\lambda = 0$, denoted $X^{(m)}$. In particular, the wavefunction and correlation energy can be expanded as follows.

$$\Psi = \sum_{m=0}^{\infty} \Psi_c^{(m)} \quad E_c = \sum_{n=0}^{\infty} E_c^{(m)} \quad \Psi^{(m)} \equiv \frac{1}{m!} \left. \frac{\partial^m \Psi(\lambda)}{\partial \lambda^m} \right|_{\lambda=0} \quad E_c^{(m)} \equiv \frac{1}{m!} \left. \frac{\partial^m E(\lambda)}{\partial \lambda^m} \right|_{\lambda=0} \quad H(\lambda)\Psi(\lambda) = E(\lambda)\Psi(\lambda) \tag{6.3}$$

The order(s) at which a term contributes to the wavefunction or energy provides one measure of its relative importance.

Remark 6.1. Projecting the Schrödinger equation by Φ and using eq 6.2, along with intermediate normalization, implies

$$E_c = \langle \Phi | V_c | \Psi \rangle \qquad\Longrightarrow\qquad E_c^{(m+1)} = \langle \Phi | V_c | \Psi^{(m)} \rangle \tag{6.4}$$

where the equation on the right follows from generalizing the energy expression to $E(\lambda) = \langle \Phi | \lambda V_c | \Psi(\lambda) \rangle$. In words, this says that the mth-order wavefunction contribution determines the $(m+1)^{\text{th}}$-order energy contribution. This immediately identifies the first-order energy as $E_c^{(1)} = \langle \Phi | V_c | \Phi \rangle = 0$, since V_c consists of Φ-normal-ordered operators.

Definition 6.3. ***Model space projection operator.*** The projection onto the reference determinant, $P = |\Phi\rangle\langle\Phi|$, is termed the *model space projection operator*. Its complement is the *orthogonal space projection operator*.

$$Q \equiv 1 - P = \sum_k \left(\tfrac{1}{k!} \right)^2 \sum_{\substack{a_1 \cdots a_k \\ i_1 \cdots i_k}} |\Phi_{i_1 \cdots i_k}^{a_1 \cdots a_k}\rangle\langle\Phi_{i_1 \cdots i_k}^{a_1 \cdots a_k}| \tag{6.5}$$

Note that P and Q satisfy the following relationships, which are characteristic of complementary projection operators.

$$P + Q = 1 \qquad\qquad P^2 = P \qquad\qquad Q^2 = Q \qquad\qquad PQ = QP = 0 \tag{6.6}$$

Due to intermediate normalization, we also have that $P\Psi = \Phi$ and $Q\Psi = \Psi - \Phi$.

Definition 6.4. ***Resolvent.*** The *resolvent*, $R_0 \equiv (-H_0)^{-1} Q$, is the negative[3] inverse of H_0 in the orthogonal space.

$$R_0 \Phi = 0\Phi \qquad\qquad R_0 \Phi_{i_1 \cdots i_k}^{a_1 \cdots a_k} = (\mathcal{E}_{a_1 \cdots a_k}^{i_1 \cdots i_k})^{-1} \Phi_{i_1 \cdots i_k}^{a_1 \cdots a_k} \qquad\qquad R_0 = \sum_k \left(\tfrac{1}{k!} \right)^2 \sum_{\substack{a_1 \cdots a_k \\ i_1 \cdots i_k}} \frac{|\Phi_{i_1 \cdots i_k}^{a_1 \cdots a_k}\rangle\langle\Phi_{i_1 \cdots i_k}^{a_1 \cdots a_k}|}{\mathcal{E}_{a_1 \cdots a_k}^{i_1 \cdots i_k}} \tag{6.7}$$

The equation on the right is the spectral decomposition of the resolvent.[4] Restriction to the orthogonal space is necessary because H_0 is singular in the model space, which means that H_0^{-1} does not exist there.

---

[1] For the sake of brevity I will here refer to H_c as "the electronic Hamiltonian". We could also use $H_e = E_0 + H_c$, which will simply shift some of the equations by a constant.

[2] See https://en.wikipedia.org/wiki/Diagonally_dominant_matrix.

[3] The annoying sign factor is required for consistency with $R(\zeta) \equiv (\zeta - H_0)^{-1} Q$, which is a more general definition of the resolvent.

[4] This follows from the eigenvalue equations, but you can derive it explicitly by substituting equation 6.5 into $R_0 = (-H_0)^{-1} Q$.

Remark 6.2. *A recursive solution to the Schrödinger equation.* Operating R_0 on $H(\lambda)\Psi(\lambda) = E(\lambda)\Psi(\lambda)$ gives[5]

$$\Psi(\lambda) = \Phi + R_0(\lambda V_{\rm c} - E(\lambda))\Psi(\lambda) \tag{6.8}$$

which provides a recursive equation for $\Psi(\lambda)$ that can be used solve for wavefunction contributions order by order.

Example 6.1. The first two derivatives of equation 6.8 are given by

$$\frac{\partial\Psi(\lambda)}{\partial\lambda} = R_0\left(V_{\rm c} - \frac{\partial E(\lambda)}{\partial\lambda}\right)\Psi(\lambda) + R_0(\lambda V_{\rm c} - E(\lambda))\frac{\partial\Psi(\lambda)}{\partial\lambda}$$

$$\frac{\partial^2\Psi(\lambda)}{\partial\lambda^2} = -R_0\frac{\partial^2 E(\lambda)}{\partial\lambda^2}\Psi(\lambda) + 2R_0\left(V_{\rm c} - \frac{\partial E(\lambda)}{\partial\lambda}\right)\frac{\partial\Psi(\lambda)}{\partial\lambda} + R_0(\lambda V_{\rm c} - E(\lambda))\frac{\partial^2\Psi(\lambda)}{\partial\lambda^2}$$

which can be used to determine the first- and second-order wavefunction contributions.

$$\Psi^{(1)} = \left.\frac{\partial\Psi(\lambda)}{\partial\lambda}\right|_{\lambda=0} = R_0 V_{\rm c}\Phi \qquad\qquad \Psi^{(2)} = \left.\frac{1}{2}\frac{\partial^2\Psi(\lambda)}{\partial\lambda^2}\right|_{\lambda=0} = R_0 V_{\rm c}\Psi^{(1)} = R_0 V_{\rm c} R_0 V_{\rm c}\Phi$$

Here we have used $E_{\rm c}^{(0)} = E_{\rm c}^{(1)} = 0$ and $R_0\Phi = 0$ to simplify the result.

Example 6.2. Expanding the resolvent according to eq (6.7) allows us to expand $\Psi^{(1)}$ in the determinant basis.

$$\Psi^{(1)} = R_0 V_{\rm c}\Phi = \sum_{\substack{a \\ i}} \Phi_i^a \frac{\langle\Phi_i^a|V_{\rm c}|\Phi\rangle}{\mathcal{E}_a^i} + \left(\frac{1}{2!}\right)^2 \sum_{\substack{ab \\ ij}} \Phi_{ij}^{ab}\frac{\langle\Phi_{ij}^{ab}|V_{\rm c}|\Phi\rangle}{\mathcal{E}_{ab}^{ij}} \tag{6.9}$$

The expansion truncates at double excitations because the maximum excitation level of $V_{\rm c}$ is $+2$.

Example 6.3. The numerators in example 6.2 are easily evaluated using Slater's rules, which leads to the following.

$$\Psi^{(1)} = \sum_{\substack{a \\ i}} \Phi_i^a \frac{f_a^i}{\mathcal{E}_a^i} + \left(\frac{1}{2!}\right)^2 \sum_{\substack{ab \\ ij}} \Phi_{ij}^{ab}\frac{\bar{g}_{ab}^{ij}}{\mathcal{E}_{ab}^{ij}} \qquad\Longrightarrow\qquad E_{\rm c}^{(2)} = \langle\Phi|V_{\rm c}|\Psi^{(1)}\rangle = \sum_{\substack{a \\ i}} \frac{f_i^a f_a^i}{\mathcal{E}_a^i} + \left(\frac{1}{2!}\right)^2 \sum_{\substack{ab \\ ij}} \frac{\bar{g}_{ij}^{ab}\bar{g}_{ab}^{ij}}{\mathcal{E}_{ab}^{ij}}$$

Note that the singles contribution vanishes for canonical Hartree-Fock references, since $f_a^i = 0$. These extra terms are required for non-canonical orbitals, such as those obtained from restricted open-shell Hartree-Fock (ROHF) theory.

Example 6.4. Diagrammatically, $\Psi^{(1)}$ and $E_{\rm c}^{(2)}$ can be expressed as follows.

$$\Psi^{(1)} = \quad\text{}\quad + \quad\text{}\qquad\qquad E_{\rm c}^{(2)} = \quad\text{}\quad + \quad\text{} \tag{6.10}$$

Remark 6.3. Notice that a fully contracted operator of the form $\sum_i v_i\, a_i|\Phi\rangle\langle\Phi|a_i^\dagger$ can be simplified as

$$\sum_i v_i\, \overline{a_p^\dagger a_i}|\Phi\rangle\langle\Phi|\overline{a_i^\dagger a_q} = v_p\, \overline{a_p^\dagger a_q}|\Phi\rangle\langle\Phi|$$

where v_p is required to have an occupied index by the hole contraction. A similar argument can be made for completely contracted operators of the form $\sum_a v_a\, a_a^\dagger|\Phi\rangle\langle\Phi|a_a$ or in general $\sum_{\substack{abc\cdots \\ ijk\cdots}} v_{abc\cdots}^{ijk\cdots} a_{abc\cdots}^{ijk\cdots}|\Phi\rangle\langle\Phi|a_{ijk\cdots}^{abc\cdots}$, which leads to[6]

$$\langle\Phi|X R_0 X_1 R_0 X_2 \cdots R_0 X_n|\Phi\rangle = \vdots\overline{X\,\vdots\,X_1\,\vdots\,X_2\cdots\,\vdots\,X_n}\vdots \tag{6.11}$$

for a general series of excitation operators $X, X_1, \ldots, X_n$. This makes a convenient corollary to Wick's theorem for operator products with resolvents.

Example 6.5. Expanding the resolvent in $\Psi^{(2)} = R_0 V_{\rm c}\Psi^{(1)}$ gives an expansion that includes up to quadruples

$$\Psi^{(2)} = \sum_{\substack{a \\ i}} \Phi_i^a \frac{\langle\Phi_i^a|V_{\rm c}|\Psi^{(1)}\rangle}{\mathcal{E}_a^i} + \left(\frac{1}{2!}\right)^2 \sum_{\substack{ab \\ ij}} \Phi_{ij}^{ab}\frac{\langle\Phi_{ij}^{ab}|V_{\rm c}|\Psi^{(1)}\rangle}{\mathcal{E}_{ab}^{ij}} + \left(\frac{1}{3!}\right)^2 \sum_{\substack{abc \\ ijk}} \Phi_{ijk}^{abc}\frac{\langle\Phi_{ijk}^{abc}|V_{\rm c}|\Psi^{(1)}\rangle}{\mathcal{E}_{abc}^{ijk}} + \left(\frac{1}{4!}\right)^2 \sum_{\substack{abcd \\ ijkl}} \Phi_{ijkl}^{abcd}\frac{\langle\Phi_{ijkl}^{abcd}|V_{\rm c}|\Psi^{(1)}\rangle}{\mathcal{E}_{abcd}^{ijkl}}$$

since the two-electron part of $\Psi^{(1)}$ has an excitation level of $+2$ and the excitation level of $V_{\rm c}$ ranges from -2 to $+2$.

---

[5]This follows from $R_0 H_0\Psi = -Q\Psi = -\Psi + \Phi$.

[6]The dotted lines are resolvent lines.

Example 6.6. The expansion for $\Psi^{(2)}$ can be evaluated graphically. Assuming Brillouin's theorem for simplicity,

$$\Psi^{(2)} = \text{[diagrams]} + \text{[diagrams]} + \text{[diagrams]} + \text{[diagrams]} + \text{[diagrams]} \tag{6.12}$$

$$+ \text{[diagrams]} + \text{[diagrams]} + \text{[diagrams]}$$

$$= \frac{1}{2}\sum_{\substack{abc \\ ij}} \Phi_i^a \frac{\overline{g}_{aj}^{bc}\overline{g}_{bc}^{ij}}{\mathcal{E}_a^i\mathcal{E}_{bc}^{ij}} + \frac{1}{2}\sum_{\substack{ab \\ ijk}} \Phi_i^a \frac{\overline{g}_{jk}^{ib}\overline{g}_{ab}^{jk}}{\mathcal{E}_a^i\mathcal{E}_{ab}^{jk}} + \frac{1}{2^3}\sum_{\substack{abcd \\ ij}} \Phi_{ij}^{ab} \frac{\overline{g}_{ab}^{cd}\overline{g}_{cd}^{ij}}{\mathcal{E}_{ab}^{ij}\mathcal{E}_{cd}^{ij}} + \frac{1}{2^3}\sum_{\substack{ab \\ ijkl}} \Phi_{ij}^{ab} \frac{\overline{g}_{kl}^{ij}\overline{g}_{ab}^{kl}}{\mathcal{E}_{ab}^{ij}\mathcal{E}_{ab}^{kl}} + \sum_{\substack{abc \\ ijk}} \Phi_{ij}^{ab} \frac{\overline{g}_{ac}^{ik}\overline{g}_{kb}^{cj}}{\mathcal{E}_{ab}^{ij}\mathcal{E}_{ac}^{ik}}$$

$$+ \frac{1}{2^2}\sum_{\substack{abcd \\ ijk}} \Phi_{ijk}^{abc} \frac{\overline{g}_{ad}^{ij}\overline{g}_{bc}^{dk}}{\mathcal{E}_{abc}^{ijk}\mathcal{E}_{ad}^{ij}} - \frac{1}{2^2}\sum_{\substack{abc \\ ijkl}} \frac{\overline{g}_{ab}^{il}\overline{g}_{lc}^{jk}}{\mathcal{E}_{abc}^{ijk}\mathcal{E}_{ab}^{il}} + \frac{1}{2^4}\sum_{\substack{abcd \\ ijkl}} \Phi_{ijkl}^{abcd} \frac{\overline{g}_{ab}^{ij}\overline{g}_{cd}^{kl}}{\mathcal{E}_{abcd}^{ijkl}\mathcal{E}_{ab}^{ij}} \tag{6.13}$$

where the operators in the final diagram do not form an equivalent pair because they pass through different resolvent lines. This can be used to determine the third-order correlation energy

$$E_{\mathrm{c}}^{(3)} = \langle\Phi|V_{\mathrm{c}}|\Psi^{(2)}\rangle = \text{[diagrams]} + \text{[diagrams]} + \text{[diagrams]} = \frac{1}{2^3}\sum_{\substack{abcd \\ ij}} \frac{\overline{g}_{ij}^{ab}\overline{g}_{ab}^{cd}\overline{g}_{cd}^{ij}}{\mathcal{E}_{ab}^{ij}\mathcal{E}_{cd}^{ij}} + \frac{1}{2^3}\sum_{\substack{ab \\ ijkl}} \frac{\overline{g}_{ij}^{ab}\overline{g}_{kl}^{ij}\overline{g}_{ab}^{kl}}{\mathcal{E}_{ab}^{ij}\mathcal{E}_{ab}^{kl}} + \sum_{\substack{abc \\ ijk}} \frac{\overline{g}_{ij}^{ab}\overline{g}_{ac}^{ik}\overline{g}_{kb}^{cj}}{\mathcal{E}_{ab}^{ij}\mathcal{E}_{ac}^{ik}} \tag{6.14}$$

which also equals $\langle\Psi^{(1)}|V_{\mathrm{c}}|\Psi^{(1)}\rangle$. This is an example of the *Wigner (2n+1) rule*, which says that $E_{\mathrm{c}}^{(2n+1)} = \langle\Psi^{(n)}|V_{\mathrm{c}}|\Psi^{(n)}\rangle$. Note that $E_{\mathrm{c}}^{(m+1)}$ always only depends on the doubles contribution to $\Psi^{(m)}$, but that the doubles coefficients themselves may involve triples, quadruples and higher contributions from wavefunction components of order less than m.

Example 6.7. Using $^{(m)}c_{ab\cdots}^{ij\cdots} = \langle\Phi_{ij\cdots}^{ab\cdots}|\Psi^{(m)}\rangle$, the second order CI coefficients can be determined from eq 6.12 by contracting a bare excitation operator with the top of each diagram. Interpreting these graphs gives the following.

$$^{(2)}c_a^i = \frac{1}{2}\sum_{\substack{bc \\ j}} \frac{\overline{g}_{aj}^{bc}\overline{g}_{bc}^{ij}}{\mathcal{E}_a^i\mathcal{E}_{bc}^{ij}} + \frac{1}{2}\sum_{\substack{b \\ jk}} \frac{\overline{g}_{jk}^{ib}\overline{g}_{ab}^{jk}}{\mathcal{E}_a^i\mathcal{E}_{ab}^{jk}}$$

$$^{(2)}c_{ab}^{ij} = \frac{1}{2}\sum_{cd} \frac{\overline{g}_{ab}^{cd}\overline{g}_{cd}^{ij}}{\mathcal{E}_{ab}^{ij}\mathcal{E}_{cd}^{ij}} + \frac{1}{2}\sum_{kl} \frac{\overline{g}_{kl}^{ij}\overline{g}_{ab}^{kl}}{\mathcal{E}_{ab}^{ij}\mathcal{E}_{ab}^{kl}} + \hat{P}_{(a/b)}^{(i/j)}\sum_{\substack{c \\ k}} \frac{\overline{g}_{ac}^{ik}\overline{g}_{kb}^{cj}}{\mathcal{E}_{ab}^{ij}\mathcal{E}_{ac}^{ik}}$$

$$^{(2)}c_{abc}^{ijk} = \hat{P}_{(a/bc)}^{(ij/k)}\sum_d \Phi_{ijk}^{abc} \frac{\overline{g}_{ad}^{ij}\overline{g}_{bc}^{dk}}{\mathcal{E}_{abc}^{ijk}\mathcal{E}_{ad}^{ij}} - \hat{P}_{(a/bc)}^{(i/jk)}\sum_l \frac{\overline{g}_{ab}^{il}\overline{g}_{lc}^{jk}}{\mathcal{E}_{abc}^{ijk}\mathcal{E}_{ab}^{il}}$$

$$^{(2)}c_{abcd}^{ijkl} = \hat{P}_{(ab/cd)}^{(ij/kl)} \frac{\overline{g}_{ab}^{ij}\overline{g}_{cd}^{kl}}{\mathcal{E}_{abcd}^{ijkl}\mathcal{E}_{ab}^{ij}}$$

Note that the second order quadruples coefficient is disconnected. Prop. 6.1 shows that the second-order quadruples operator is actually a simple product of first-order doubles operators. This fact was an early motivation for coupled-pair many-electron theory,[7] since it justifies approximating $\Psi_{\mathrm{CIDQ}} = (1 + C_2 + C_4)\Phi$ by $\Psi_{\mathrm{CPMET}} = (1 + C_2 + \frac{1}{2}C_2^2)\Phi$.

Proposition 6.1. $^{(2)}C_4 = \frac{1}{2}{}^{(1)}C_2^2$
   Proof: This follows from rearranging the resolvent denominator.

$$\frac{1}{\mathcal{E}_{abcd}^{ijkl}\mathcal{E}_{ab}^{ij}} + \frac{1}{\mathcal{E}_{abcd}^{ijkl}\mathcal{E}_{cd}^{kl}} = \frac{\mathcal{E}_{cd}^{kl} + \mathcal{E}_{ab}^{ij}}{\mathcal{E}_{abcd}^{ijkl}\mathcal{E}_{ab}^{ij}\mathcal{E}_{cd}^{kl}} = \frac{1}{\mathcal{E}_{ab}^{ij}\mathcal{E}_{cd}^{kl}} \implies {}^{(2)}C_4 = \left(\frac{1}{2}\right)^4 \sum_{\substack{abcd \\ ijkl}} \tilde{a}_{abcd}^{ijkl} \frac{\overline{g}_{ab}^{ij}\overline{g}_{cd}^{kl}}{\mathcal{E}_{abcd}^{ijkl}\mathcal{E}_{ab}^{ij}} = \frac{1}{2}\cdot\left(\frac{1}{2}\right)^4 \sum_{\substack{abcd \\ ijkl}} \tilde{a}_{abcd}^{ijkl} \frac{\overline{g}_{ab}^{ij}\overline{g}_{cd}^{kl}}{\mathcal{E}_{ab}^{ij}\mathcal{E}_{cd}^{kl}} = \frac{1}{2}{}^{(1)}C_2^2$$

---

[7]This is the original name for coupled-cluster doubles.