

7 Perturbative analysis

Definition 7.1. Correct to order p . Define $X^{(p-)} \equiv X^{(0)} + X^{(1)} + \dots + X^{(p)}$ as well as $X^{(p+)} \equiv X^{(p)} + X^{(p+1)} + \dots + X^{(\infty)}$. Then we say an approximation to X is *correct to order p* if it contains all of the contributions in $X^{(p-)}$.

Definition 7.2. Truncated CI and CC. Let CISDTQPH78 $\dots m$ denote truncation of the CI Ansatz at m -tuples. Similarly, CCS $\dots m$ means that we omit cluster operators of excitation level $m \leq k \leq n$. Note that, unlike C_k , the cluster operator T_k in truncated CC describes only *connected k -tuples*, which are residual correlations that cannot be decomposed into products of smaller clusters. As a result, truncated CC actually contains much higher excitations than CI.

Notation 7.1. Let Φ_k be a row vector containing all unique k -fold substitutions of Φ ,¹ so that $\Phi = [\Phi \ \Phi_1 \ \dots \ \Phi_n]$ spans \mathcal{F}_n . Using \mathbf{c} and \mathbf{t} to denote column vectors of CI coefficients and CC amplitudes leads to the following relationships.²

$$\Phi \cdot \mathbf{c} = (C_0 + C_1 + \dots + C_n)\Phi \quad \Phi \cdot \mathbf{t} = (1 + T_1 + \dots + T_n)\Phi \quad (7.1)$$

In Dirac notation the bra $\langle \Phi |$ is transposed, so that $\langle \Phi | W | \Phi \rangle = [\langle \Phi_\sigma | W | \Phi_\tau \rangle]$ is the matrix representation of W in \mathcal{F}_n . Such matrix representations will be denoted with bolded letters, $\mathbf{W} \equiv \langle \Phi | W | \Phi \rangle$.

Remark 7.1. Perturbative analysis of CI. Writing the CI eigenvalue equation $\mathbf{H}_c \mathbf{c} = E_c \mathbf{c}$ in terms of model-Hamiltonian and fluctuation-potential matrices and rearranging yields a new matrix equation

$$(-\mathbf{H}_0 + E_c) \mathbf{c} = \mathbf{V}_c \mathbf{c} \quad (7.2)$$

which provides a convenient starting point for a perturbative analysis and for comparison to the coupled-cluster equations. The matrix elements of the model Hamiltonian are given by $\langle \Phi_\sigma | H_0 | \Phi_\tau \rangle = \mathcal{E}_\sigma \delta_{\sigma\tau}$, so the matrix on the left is diagonal with eigenvalues $-\mathcal{E}_{i_1 \dots i_k}^{a_1 \dots a_k} + E_c$. The rows of this equation can be written in terms of CI operators as follows³

$$\begin{matrix} c_0 & E_c \\ (0^+) & (2^+) \end{matrix} = \langle \Phi | V_c (\begin{matrix} C_1 \\ (1) \end{matrix} + \begin{matrix} C_2 \\ ((1^+)) \end{matrix}) | \Phi \rangle \quad (7.3)$$

$$\begin{matrix} c_a^i & (\mathcal{E}_a^i + E_c) \\ (1^+) & (0) \end{matrix} \begin{matrix} (2^+) \\ (2^+) \end{matrix} = \langle \Phi_i^a | V_c (\begin{matrix} C_0 \\ (1) \end{matrix} + \begin{matrix} C_1 \\ ((0^+)) \end{matrix} + \begin{matrix} C_2 \\ (1^+) \end{matrix} + \begin{matrix} C_3 \\ (1^+) \end{matrix} + \begin{matrix} C_4 \\ (2^+) \end{matrix}) | \Phi \rangle \quad (7.4)$$

$$\begin{matrix} c_{ab}^{ij} & (\mathcal{E}_{ab}^{ij} + E_c) \\ (1^+) & (0) \end{matrix} \begin{matrix} (2^+) \\ (2^+) \end{matrix} = \langle \Phi_{ij}^{ab} | V_c (\begin{matrix} C_0 \\ (1) \end{matrix} + \begin{matrix} C_1 \\ (0^+) \end{matrix} + \begin{matrix} C_2 \\ (1^+) \end{matrix} + \begin{matrix} C_3 \\ (1^+) \end{matrix} + \begin{matrix} C_4 \\ (2^+) \end{matrix} + \begin{matrix} C_5 \\ (2^+) \end{matrix}) | \Phi \rangle \quad (7.5)$$

$$\begin{matrix} c_{abc}^{ijk} & (\mathcal{E}_{abc}^{ijk} + E_c) \\ (2^+) & (0) \end{matrix} \begin{matrix} (2^+) \\ (2^+) \end{matrix} = \langle \Phi_{ijk}^{abc} | V_c (\begin{matrix} C_1 \\ (1) \end{matrix} + \begin{matrix} C_2 \\ (1^+) \end{matrix} + \begin{matrix} C_3 \\ (1^+) \end{matrix} + \begin{matrix} C_4 \\ (2^+) \end{matrix} + \begin{matrix} C_5 \\ (2^+) \end{matrix} + \begin{matrix} C_6 \\ (3^+) \end{matrix}) | \Phi \rangle \quad (7.6)$$

$$\begin{matrix} c_{abcd}^{ijkl} & (\mathcal{E}_{abcd}^{ijkl} + E_c) \\ (2^+) & (0) \end{matrix} \begin{matrix} (2^+) \\ (2^+) \end{matrix} = \langle \Phi_{ijkl}^{abcd} | V_c (\begin{matrix} C_2 \\ (1) \end{matrix} + \begin{matrix} C_3 \\ (1^+) \end{matrix} + \begin{matrix} C_4 \\ (2^+) \end{matrix} + \begin{matrix} C_5 \\ (2^+) \end{matrix} + \begin{matrix} C_6 \\ (3^+) \end{matrix} + \begin{matrix} C_7 \\ (3^+) \end{matrix}) | \Phi \rangle \quad (7.7)$$

$$\begin{matrix} c_{a_1 \dots a_k}^{i_1 \dots i_k} & (\mathcal{E}_{a_1 \dots a_k}^{i_1 \dots i_k} + E_c) \\ ([k/2]^+) & (0) \end{matrix} \begin{matrix} (2^+) \\ (2^+) \end{matrix} = \langle \Phi_{i_1 \dots i_k}^{a_1 \dots a_k} | V_c (\begin{matrix} C_{k-2} \\ (1) \end{matrix} + \begin{matrix} C_{k-1} \\ ([k/2]^+ - 1) \end{matrix} + \begin{matrix} C_k \\ (([k-1])/2]^+) \end{matrix} + \begin{matrix} C_{k+1} \\ ([k/2]^+) \end{matrix} + \begin{matrix} C_{k+2} \\ (([k+1])/2]^+) \end{matrix} + \begin{matrix} C_{k+3} \\ ([k/2]^+ + 1) \end{matrix}) | \Phi \rangle \quad (7.8)$$

where the numbers in parentheses denote orders in perturbation theory and the double parentheses denote terms which vanish under Brillouin's theorem. The orders of the CI operators follow from the fact that each order in perturbation theory increases the maximum excitation level of the wavefunction by +2, starting from $\Psi^{(1)}$ which contains up to doubles. Therefore the leading contributions to C_k have order $[k/2]$. If Brillouin's theorem holds, the first-order contribution to C_1 vanishes and singles contribute at orders 2^+ in perturbation theory.

Example 7.1. The analysis of remark 7.1 shows that CISD is only correct to first order in the wavefunction, since triples contribute at second order. However, C_1 and C_2 are both correct to second order, since the truncation error in equations 7.4 and 7.5 is $\mathcal{O}(V_c^3)$, making the CISD correlation energy correct to third order. In order to gain an order in perturbation theory we have to increase the truncation level by at least two, since triples and quadruples contribute at the same order. CISDTQ is correct to second order in the wavefunction and fifth order in the energy.

Proposition 7.1. CIS $\dots m$ is correct to order $\lfloor m/2 \rfloor$ in the wavefunction and order $2 \lfloor m/2 \rfloor + 1$ in the energy.

Proof: According to rmk 7.1, C_{m+1} contributes at order $\lceil (m+1)/2 \rceil$, implying that the wavefunction is correct to $\lceil (m+1)/2 \rceil - 1 = \lfloor m/2 \rfloor$. Truncation also leaves C_m and C_{m-1} correct to $\lceil (m+1)/2 \rceil = \lfloor m/2 \rfloor + 1$, and propagating these truncation errors down to C_{m-2h} and C_{m-1-2h} makes the latter correct to $\lfloor m/2 \rfloor + 1 + h$. One of these operators is C_2 when $h = \lfloor m/2 \rfloor - 1$. Since C_2 limits the error in equation 7.3,⁴ the energy is correct to $2 \lfloor m/2 \rfloor + 1$.

¹That is, Φ_k contains all unique $\Phi_{i_1 \dots i_k}^{a_1 \dots a_k}$. Uniqueness can be enforced by requiring $i_1 < \dots < i_k$ and $a_1 < \dots < a_k$.

²For the sake of generality we are not assuming intermediate normalization for CI. The C_0 operator scales functions by c_0 .

³See https://en.wikipedia.org/wiki/Floor_and_ceiling_functions for details floor and ceiling functions, $\lfloor x \rfloor$ and $\lceil x \rceil$.

⁴The error propagation ensures that errors decrease monotonically with excitation level.

Remark 7.2. Perturbative analysis of CC. The CC equations can be written as a non-linear matrix equation similar.

$$E_c \langle \Phi | \Phi \rangle - \mathbf{H}_0 \mathbf{t} = \langle \Phi | V_c \exp(T(\mathbf{t})) | \Phi \rangle_C \quad (7.9)$$

To make the comparison with equation 7.2 more transparent, this can be written as $(-\mathbf{H}_0 + E_c \langle \Phi | \Phi \rangle) \mathbf{t} = (\mathbf{V}_c \mathbf{t})_C + \mathcal{O}(\mathbf{t}^2)$. Non-vanishing contributions to the rows of this matrix equation can be expressed in terms of cluster operators as follows

$$E_c = \langle \Phi | V_c (\underset{(2^+)}{T_1} + \underset{(1) ((1^+))}{T_2} + \underset{(1^+)}{T_2} + \underset{(2^+)}{\frac{1}{2} T_1^2}) | \Phi \rangle_C \quad (7.10)$$

$$t_{(1^+)(0)}^i \mathcal{E}_a^i = \langle \Phi_i^a | V_c (\underset{(1)}{1} + \underset{((0))}{T_1} + \underset{(1^+)}{T_2} + \underset{(1^+)}{T_2} + \underset{(2^+)}{T_3} + \underset{(2^+)}{\frac{1}{2} T_1^2} + \underset{(2^+)}{T_1 T_2} + \underset{(3^+)}{\frac{1}{3!} T_1^3}) | \Phi \rangle_C \quad (7.11)$$

$$t_{(1^+)(0)}^{ij} \mathcal{E}_{ab}^{ij} = \langle \Phi_{ij}^{ab} | V_c (\underset{(1)}{1} + \underset{(0)}{T_1} + \underset{(1^+)}{T_2} + \underset{(1^+)}{T_2} + \underset{(2^+)}{T_3} + \underset{(3^+)}{T_4} + \underset{(2^+)}{\frac{1}{2} T_1^2} + \underset{(2^+)}{T_1 T_2} + \underset{(3^+)}{T_1 T_3} + \underset{(2^+)}{\frac{1}{2} T_2^2} + \underset{(3^+)}{\frac{1}{3!} T_1^3} + \underset{(3^+)}{\frac{1}{2} T_1^2 T_2} + \underset{(4^+)}{\frac{1}{4!} T_1^4}) | \Phi \rangle_C \quad (7.12)$$

$$t_{(2^+)(0)}^{ijk} \mathcal{E}_{abc}^{ijk} = \langle \Phi_{ijk}^{abc} | V_c (\underset{(1)}{T_2} + \underset{(1^+)}{T_3} + \underset{(2^+)}{T_4} + \underset{(3^+)}{T_5} + \underset{(2^+)}{T_1 T_2} + \underset{(3^+)}{T_1 T_3} + \underset{(2^+)}{\frac{1}{2} T_2^2} + \underset{(4^+)}{T_1 T_4} + \underset{(3^+)}{T_2 T_3} + \underset{(3^+)}{\frac{1}{2} T_1^2 T_2} + \underset{(4^+)}{\frac{1}{2} T_1^2 T_3} + \underset{(3^+)}{\frac{1}{2} T_1 T_2^2} + \underset{(4^+)}{\frac{1}{3!} T_1^3 T_2}) | \Phi \rangle_C \quad (7.13)$$

$$t_{(3^+)(0)}^{ijkl} \mathcal{E}_{abcd}^{ijkl} = \langle \Phi_{ijkl}^{abcd} | V_c (\underset{(1)}{T_3} + \underset{(2^+)}{T_4} + \underset{(3^+)}{T_5} + \underset{(4^+)}{T_6} + \underset{(5^+)}{T_1 T_3} + \underset{(2^+)}{\frac{1}{2} T_2^2} + \underset{(4^+)}{T_1 T_4} + \underset{(3^+)}{T_2 T_3} + \underset{(5^+)}{T_1 T_5} + \underset{(4^+)}{T_2 T_4} + \underset{(4^+)}{\frac{1}{2} T_3^2} + \underset{(4^+)}{\frac{1}{2} T_1^2 T_3} + \underset{(3^+)}{\frac{1}{2} T_1 T_2^2} + \underset{(5^+)}{\frac{1}{2} T_1^2 T_4} + \underset{(4^+)}{T_1 T_2 T_3} + \underset{(3^+)}{\frac{1}{3!} T_2^3} + \underset{(5^+)}{\frac{1}{3!} T_1^3 T_3} + \underset{(4^+)}{\frac{1}{2!2!} T_1^2 T_2^2}) | \Phi \rangle_C \quad (7.14)$$

⋮

$$t_{((k-1)^+)(0)}^{i_1 \dots i_k} \mathcal{E}_{a_1 \dots a_k}^{i_1 \dots i_k} = \langle \Phi_{i_1 \dots i_k}^{a_1 \dots a_k} | V_c (\underset{(1)}{T_{k-1}} + \underset{((k-2)^+)}{T_k} + \underset{((k-1)^+)}{T_{k+1}} + \underset{(k^+)}{T_{k+2}} + \sum_{p=2}^4 \frac{1}{p!} \sum_{h=p-2}^2 \sum_{\mathbf{k}}^{C_p(k+h)} T_{k_1} \dots T_{k_p}) | \Phi \rangle_C \quad (7.15)$$

where $\mathcal{C}_p(m)$ is the set of k -part compositions of m and $\phi_{\mathbf{k}}$ is the number of 1's in $\mathbf{k} = (k_1, \dots, k_p)$. The orders of the cluster operators follow from straightforward induction on the fact that the lowest order contribution to T_k always comes from $(V_c T_{k-1})_C$, which means that each T_k contributes at one order above T_{k-1} , starting from $k = 2$. If Brillouin's theorem holds, T_1 contributes at orders 2^+ and the orders of the disconnected products become $(k + h - p + 2 \cdot \phi_{\mathbf{k}})^+$.

Proposition 7.2. *CCS $\dots m$ is correct to order $m - 1$ in the wavefunction and order $m + \lfloor m/2 \rfloor$ in the energy.*

Proof: According to rmk 7.2, T_{m+1} contributes at order m , implying that the wavefunction is correct to order $m - 1$. Truncation also leaves T_m and T_{m-1} correct to order m , and propagating these truncation errors down to T_{m-2h} and T_{m-1-2h} makes the latter correct to $m + h$. One of these operators is T_2 when $h = \lfloor m/2 \rfloor - 1$. Since T_2 limits the error in equation 7.10, the energy is correct to order $m + \lfloor m/2 \rfloor$.

Example 7.2. Props 7.1 and 7.2 allow us to compare the accuracies of CI and CC in perturbation theory. Truncating at doubles, CI and CC are both correct to first order in the wavefunction and third order in the energy. Triples yield no improvement for CI, whereas CC gains an order in both wavefunction and energy. In general, the CC wavefunction and energy improves upon the CI wavefunction and energy by $m - \lfloor m/2 \rfloor - 1 = \lfloor (m - 1)/2 \rfloor$ orders in perturbation theory.

Definition 7.3. Order p truncation. If X is a polynomial in T_1, T_2, \dots, T_n , its *order p truncation*, denoted $X^{[p]}$, retains all terms in the polynomial with leading contributions of order p or less. This makes $X^{[p]}$ correct to order p without isolating specific orders in the cluster operators, which will generally involve contributions up to infinite order.

Example 7.3. The [T] correction. Assuming Brillouin's theorem, we can complete the energy to fourth order using

$$t_{ab}^{ij} = \langle \Phi_{ij}^{ab} | R_0 V_c (1 + T_1 + T_2 + T_3^{[2]} + \frac{1}{2} T_1^2 + T_1 T_2 + \frac{1}{2} T_2^2 + \frac{1}{3!} T_1^3 + \frac{1}{2} T_1^2 T_2 + \frac{1}{4!} T_1^4) | \Phi \rangle_C \quad [2] t_{abc}^{ijk} = \langle \Phi_{ijk}^{abc} | R_0 V_c T_2 | \Phi \rangle_C \quad (7.16)$$

where the resulting energy correction is $E_e - E_e^{\text{CCSD}} = \langle \Phi | V_c R_0 V_c T_3^{[2]} | \Phi \rangle$. We can introduce additional infinite order contributions by noting that $T_2 \Phi = R_0 V_c \Phi + \mathcal{O}(V_c^2)$ and that the additional terms in $\langle \Phi | T_2^\dagger V_c T_3^{[2]} | \Phi \rangle$ are also valid energy contributions in perturbation theory. There is no risk of double counting since all of these contributions involve connected triples, which are absent in CCSD. With converged CCD or CCSD T_2 -amplitudes, this defines the “brackets” T correction

$$E_{[T]} = \langle \Phi | T_2^\dagger V_c T_3^{[2]} | \Phi \rangle = (\frac{1}{3!})^2 \sum_{\substack{abc \\ ijk}}^{[2]} t_{abc}^{ijk*} \mathcal{E}_{abc}^{ijk[2]} t_{abc}^{ijk} \quad \equiv \quad \text{diagrams} \quad (7.17)$$

which is $\langle \Phi | T_2^\dagger V_c T_3^{[2]} | \Phi \rangle = -\langle \Phi | T_3^{[2]\dagger} H_0 T_3^{[2]} | \Phi \rangle$. The dash-dotted line represents $-H_0$ and is analogous to a resolvent line.

Remark 7.3. *CC excited states via equation-of-motion theory.* Similarity transformations preserve operator eigenvalues, so \bar{H}_e has the same spectrum as the original Hamiltonian, H_e . The corresponding eigenvalue equations

$$\bar{\mathbf{H}}_e \mathbf{r}_k = E_k \mathbf{r}_k \quad \mathbf{l}_k^\dagger \bar{\mathbf{H}}_e = \mathbf{l}_k^\dagger E_k \quad \mathbf{l}_k^* \cdot \mathbf{r}_l = \delta_{kl} \quad (7.18)$$

involve both right- and left-eigenvectors since $\bar{\mathbf{H}}_e = \langle \Phi | (H_e \exp(T))_C | \Phi \rangle$ is non-Hermitian. These can also be written as

$$\bar{H}_e \mathcal{R}_k | \Phi \rangle = E_k \mathcal{R}_k | \Phi \rangle \quad \langle \Phi | \bar{H}_e \mathcal{L}_k = \langle \Phi | \mathcal{L}_k E_k \quad \begin{aligned} \mathcal{R}_k | \Phi \rangle &= | \Phi \rangle \cdot \mathbf{r}_k \\ \langle \Phi | \mathcal{L}_k &= \mathbf{l}_k^* \cdot \langle \Phi | \end{aligned} \quad (7.19)$$

in terms of abstract states. E_k denotes the k^{th} excitation energy, which matches the corresponding CI eigenvalue unless we truncate the Ansatz. In general, the expectation value of an observable can be determined as $\langle \Psi_k | W | \Psi_k \rangle = \langle \Phi | \mathcal{L}_k \bar{W} \mathcal{R}_k | \Phi \rangle$.

Definition 7.4. *The CC Lagrangian.* Assuming we have solved equation 7.9, the right eigenvector of the ground state is simply a unit vector $\mathbf{r}_e = \langle \Phi | \Phi \rangle$. The left ground-state eigenvector is unknown, but the biorthonormality condition in equation 7.18 implies that its first entry is one. Therefore, the left and right ground-state wave operators have the form

$$\mathcal{R}_e = 1 \quad \mathcal{L}_e = 1 + \Lambda \quad \Lambda = \Lambda_1 + \dots + \Lambda_n \quad \Lambda_k \equiv \left(\frac{1}{k!}\right)^2 \lambda_{i_1 \dots i_k}^{a_1 \dots a_k} \tilde{a}_{a_1 \dots a_k}^{i_1 \dots i_k} \quad (7.20)$$

where Λ is a linear de-excitation operator analogous to C^\dagger . The ground-state expectation value is therefore

$$E_e = \langle \Psi_e | H_e | \Psi_e \rangle = \langle \Phi | (1 + \Lambda) \bar{H}_e | \Phi \rangle \quad (7.21)$$

which can be identified as a Lagrangian for the coupled-cluster energy. To see why, note that setting the $\boldsymbol{\lambda}$ -gradient equal to zero yields the CC amplitude equations: $\langle \Phi | \bar{H}_e | \Phi \rangle = 0$. If these are satisfied, then $E_e = \langle \Phi | \bar{H}_e | \Phi \rangle$ gives the CC energy. Therefore, we can view the $\boldsymbol{\lambda}$ as Lagrange multipliers enforcing equation 7.9 and reduce ground-state CC to a constrained optimization problem.

Definition 7.5. *The CC lambda equations.* Projecting the effective Schrödinger equation for the left ground-state eigenbra⁵ on the right by $|\Phi_{i_1 \dots i_k}^{a_1 \dots a_k}\rangle$ and rearranging yields the following form of the *coupled-cluster lambda equations*

$$\lambda_{i_1 \dots i_k}^{a_1 \dots a_k} \mathcal{E}_{a_1 \dots a_k}^{i_1 \dots i_k} = \langle \Phi | \Lambda H_0 T + V_c + V_c T + \Lambda V_c \exp(T) | \Phi_{i_1 \dots i_k}^{a_1 \dots a_k} \rangle_C \quad (7.22)$$

where we have used $\lambda_{i_1 \dots i_k}^{a_1 \dots a_k} E_c = \langle \Phi | \Lambda (V_c \exp(T))_C | \Phi_{i_1 \dots i_k}^{a_1 \dots a_k} \rangle_U$ to cancel the unlinked contributions and the subscript C denotes that both the operators in Λ and those in $\exp(T)$ must be connected to V_c .

Example 7.4. Assuming Brillouin's theorem holds true, the CCD lambda equations have the following form.

$$\lambda_{ij}^{ab} \mathcal{E}_{ab}^{ij} = \langle \Phi | V_c + \Lambda_2 V_c + \Lambda_2 V_c T_2 | \Phi_{ij}^{ab} \rangle_C \quad (7.23)$$

⁵That is, $\langle \Phi | (1 + \Lambda) \bar{H}_e = \langle \Phi | (1 + \Lambda) E_e$.

Definition 7.6. The [T] correction. For simplicity, let T be the CCSD cluster operator and let \bar{H}_c and E_c be the corresponding effective Hamiltonian and correlation energy. Then the “brackets” T correction can be derived by introducing a unitary exponential triples operator $e^{T_3-T_3^\dagger}$ in front of the CCSD wave operator, $\Psi_{\text{CCSD}[T]} = e^T e^{T_3-T_3^\dagger} \Phi$, and determining its lowest-order contribution to the energy. Multiplying the Schrödinger equation by the inverse wave operator, projecting by $\langle \Phi |$,⁶ subtracting off the original CCSD energy, and keeping all terms up to fifth order in perturbation theory yields

$$E_{[T]} = \langle \Phi | T_3^\dagger (V_c T_2)_C | \Phi \rangle \quad (7.24)$$

which corrects the energy

Definition 7.7. Let T , \bar{H}_c , and E_c be the cluster operator, effective Hamiltonian, and correlation energy for a truncated coupled-cluster Ansatz, CCSD $\cdots m$, with m greater than two. Then the “brackets” $m+1$ correction, denoted $E_{[m+1]}$, can be derived by introducing a unitary exponential to the wave operator, .

$$= \langle \Phi | e^{T_{m+1}^\dagger - T_{m+1}} (H_c \exp(T))_C e^{T_{m+1} - T_{m+1}^\dagger} | \Phi \rangle \approx E_c + \langle \Phi | T_{m+1}^\dagger (V_c T_2)_C | \Phi \rangle + \mathcal{O} \quad (7.25)$$

Remark 7.4. From PT, we know that the lowest order contributions to T_1 , T_2 , and T_3 (or C_1 and C_2) occur at first and second order in perturbation theory.

Now, just to be confusing, redefine some shit.

$$\mathbf{1}_i \equiv \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \mathbf{1}_e \equiv \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \quad \mathbf{H}_{xy} \equiv \mathbf{1}_x \mathbf{H} \mathbf{1}_y \quad \mathbf{c}_x \equiv \mathbf{1}_x \mathbf{c} \quad (7.26)$$

$$\mathbf{H} = \mathbf{H}_{ii} + \mathbf{H}_{ie} + \mathbf{H}_{ei} + \mathbf{H}_{ee} \quad (7.27)$$

$$\mathbf{R}_{ee} \equiv (E - \mathbf{H})^{-1} \big|_e \quad \mathbf{R}_{ee}(E - \mathbf{H})\mathbf{1}_e = \mathbf{1}_e \quad \begin{aligned} \mathbf{R}_{ee}(E - \mathbf{H}) &= -\mathbf{R}_{ee} \mathbf{H}_{ei} + \mathbf{1}_e \\ (E - \mathbf{H}) \mathbf{R}_{ee} &= -\mathbf{H}_{ie} \mathbf{R}_{ee} + \mathbf{1}_e \end{aligned} \quad (7.28)$$

Operating the upper equation on \mathbf{c} gives zero due to the Schrödinger equation, which implies $\mathbf{c}_e = \mathbf{R}_{ee} \mathbf{H}_{ei} \mathbf{c}_i$. Projecting the Schrödinger equation by $\mathbf{1}_i$ and substituting in this result leads to the following.

$$(\mathbf{H}_{ii} + \mathbf{V}_{ii})\mathbf{c}_i = E\mathbf{c}_i \quad \mathbf{V}_{ii} \equiv \mathbf{H}_{ie} \mathbf{R}_{ee} \mathbf{H}_{ei} \quad (7.29)$$

$$E = \frac{\mathbf{c}_i^\dagger (\mathbf{H}_{ii} + \mathbf{V}_{ii}) \mathbf{c}_i}{\mathbf{c}_i \cdot \mathbf{c}_i} \quad (7.30)$$

$$E = \langle \Phi | (1 + \Lambda) \bar{H} | \Phi \rangle + \langle \Phi | \Lambda \bar{H} | \mathbf{e} \rangle \langle \mathbf{e} | E - \bar{H} | \mathbf{e} \rangle^{-1} \langle \mathbf{e} | \bar{H} | \Phi \rangle$$

$$\delta E = \langle \Phi | \Lambda \bar{H} | \mathbf{e} \rangle \langle \mathbf{e} | E - \bar{H} | \mathbf{e} \rangle^{-1} \langle \mathbf{e} | \bar{H} | \Phi \rangle \approx \langle \Phi | \Lambda \bar{H}^{(1)} | \mathbf{e} \rangle \langle \mathbf{e} | E^{(0)} - \bar{H}^{(0)} | \mathbf{e} \rangle^{-1} \langle \mathbf{e} | \bar{H}^{(m)} | \Phi \rangle$$

$$\bar{H} = E_{\text{ref}} + H_0 + (H_0 T + V_c \exp(T))_C \quad (7.31)$$

$$\bar{H}^{(0)} = E_{\text{ref}} + H_0 \quad \bar{H}^{(1)} = (H_0 T^{(1)})_C + V_c \quad \bar{H}^{(2)} = (H_0 T^{(2)} + V_c T^{(1)})_C \quad (7.32)$$

Structure of matrix

$$\bar{\mathbf{H}}_c = \begin{bmatrix} E_c & \langle \Phi | \bar{H}_c | \mathbf{i} \rangle & \langle \Phi | \bar{H}_c | \mathbf{e} \rangle \\ \mathbf{0} & \langle \mathbf{i} | \bar{H}_c | \mathbf{i} \rangle & \langle \mathbf{i} | \bar{H}_c | \mathbf{e} \rangle \\ \langle \mathbf{e} | \bar{H}_c | \Phi \rangle & \langle \mathbf{e} | \bar{H}_c | \mathbf{i} \rangle & \langle \mathbf{e} | \bar{H}_c | \mathbf{e} \rangle \end{bmatrix} \quad \mathbf{i} \equiv [\mathbf{s}_1 \quad \cdots \quad \mathbf{s}_m] \quad \mathbf{e} \equiv [\mathbf{s}_{m-1} \quad \cdots \quad \mathbf{s}_n] \quad \mathbf{s}_h \equiv [\cdots \quad \Phi_{i_1 \dots i_h}^{a_1 \dots a_h} \quad \cdots]$$

⁶These two steps yield $E_c + E_{[T]} + \mathcal{O}(V_c^5) = \langle \Phi | e^{T_3^\dagger - T_3} \bar{H}_c e^{T_3 - T_3^\dagger} | \Phi \rangle = E_c + \langle \Phi | T_3^\dagger \bar{H}_c | \Phi \rangle + \langle \Phi | T_3^{\dagger 2} \bar{H}_c | \Phi \rangle + \langle \Phi | T_3^\dagger \bar{H}_c T_3 | \Phi \rangle + \mathcal{O}(T_3^3)$.

A Faà di Bruno's formula

Theorem A.1. *Faà di Bruno's formula.*

$$\frac{\partial^n}{\partial x_1 \cdots \partial x_n} f(g(\mathbf{x})) = \sum_{k=1}^n \sum_{(\mathbf{x}_1, \dots, \mathbf{x}_k) \in \mathcal{P}_k(\mathbf{x})} f^{(k)}(g(\mathbf{x})) \prod_{i=1}^k \frac{\partial^{|\mathbf{x}_i|} g(\mathbf{x})}{\partial x_{i,1} \cdots \partial x_{i,|\mathbf{x}_i|}} \quad (\text{A.1})$$