

6 Perturbation theory

Definition 6.1. Model Hamiltonian. The electronic Hamiltonian¹ 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.

$$H_0 \equiv f_p^p \tilde{a}_p^p \quad V_c \equiv f_p^q (1 - \delta_p^q) \tilde{a}_q^p + \frac{1}{4} \bar{g}_{pq}^{rs} \tilde{a}_{rs}^{pq} \quad (6.1)$$

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

$$H_0 \Phi = 0 \Phi \quad H_0 \Phi_{i_1 \dots i_k}^{a_1 \dots a_k} = \mathcal{E}_{i_1 \dots i_k}^{a_1 \dots a_k} \Phi_{i_1 \dots i_k}^{a_1 \dots a_k} \quad \mathcal{E}_{q_1 \dots q_k}^{p_1 \dots p_k} \equiv \sum_{r=1}^k f_{p_r}^{q_r} - \sum_{r=1}^k f_{q_r}^{p_r} \quad (6.2)$$

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 $\langle \Phi | \Psi \rangle \gg \langle \Phi_{i_1 \dots i_k}^{a_1 \dots a_k} | \Psi \rangle$ 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 m^{th} -order contribution to a quantity X is then defined as the m^{th} coefficient in its Taylor series about $\lambda = 0$, denoted $X^{(m)}$. In particular, the wavefunction and correlation energy can be expanded as follows.

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

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

$$E_c = \langle \Phi | V_c | \Psi \rangle \quad \implies \quad E_c^{(m+1)} = \langle \Phi | V_c | \Psi^{(m)} \rangle \quad (6.4)$$

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 m^{th} -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*.³

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

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

$$P + Q = 1 \quad P^2 = P \quad Q^2 = Q \quad PQ = QP = 0 \quad (6.6)$$

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 negative⁴ inverse of H_0 in the orthogonal space.

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

The equation on the right is the spectral decomposition of the resolvent.⁵ 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.

¹For 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.

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

³Note that these definitions are implicitly restricted to \mathcal{F}_n rather than the full Fock space.

⁴The 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.

⁵This 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 $H(\lambda)\Psi(\lambda) = E(\lambda)\Psi(\lambda)$ gives⁶

$$\Psi(\lambda) = \Phi + R_0(\lambda V_c - E(\lambda))\Psi(\lambda) \quad (6.8)$$

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

$$\begin{aligned} \frac{\partial \Psi(\lambda)}{\partial \lambda} &= R_0 \left(V_c - \frac{\partial E(\lambda)}{\partial \lambda} \right) \Psi(\lambda) + R_0(\lambda V_c - E(\lambda)) \frac{\partial \Psi(\lambda)}{\partial \lambda} \\ \frac{\partial^2 \Psi(\lambda)}{\partial \lambda^2} &= -R_0 \frac{\partial^2 E(\lambda)}{\partial \lambda^2} \Psi(\lambda) + 2R_0 \left(V_c - \frac{\partial E(\lambda)}{\partial \lambda} \right) \frac{\partial \Psi(\lambda)}{\partial \lambda} + R_0(\lambda V_c - E(\lambda)) \frac{\partial^2 \Psi(\lambda)}{\partial \lambda^2} \end{aligned}$$

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

$$\Psi^{(1)} = \left. \frac{\partial \Psi(\lambda)}{\partial \lambda} \right|_{\lambda=0} = R_0 V_c \Phi \quad \Psi^{(2)} = \left. \frac{1}{2} \frac{\partial^2 \Psi(\lambda)}{\partial \lambda^2} \right|_{\lambda=0} = R_0 V_c \Psi^{(1)} = R_0 V_c R_0 V_c \Phi$$

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

Example 6.2. Plugging in the spectral decomposition for R_0 allows us to expand $\Psi^{(1)}$ in the determinant basis.

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

The expansion truncates at double excitations because the maximum excitation level of V_c is +2.

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

$$\Psi^{(1)} = \sum_i \Phi_i^a \frac{f_a^i}{\mathcal{E}_a^i} + \left(\frac{1}{2!} \right)^2 \sum_{ij} \Phi_{ij}^{ab} \frac{\bar{g}_{ab}^{ij}}{\mathcal{E}_{ab}^{ij}} \quad \Rightarrow \quad E_c^{(2)} = \langle \Phi | V_c | \Psi^{(1)} \rangle = \sum_i \frac{f_i^a f_a^i}{\mathcal{E}_a^i} + \left(\frac{1}{2!} \right)^2 \sum_{ij} \frac{\bar{g}_{ij}^{ab} \bar{g}_{ab}^{ij}}{\mathcal{E}_{ab}^{ij}}$$

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.

Definition 6.5. Resolvent line (extended definition). Let Y be a particle-number conserving operator string⁷ and compare $R_0 Y |\Phi\rangle$ with $Y |\Phi\rangle$ by expanding this operator string using Wick's theorem. We can turn the second expansion into the first by setting each term to zero unless the contracted Y has the form $\tilde{a}_{j_1 \dots j_m}^{b_1 \dots b_m} \times \text{const}$,⁸ in which case we scale it by $(\mathcal{E}_{b_1 \dots b_m}^{j_1 \dots j_m})^{-1}$.⁹ Multiplying these expansions by Y' and applying Wick's theorem gives the following

$$Y' Y |\Phi\rangle = (\text{; } Y' Y \text{; } + \text{; } \overline{Y' Y} \text{; }) |\Phi\rangle \quad Y' R_0 Y |\Phi\rangle = (\text{; } Y' \text{; } Y \text{; } + \text{; } \overline{Y' Y} \text{; }) |\Phi\rangle$$

where the dotted lines indicate that wherever the internal contractions on Y leave a remainder of the form $\tilde{a}_{j_1 \dots j_m}^{b_1 \dots b_m} \times \text{const}$ we scale the term by $(\mathcal{E}_{b_1 \dots b_m}^{j_1 \dots j_m})^{-1}$. All other terms are omitted. This builds upon our earlier concept of a *resolvent line*, which was limited to a particular case: $\langle \Phi | Y' R_0 Y |\Phi\rangle = \text{; } \overline{Y' Y} \text{; }$ with $Y' = \tilde{a}_{a_1 \dots a_k}^{i_1 \dots i_k}$ and $Y = \tilde{a}_{j_1 \dots j_k}^{b_1 \dots b_k}$. The extended definition implies $\text{; } |\Phi_{i_1 \dots i_k}^{a_1 \dots a_k}\rangle = (\mathcal{E}_{i_1 \dots i_k}^{a_1 \dots a_k})^{-1} |\Phi_{i_1 \dots i_k}^{a_1 \dots a_k}\rangle$ and $\text{; } |\Phi\rangle = 0$, which makes an uncontracted resolvent line interchangeable with R_0 .¹⁰ This definition of has a simple graphical interpretation: A graph with hole lines i_1, \dots, i_h and particle lines a_1, \dots, a_k passing through a resolvent line is scaled by $(\mathcal{E}_{a_1 \dots a_k}^{i_1 \dots i_k})^{-1}$ when $h = k$ and vanishes when $h \neq k$. A graph with no operators separating a resolvent line from the bra or ket of the reference state equals zero.

Corollary 6.1. Wick's theorem for perturbation theory. $Y R_0 Y_1 \dots R_0 Y_m |\Phi\rangle = (\text{; } Y \text{; } Y_1 \dots Y_m \text{; } + \text{; } \overline{Y Y_1 \dots Y_m} \text{; }) |\Phi\rangle$

Proof: This follows directly from Wick's theorem and the definition of a resolvent line.

⁶This follows from $R_0 H_0 \Psi = -Q \Psi = -\Psi + \Phi$.

⁷That is, it contains a balanced number of creation and annihilation operators so that $Y |\Phi\rangle \in \mathcal{F}_n$.

⁸Since $R_0 X |\Phi\rangle = 0$ unless $\text{; } X \text{; }$ can fully contract one of the bras $\langle \Phi_{i_1 \dots i_k}^{a_1 \dots a_k} |$ in the resolvent.

⁹Since $R_0 \tilde{a}_{j_1 \dots j_m}^{b_1 \dots b_m} |\Phi\rangle = (\mathcal{E}_{b_1 \dots b_m}^{j_1 \dots j_m})^{-1} \tilde{a}_{j_1 \dots j_m}^{b_1 \dots b_m} |\Phi\rangle$.

¹⁰The difference between resolvents and resolvent lines is that we haven't defined a "contraction through R_0 " operation.

Example 6.4. In diagram notation, $\Psi^{(1)}$ and $E_c^{(2)}$ can be expressed as follows.

$$\Psi^{(1)} = \text{diagram 1} + \text{diagram 2} \quad E_c^{(2)} = \text{diagram 3} + \text{diagram 4} \quad (6.10)$$

Example 6.5. Using the spectral decomposition of R_0 in $\Psi^{(2)} = R_0 V_c \Psi^{(1)}$ gives an expansion that truncates at quadruples

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

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

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

$$\Psi^{(2)} = \text{diagram 1} + \text{diagram 2} + \text{diagram 3} + \text{diagram 4} + \text{diagram 5} + \text{diagram 6} + \text{diagram 7} + \text{diagram 8} + \text{diagram 9} \quad (6.11)$$

$$\begin{aligned} &= \frac{1}{2} \sum_{abc} \Phi_i^a \frac{\bar{g}_{aj}^{bc} \bar{g}_{bc}^{ij}}{\mathcal{E}_a^i \mathcal{E}_{bc}^{ij}} + \frac{1}{2} \sum_{ab} \Phi_i^a \frac{\bar{g}_{jk}^{ib} \bar{g}_{ab}^{jk}}{\mathcal{E}_a^i \mathcal{E}_{ab}^{jk}} + \frac{1}{2^3} \sum_{abcd} \Phi_{ij}^{ab} \frac{\bar{g}_{ab}^{cd} \bar{g}_{cd}^{ij}}{\mathcal{E}_{ab}^{ij} \mathcal{E}_{cd}^{ij}} + \frac{1}{2^3} \sum_{ab} \Phi_{ij}^{ab} \frac{\bar{g}_{kl}^{ij} \bar{g}_{ab}^{kl}}{\mathcal{E}_{ab}^{ij} \mathcal{E}_{kl}^{ij}} + \sum_{ijk} \Phi_{ij}^{ab} \frac{\bar{g}_{ac}^{ik} \bar{g}_{kb}^{cj}}{\mathcal{E}_{ab}^{ij} \mathcal{E}_{ac}^{ik}} \\ &+ \frac{1}{2^2} \sum_{abcd} \Phi_{ijk}^{abc} \frac{\bar{g}_{ad}^{ij} \bar{g}_{bc}^{dk}}{\mathcal{E}_{abc}^{ijk} \mathcal{E}_{ad}^{ij}} - \frac{1}{2^2} \sum_{abc} \frac{\bar{g}_{il}^{ij} \bar{g}_{lc}^{jk}}{\mathcal{E}_{abc}^{ijk} \mathcal{E}_{ab}^{ij}} + \frac{1}{2^4} \sum_{abcd} \Phi_{ijkl}^{abcd} \frac{\bar{g}_{ab}^{ij} \bar{g}_{cd}^{kl}}{\mathcal{E}_{abcd}^{ijkl} \mathcal{E}_{ab}^{ij}} \end{aligned} \quad (6.12)$$

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

$$E_c^{(3)} = \langle \Phi | V_c | \Psi^{(2)} \rangle = \text{diagram 1} + \text{diagram 2} + \text{diagram 3} = \frac{1}{2^3} \sum_{abcd} \frac{\bar{g}_{ij}^{ab} \bar{g}_{ab}^{cd} \bar{g}_{cd}^{ij}}{\mathcal{E}_{ab}^{ij} \mathcal{E}_{cd}^{ij}} + \frac{1}{2^3} \sum_{ab} \frac{\bar{g}_{ij}^{ab} \bar{g}_{kl}^{ij} \bar{g}_{ab}^{kl}}{\mathcal{E}_{ab}^{ij} \mathcal{E}_{kl}^{ij}} + \sum_{abc} \frac{\bar{g}_{ij}^{ab} \bar{g}_{ac}^{ik} \bar{g}_{kb}^{cj}}{\mathcal{E}_{ab}^{ij} \mathcal{E}_{ac}^{ik}} \quad (6.13)$$

which also equals $\langle \Psi^{(1)} | V_c | \Psi^{(1)} \rangle$. This is an example of the *Wigner (2n+1) rule*, which says that $E_c^{(2n+1)} = \langle \Psi^{(n)} | V_c | \Psi^{(n)} \rangle$. Note that $E_c^{(m+1)}$ 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 ${}^{(m)}c_{ab\dots}^{ij\dots} = \langle \Phi_{ij\dots}^{ab\dots} | \Psi^{(m)} \rangle$, the second order CI coefficients can be determined from eq 6.11 by contracting a bare excitation operator with the top of each diagram. Interpreting these graphs gives the following.

$$\begin{aligned} {}^{(2)}c_a^i &= \frac{1}{2} \sum_{bc} \frac{\bar{g}_{aj}^{bc} \bar{g}_{bc}^{ij}}{\mathcal{E}_a^i \mathcal{E}_{bc}^{ij}} + \frac{1}{2} \sum_{jk} \frac{\bar{g}_{jk}^{ib} \bar{g}_{ab}^{jk}}{\mathcal{E}_a^i \mathcal{E}_{ab}^{jk}} \\ {}^{(2)}c_{ab}^{ij} &= \frac{1}{2} \sum_{cd} \frac{\bar{g}_{ab}^{cd} \bar{g}_{cd}^{ij}}{\mathcal{E}_{ab}^{ij} \mathcal{E}_{cd}^{ij}} + \frac{1}{2} \sum_{kl} \frac{\bar{g}_{kl}^{ij} \bar{g}_{ab}^{kl}}{\mathcal{E}_{ab}^{ij} \mathcal{E}_{kl}^{ij}} + \hat{P}_{(a/b)} \sum_c \frac{\bar{g}_{ac}^{ik} \bar{g}_{kb}^{cj}}{\mathcal{E}_{ab}^{ij} \mathcal{E}_{ac}^{ik}} \\ {}^{(2)}c_{abc}^{ijk} &= \hat{P}_{(a/bc)} \sum_d \Phi_{ijk}^{abc} \frac{\bar{g}_{ad}^{ij} \bar{g}_{bc}^{dk}}{\mathcal{E}_{abc}^{ijk} \mathcal{E}_{ad}^{ij}} - \hat{P}_{(a/bc)} \sum_l \frac{\bar{g}_{ab}^{ij} \bar{g}_{lc}^{jk}}{\mathcal{E}_{abc}^{ijk} \mathcal{E}_{ab}^{ij}} \\ {}^{(2)}c_{abcd}^{ijkl} &= \hat{P}_{(ab/cd)} \frac{\bar{g}_{ab}^{ij} \bar{g}_{cd}^{kl}}{\mathcal{E}_{abcd}^{ijkl} \mathcal{E}_{ab}^{ij}} \end{aligned}$$

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,¹¹ since it justifies approximating $\Psi_{\text{CIDQ}} = (1 + C_2 + C_4)\Phi$ by $\Psi_{\text{CPMET}} = (1 + C_2 + \frac{1}{2}C_2^2)\Phi$.

¹¹This is the original name for coupled-cluster doubles.

Proposition 6.1. $^{(2)}C_4 = \frac{1}{2} {}^{(1)}C_2^2$

Proof: This follows from rearranging the resolvent denominator.

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