7 Perturbative analysis

Definition 7.1. *Correct to order p.* Let $X^{(p^-)}$ and $X^{(p^+)}$ denote $X^{(0)}+X^{(1)}+\cdots+X^{(p)}$ and $X^{(p)}+X^{(p+1)}+\cdots+X^{(\infty)}$. We say that 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 CISDTQPH7$\cdots m$ denote truncation of the CI Ansatz at m-tuples. Similarly, CCS$\cdots m$ means that we omit cluster operators of excitation level $k > m$. 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 $\boldsymbol{\Phi}_k$ be a row vector containing all unique k-fold substitutions of Φ,[1] so that $\boldsymbol{\Phi} = \begin{bmatrix} \Phi & \boldsymbol{\Phi}_1 & \cdots & \boldsymbol{\Phi}_n \end{bmatrix}$ 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]

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

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

Remark 7.1. *Perturbative analysis of CI.* Writing the CI eigenvalue equation $\mathbf{H}_\mathrm{c}\,\mathbf{c} = E_\mathrm{c}\,\mathbf{c}$ in terms of model-Hamiltonian and fluctuation-potential matrices leads to the following rearranged matrix equation

$$\left(-\mathbf{H}_0 + E_\mathrm{c}\right)\mathbf{c} = \mathbf{V}_\mathrm{c}\,\mathbf{c} \qquad (7.2)$$

which provide a convenient starting point for perturbative analysis and for comparison to the coupled-cluster equations. The matrix elements of the model Hamiltonian are given by $\langle \Phi_\sigma | H_0 | \Phi_\tau \rangle = \mathcal{E}_\tau \delta_{\sigma\tau}$, so the matrix on the left is diagonal with eigenvalues $-\mathcal{E}_{i_1 \cdots i_k}^{a_1 \cdots a_k} + E_\mathrm{c}$. The rows of this equation can be written in terms of CI operators as follows[3]

$$\underset{(0^+)(2^+)}{c_0\ E_\mathrm{c}} = \langle \Phi | V_\mathrm{c} (\underset{(1)}{C_1}\ +\ \underset{((1^+))}{C_2})|\Phi\rangle \qquad (7.3)$$

$$\underset{(1^+)\ (0)}{c_a^i\ (\mathcal{E}_a^i}\ +\ \underset{(2^+)}{E_\mathrm{c})} = \langle \Phi_i^a | V_\mathrm{c} (\underset{(1)}{C_0}\ +\ \underset{((0^+))}{C_1}\ +\ \underset{(1^+)}{C_2}\ +\ \underset{(1^+)}{C_3})|\Phi\rangle \qquad (7.4)$$

$$\underset{(1^+)\ (0)}{c_{ab}^{ij}\ (\mathcal{E}_{ab}^{ij}}\ +\ \underset{(2^+)}{E_\mathrm{c})} = \langle \Phi_{ij}^{ab} | V_\mathrm{c} (\underset{(1)}{C_0}\ +\ \underset{(0^+)}{C_1}\ +\ \underset{(1^+)}{C_2}\ +\ \underset{(1^+)}{C_3}\ +\ \underset{(2^+)}{C_4})|\Phi\rangle \qquad (7.5)$$

$$\underset{(2^+)\ (0)}{c_{abc}^{ijk}(\mathcal{E}_{abc}^{ijk}}\ +\ \underset{(2^+)}{E_\mathrm{c})} = \langle \Phi_{ijk}^{abc} | V_\mathrm{c} (\underset{(1)}{C_1}\ +\ \underset{(1^+)}{C_2}\ +\ \underset{(1^+)}{C_3}\ +\ \underset{(2^+)}{C_4}\ +\ \underset{(2^+)}{C_5})|\Phi\rangle \qquad (7.6)$$

$$\underset{(2^+)\ (0)}{c_{abcd}^{ijkl}(\mathcal{E}_{abcd}^{ijkl}}\ +\ \underset{(2^+)}{E_\mathrm{c})} = \langle \Phi_{ijkl}^{abcd} | V_\mathrm{c} (\underset{(1)}{C_2}\ +\ \underset{(1^+)}{C_3}\ +\ \underset{(2^+)}{C_4}\ +\ \underset{(2^+)}{C_5}\ +\ \underset{(3^+)}{C_6})|\Phi\rangle \qquad (7.7)$$

$$\vdots$$

$$\underset{(\lceil k/2\rceil^+)}{c_{a_1\cdots a_k}^{i_1\cdots i_k}}\ (\underset{(0)}{\mathcal{E}_{a_1\cdots a_k}^{i_1\cdots i_k}}\ +\ \underset{(2^+)}{E_\mathrm{c})} = \langle \Phi_{i_1\cdots i_k}^{a_1\cdots a_k} | V_\mathrm{c} (\underset{(1)}{C_{k-2}}\ +\ \underset{(\lceil k/2\rceil^+-1)}{C_{k-1}}\ +\ \underset{(\lceil (k-1)/2\rceil^+)}{C_k}\ +\ \underset{(\lceil k/2\rceil^+)}{C_{k+1}}\ +\ \underset{(\lceil (k+1)/2\rceil^+)}{C_{k+2}})|\Phi\rangle \quad (7.8)$$

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 $\lceil k/2 \rceil$. If Brillouin's theorem holds, the first-order contribution to C_1 vanishes and singles contribute at order 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, ensuring that the the CISD correlation energy is 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$\cdots m$ 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, the leading term in C_{m+1} has order $\lceil (m+1)/2 \rceil$. This implies that the wavefunction is correct to $\lceil (m+1)/2 \rceil - 1 = \lfloor m/2 \rfloor$ and that C_m and C_{m-1} are correct to $\lceil (m+1)/2 \rceil = \lfloor m/2 \rfloor + 1$. Propagating these truncation errors down to C_{m-2h} and C_{m-1-2h} leaves the latter correct to $\lfloor m/2 \rfloor + 1 + h$. One of these operators equals 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 $2\lfloor m/2 \rfloor + 1$.

---

[1] That is, $\boldsymbol{\Phi}_k = [\Phi_{i_1\cdots i_k}^{a_1\cdots a_k}]$. Uniqueness can be enforced by requiring $i_1 < \cdots < i_k$ and $a_1 < \cdots < a_k$.

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

[3] See `https://en.wikipedia.org/wiki/Floor_and_ceiling_functions` for details on the floor and ceiling functions.

[4] The 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.

$$E_{\mathrm{c}}\langle\boldsymbol{\Phi}|\Phi\rangle - \mathbf{H}_0\mathbf{t} = \langle\boldsymbol{\Phi}|V_{\mathrm{c}}\exp(T(\mathbf{t}))|\Phi\rangle_{\mathrm{C}} \tag{7.9}$$

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

$$\underset{(2^+)}{E_{\mathrm{c}}} = \langle\Phi|V_{\mathrm{c}}(\underset{(1)}{T_1} + \underset{((1^+))}{T_2} + \underset{(1^+)}{\tfrac{1}{2}T_1^2})|\Phi\rangle_{\mathrm{C}} \tag{7.10}$$

$$\underset{(1^+)(0)}{t_a^i\,\mathcal{E}_a^i} = \langle\Phi_i^a|V_{\mathrm{c}}(\underset{(1)}{1} + \underset{((0))}{T_1} + \underset{(1^+)}{T_2} + \underset{(1^+)}{T_3} + \underset{(2^+)}{\tfrac{1}{2}T_1^2} + \underset{(2^+)}{T_1T_2} + \underset{(3^+)}{\tfrac{1}{3!}T_1^3})|\Phi\rangle_{\mathrm{C}} \tag{7.11}$$

$$\underset{(1^+)(0)}{t_{ab}^{ij}\,\mathcal{E}_{ab}^{ij}} = \langle\Phi_{ij}^{ab}|V_{\mathrm{c}}(\underset{(1)}{1} + \underset{(0)}{T_1} + \underset{(1^+)}{T_2} + \underset{(1^+)}{T_3} + \underset{(2^+)}{T_4} + \underset{(3^+)}{\tfrac{1}{2}T_1^2} + \underset{(2^+)}{T_1T_2} + \underset{(2^+)}{T_1T_3}$$
$$+ \underset{(2^+)}{\tfrac{1}{2}T_2^2} + \underset{(3^+)}{\tfrac{1}{3!}T_1^3} + \underset{(3^+)}{\tfrac{1}{2}T_1^2T_2} + \underset{(4^+)}{\tfrac{1}{4!}T_1^4})|\Phi\rangle_{\mathrm{C}} \tag{7.12}$$

$$\underset{(2^+)(0)}{t_{abc}^{ijk}\,\mathcal{E}_{abc}^{ijk}} = \langle\Phi_{ijk}^{abc}|V_{\mathrm{c}}(\underset{(1)}{T_2} + \underset{(1^+)}{T_3} + \underset{(2^+)}{T_4} + \underset{(3^+)}{T_5} + \underset{(4^+)}{T_1T_2} + \underset{(2^+)}{T_1T_3} + \underset{(3^+)}{\tfrac{1}{2}T_2^2} + \underset{(2^+)}{T_1T_4} + \underset{(4^+)}{T_2T_3}$$
$$+ \underset{(3^+)}{\tfrac{1}{2}T_1^2T_2} + \underset{(4^+)}{\tfrac{1}{2}T_1^2T_3} + \underset{(3^+)}{\tfrac{1}{2}T_1T_2^2} + \underset{(4^+)}{\tfrac{1}{3!}T_1^3T_2})|\Phi\rangle_{\mathrm{C}} \tag{7.13}$$

$$\underset{(3^+)(0)}{t_{abcd}^{ijkl}\,\mathcal{E}_{abcd}^{ijkl}} = \langle\Phi_{ijkl}^{abcd}|V_{\mathrm{c}}(\underset{(1)}{T_3} + \underset{(2^+)}{T_4} + \underset{(3^+)}{T_5} + \underset{(4^+)}{T_6} + \underset{(5^+)}{T_1T_3} + \underset{(2^+)}{\tfrac{1}{2}T_2^2} + \underset{(2^+)}{T_1T_4} + \underset{(4^+)}{T_2T_3} + \underset{(3^+)}{T_1T_5} + \underset{(5^+)}{T_2T_4} + \underset{(4^+)}{\tfrac{1}{2}T_3^2}$$
$$+ \underset{(4^+)}{\tfrac{1}{2}T_1^2T_3} + \underset{(3^+)}{\tfrac{1}{2}T_1T_2^2} + \underset{(5^+)}{\tfrac{1}{2}T_1^2T_4} + \underset{(4^+)}{T_1T_2T_3} + \underset{(3^+)}{\tfrac{1}{3!}T_2^3} + \underset{(5^+)}{\tfrac{1}{3!}T_1^3T_3} + \underset{(4^+)}{\tfrac{1}{2!2!}T_1^2T_2^2})|\Phi\rangle_{\mathrm{C}} \tag{7.14}$$

$$\vdots$$

$$\underset{((k-1)^+)\,(0)}{t_{a_1\cdots a_k}^{i_1\cdots i_k}\,\mathcal{E}_{a_1\cdots a_k}^{i_1\cdots i_k}} = \langle\Phi_{i_1\cdots i_k}^{a_1\cdots a_k}|V_{\mathrm{c}}(\underset{(1)}{T_{k-1}} + \underset{((k-2)^+)}{T_k} + \underset{((k-1)^+)}{T_{k+1}} + \underset{(k^+)}{T_{k+2}} + \sum_{p=2}^{4}\frac{1}{p!}\sum_{h=p-2}^{2}\underset{((k+h-p+\phi_k)^+)}{\sum_{\boldsymbol{k}}^{\mathcal{C}_p(k+h)}T_{k_1}\cdots T_{k_p}})|\Phi\rangle_{\mathrm{C}} \tag{7.15}$$

where $\mathcal{C}_k(m)$ denotes the set of k-part integer compositions of m and $\phi_{\boldsymbol{k}}$ is the number of 1's in $\boldsymbol{k} = (k_1,\ldots,k_p)$. The orders of the cluster operators follow from straightforward induction on the fact that the lowest order contribution to T_k always comes from $(V_{\mathrm{c}}T_{k-1})_{\mathrm{C}}$, 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 has order 2^+ and the orders of the disconnected products become $(k + h - p + 2\phi_{\boldsymbol{k}})^+$.

Proposition 7.2. CCS$\cdots m$ *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. At double excitations, 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 CCS$\cdots m$ wavefunction and energy improve upon the wavefunction and energy obtained from CIS$\cdots m$ 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 $T_1, T_2, \ldots, T_n$, we define its *order p truncation*, denoted $X^{[p]}$, to include 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 in general have infinite-order contributions.

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

$$t_{ab}^{ij} = \langle\Phi_{ij}^{ab}|R_0V_{\mathrm{c}}(1 + T_1 + T_2 + T_3^{[2]} + \tfrac{1}{2}T_1^2 + T_1T_2 + \tfrac{1}{2}T_2^2 + \tfrac{1}{3!}T_1^3 + \tfrac{1}{2}T_1^2T_2 + \tfrac{1}{4!}T_1^4)|\Phi\rangle_{\mathrm{C}} \qquad {}^{[2]}t_{abc}^{ijk} = \langle\Phi_{ijk}^{abc}|R_0V_{\mathrm{c}}T_2|\Phi\rangle_{\mathrm{C}} \tag{7.16}$$

where the resulting energy correction is $E_e - E_e^{\mathrm{CCSD}} = \langle\Phi|V_{\mathrm{c}}R_0V_{\mathrm{c}}T_3^{[2]}|\Phi\rangle$. We can introduce additional infinite order contributions by noting that $T_2\Phi = R_0V_{\mathrm{c}}\Phi + \mathcal{O}(V_{\mathrm{c}}^2)$ and that the additional terms in $\langle\Phi|T_2^\dagger V_{\mathrm{c}}T_3^{[2]}|\Phi\rangle$ 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 [T] *correction.*

$$E_{[\mathrm{T}]} \equiv \langle\Phi|T_2^\dagger V_{\mathrm{c}}T_3^{[2]}|\Phi\rangle =$$

$$\tag{7.17}$$

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

$$\overline{\mathbf{H}}_e \mathbf{r}_k = E_k \mathbf{r}_k \qquad\qquad \mathbf{l}_k^\dagger \overline{\mathbf{H}}_e = \mathbf{l}_k^\dagger E_k \qquad\qquad \mathbf{l}_k^* \cdot \mathbf{r}_l = \delta_{kl} \tag{7.18}$$

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

$$\overline{H}_e\,{}^kR|\Phi\rangle = E_k\,{}^kR|\Phi\rangle \qquad \langle\Phi|\overline{H}_e\,{}^kL = \langle\Phi|\,{}^kLE_k \qquad R = R_0 + R_1 + \cdots + R_n \qquad L = L_0 + L_1 + \cdots + L_n \tag{7.19}$$

where R and L are linear excitation and de-excitation operators analogous to $(C_0 + C)$ and $(C_0 + C)^\dagger$, respectively. These are the *equation-of-motion coupled-cluster (EOM-CC) equations*. E_k is the kth excited state energy, which matches the corresponding CI eigenvalue unless we truncate the basis. In general, the expectation value of an observable can be determined as $\langle\Psi_k|W|\Psi_k\rangle = \langle\Phi|\,{}^kL\overline{W}\,{}^kR|\Phi\rangle$ and its transition matrix element is given by $\langle\Psi_k|W|\Psi_l\rangle = \langle\Phi|\,{}^kL\overline{W}\,{}^lR|\Phi\rangle$.[5]

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 $\mathbf{r}_0 = \langle\mathbf{\Phi}|\mathbf{\Phi}\rangle$. The left ground-state eigenvector is unknown, but the biorthonormality condition in equation 7.18 tells us that its first entry equals 1. Therefore, the left and right ground-state wave operators have the form

$$^0R = 1 \qquad\qquad ^0L = 1 + \Lambda \qquad\qquad \Lambda = \Lambda_1 + \cdots + \Lambda_n \tag{7.20}$$

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

$$E_e = \langle\Psi_e|H_e|\Psi_e\rangle = \langle\Phi|(1+\Lambda)\overline{H}_e|\Phi\rangle = \mathcal{L}(\mathbf{t},\boldsymbol{\lambda}) \tag{7.21}$$

which is the *coupled-cluster Lagrangian*. To see why this constitutes a Lagrangian, note that setting its $\boldsymbol{\lambda}$-gradient equal to zero yields the CC amplitude equations: $\langle\mathbf{\Phi}|\overline{H}_e|\Phi\rangle = 0$. If these are satisfied, then $E_e = \langle\Phi|\overline{H}_e|\Phi\rangle$ gives the CC energy. Therefore, we can view the $\boldsymbol{\lambda}$ coefficients as Lagrange multipliers enforcing equation 7.9 as a constraint.

Definition 7.5. *The CC Lambda equations.* Setting the $\mathbf{t}$-gradient of the CC Lagrangian equal to zero gives

$$\langle\Phi|(1+\Lambda)H_{\mathrm{c}}\exp(T)|\Phi_{i_1\cdots i_k}^{a_1\cdots a_k}\rangle_{\mathrm{C}} \overset{!}{=} 0 \tag{7.22}$$

which are the CC Lambda equations.[6] Rearranging this equation gives an alternative, more convenient form

$$\lambda_{i_1\cdots i_k}^{a_1\cdots a_k}\mathcal{E}_{a_1\cdots a_k}^{i_1\cdots i_k} = \langle\Phi|\Lambda H_0 T + (1+\Lambda)V_{\mathrm{c}}\exp(T)|\Phi_{i_1\cdots i_k}^{a_1\cdots a_k}\rangle_{\mathrm{C}} \tag{7.23}$$

which sets up the iterative procedure for determining $\lambda_{i_1\cdots i_k}^{a_1\cdots a_k}$ from a given set of amplitudes.

Example 7.4. Assuming Brillouin's theorem holds, the CCD lambda equations are as follows.

$$\lambda_{ij}^{ab}\mathcal{E}_{ab}^{ij} = \langle\Phi|V_{\mathrm{c}} + \Lambda_2 V_{\mathrm{c}} + \Lambda_2 V_{\mathrm{c}} T_2|\Phi_{ij}^{ab}\rangle_{\mathrm{C}}$$

$$= \overline{g}_{ij}^{ab} + \tfrac{1}{2}\lambda_{ij}^{cd}\overline{g}_{cd}^{ab} + \tfrac{1}{2}\lambda_{kl}^{ab}\overline{g}_{ij}^{kl} + P_{(i/j)}^{(a/b)}\lambda_{ik}^{ac}\overline{g}_{cj}^{kb} - \tfrac{1}{2}P_{(i/j)}\lambda_{ik}^{ab}t_{cd}^{kl}\overline{g}_{jl}^{cd} - \tfrac{1}{2}P^{(a/b)}\lambda_{ij}^{ac}t_{cd}^{kl}\overline{g}_{kl}^{bd}$$
$$+ \tfrac{1}{2^2}\lambda_{ij}^{cd}t_{cd}^{kl}\overline{g}_{kl}^{ab} + P_{(i/j)}^{(a/b)}\lambda_{ik}^{ac}t_{cd}^{kl}\overline{g}_{lj}^{db} + \tfrac{1}{2^2}\overline{g}_{ij}^{cd}t_{cd}^{kl}\lambda_{kl}^{ab} - \tfrac{1}{2}P_{(i/j)}\overline{g}_{ik}^{ab}t_{cd}^{kl}\lambda_{jl}^{cd} - \tfrac{1}{2}P^{(a/b)}\overline{g}_{ij}^{ac}t_{cd}^{kl}\lambda_{kl}^{bd}$$

Remark 7.4. *The Hellmann-Feynman theorem.* Variational methods satisfy the *Hellmann-Feynman theorem*, which says that the derivative of the energy with respect to a perturbation parameter ξ is independent of the first-order wave-function response.[7] This follows from the fact that the parameter gradient vanishes by the stationarity condition.

$$\left.\frac{dE_e(\xi)}{d\xi}\right|_0 = \left.\frac{\partial E_e(\xi)}{\partial\xi}\right|_0 + \frac{\partial E_e}{\partial\mathbf{c}}\cdot\frac{d\mathbf{c}(\xi)}{d\xi}\bigg|_0 = \langle\Psi|H^{(1)}|\Psi\rangle \qquad H^{(1)} \equiv \left.\frac{\partial H(\xi)}{\partial\xi}\right|_0 \tag{7.24}$$

This is not the case for non-variational methods, but the CC Lagrangian satisfies a *generalized Hellmann-Feynman theorem*

$$\left.\frac{d\mathcal{L}(\xi)}{d\xi}\right|_0 = \left.\frac{\partial\mathcal{L}(\xi)}{\partial\xi}\right|_0 + \frac{\partial\mathcal{L}}{\partial\mathbf{t}}\cdot\frac{d\mathbf{t}(\xi)}{d\xi}\bigg|_0 + \frac{\partial\mathcal{L}}{\partial\boldsymbol{\lambda}}\cdot\frac{d\boldsymbol{\lambda}(\xi)}{d\xi}\bigg|_0 = \langle\Phi|(1+\Lambda)H^{(1)}\exp(T)|\Phi\rangle_{\mathrm{C}} \tag{7.25}$$

which follows from the fact that $\mathcal{L}$ is stationary in the amplitudes $\mathbf{t}$ and their corresponding Lagrange multipliers, $\boldsymbol{\lambda}$.

---

[5] $\overline{W} \equiv \exp(-T)W\exp(T)$

[6] The subscript C here denotes that H_{c} is connected both to the ket and to the T operators.

[7] By "first-order wavefunction response", we mean $\Psi^{(1)} = \left.\frac{d\Psi(\xi)}{d\xi}\right|_0 = \frac{\partial\Psi}{\partial\mathbf{c}}\cdot\frac{d\mathbf{c}(\xi)}{d\xi}\big|_0$. Note that this $\Psi^{(1)}$ is generally not the first-order correction to the Møller-Plesset model Hamiltonian. Instead, we are usually interested in the energy derivative with respect to a deviation from the full Born-Oppenheimer electronic Hamiltonian, $H(\xi) = H_e + \sum_{m=1}^\infty \xi^m H^{(m)}$, controlled by a strength parameter ξ. For example, ξ might be a geometric parameter or a strength parameter for an electromagnetic field strength.

Remark 7.5. *Löwdin partitioning.* For a given truncation level m, let us refer to the span of $\boldsymbol{\Phi}_{\rm i} = \begin{bmatrix} \Phi\ \boldsymbol{\Phi}_1\ \cdots\ \boldsymbol{\Phi}_m \end{bmatrix}$ as the *internal space* and that of $\boldsymbol{\Phi}_{\rm e} = \begin{bmatrix} \boldsymbol{\Phi}_{m+1}\ \cdots\ \boldsymbol{\Phi}_n \end{bmatrix}$ as the *external space*, so that $|\boldsymbol{\Phi}_{\rm i}\rangle\langle\boldsymbol{\Phi}_{\rm i}| + |\boldsymbol{\Phi}_{\rm e}\rangle\langle\boldsymbol{\Phi}_{\rm e}| = 1_n$. In the coordinate space over $\boldsymbol{\Phi}$ this reads $\mathbf{1}_{\rm i} + \mathbf{1}_{\rm e} = \mathbf{1}$, in terms of the following projection matrices.

$$\mathbf{1}_{\rm i} \equiv \langle\boldsymbol{\Phi}|\boldsymbol{\Phi}_{\rm i}\rangle\langle\boldsymbol{\Phi}_{\rm i}|\boldsymbol{\Phi}\rangle = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad\qquad \mathbf{1}_{\rm e} \equiv \langle\boldsymbol{\Phi}|\boldsymbol{\Phi}_{\rm e}\rangle\langle\boldsymbol{\Phi}_{\rm e}|\boldsymbol{\Phi}\rangle = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \qquad (7.26)$$

This allows us to write vector decompositions as $\mathbf{c} = \mathbf{c}_{\rm i} + \mathbf{c}_{\rm e}$ and matrix decompositions as $\mathbf{H} = \mathbf{H}_{\rm ii} + \mathbf{H}_{\rm ie} + \mathbf{H}_{\rm ei} + \mathbf{H}_{\rm ee}$ in terms of $\mathbf{c}_{\rm x} \equiv \mathbf{1}_{\rm x}\,\mathbf{c}$ and $\mathbf{H}_{\rm xy} \equiv \mathbf{1}_{\rm x}\,\mathbf{H}\,\mathbf{1}_{\rm y}$.[8] Finally, note that the *external space resolvent* $\mathbf{R}_{\rm ee} \equiv (E - \mathbf{H})^{-1}\big|_{\rm e}$ satisfies

$$\mathbf{R}_{\rm ee}\,(E - \mathbf{H}) = -\mathbf{R}_{\rm ee}\,\mathbf{H}_{\rm ei} + \mathbf{1}_{\rm e} \qquad\qquad (E - \mathbf{H})\,\mathbf{R}_{\rm ee} = -\mathbf{H}_{\rm ie}\,\mathbf{R}_{\rm ee} + \mathbf{1}_{\rm e} \qquad (7.27)$$

and operating the left equation on $\mathbf{c}$ gives zero due to due to the Schrödinger equation, implying that $\mathbf{c}_{\rm e} = \mathbf{R}_{\rm ee}\mathbf{H}_{\rm ei}\mathbf{c}_{\rm i}$. Projecting the Schrödinger equation by $\mathbf{1}_{\rm i}$ and substituting in this expression for $\mathbf{c}_{\rm e}$ then leads to

$$(\mathbf{H}_{\rm ii} + \mathbf{V}_{\rm ii})\mathbf{c}_{\rm i} = E\mathbf{c}_{\rm i} \qquad\qquad \mathbf{V}_{\rm ii} \equiv \mathbf{H}_{\rm ie}\mathbf{R}_{\rm ee}\mathbf{H}_{\rm ei} \qquad (7.28)$$

which reduces the Schrödinger equation on $\mathcal{F}_n$ to an effective Schrödinger equation in the internal space. This gives

$$E = \frac{\mathbf{c}_{\rm i}^{\dagger}(\mathbf{H}_{\rm ii} + \mathbf{V}_{\rm ii})\mathbf{c}_{\rm i}}{\mathbf{c}_{\rm i}^{*}\cdot\mathbf{c}_{\rm i}} \qquad (7.29)$$

which expresses the exact energy in terms of internal-space coefficients. Let us refer to this energy expression as the *Löwdin functional*. The Löwdin functional is the central equation the *Löwdin partitioning* method, which can be used to eliminate the leading error incurred by truncating at a given excitation level $m < n$.

Remark 7.6. *Löwdin partitioning for CC.* The Löwdin functional generalizes to the CC effective Hamiltonian as

$$E = \frac{\mathbf{l}_{\rm i}^{\dagger}(\overline{\mathbf{H}}_{\rm ii} + \overline{\mathbf{V}}_{\rm ii})\mathbf{r}_{\rm i}}{\mathbf{l}_{\rm i}^{*}\cdot\mathbf{r}_{\rm i}} = \langle\Phi|L\overline{H}R|\Phi\rangle + \langle\Phi|L\overline{H}|\boldsymbol{\Phi}_{\rm e}\rangle\langle\boldsymbol{\Phi}_{\rm e}|E - \overline{H}|\boldsymbol{\Phi}_{\rm e}\rangle^{-1}\langle\boldsymbol{\Phi}_{\rm e}|\overline{H}R|\Phi\rangle \qquad (7.30)$$

which leads to the following correction to the ground-state correlation energy.

$$E_{\rm c} - E_{\rm c}^{{\rm CCS}\cdots m} = \langle\Phi|\Lambda\overline{H}_{\rm c}|\boldsymbol{\Phi}_{\rm e}\rangle\langle\boldsymbol{\Phi}_{\rm e}|E_{\rm c} - \overline{H}_{\rm c}|\boldsymbol{\Phi}_{\rm e}\rangle^{-1}\langle\boldsymbol{\Phi}_{\rm e}|\overline{V}_{\rm c}|\Phi\rangle \qquad (7.31)$$

We can identify the leading term in this correction as follows. The leftmost factor requires $\overline{H}_{\rm c}$ to have an excitation level of -1 or less to fully contract Λ_m and $\boldsymbol{\Phi}_{\rm e}$, so its leading contribution is $\overline{H}_{\rm c}^{(1)} = V_{\rm c}$. The leading contribution to the middle factor enters in zeroth order, yielding $-|\boldsymbol{\Phi}_{\rm e}\rangle\langle\boldsymbol{\Phi}_{\rm e}|\overline{H}_{\rm c}^{(0)}|\boldsymbol{\Phi}_{\rm e}\rangle^{-1}\langle\boldsymbol{\Phi}_{\rm e}| = R_0|_{\rm e}$. The rightmost factor requires $\overline{V}_{\rm c}$ to have an excitation level of $m + 1$ in order to fully contract the bra, which first occurs at order $\overline{V}_{\rm c}^{[m]} = (V_{\rm c}\,T_m)_{\rm C}$.[9] This gives

$$E_{(m+1)_{\Lambda}} = \langle\Phi|(\Lambda_{m-1} + \Lambda_m)V_{\rm c}T_{m+1}^{[m]}|\Phi\rangle \qquad\qquad \mathbf{t}_{m+1}^{[m]} = \langle\boldsymbol{\Phi}_{m+1}|R_0V_{\rm c}T_m|\Phi\rangle_{\rm C} \qquad (7.32)$$

where we have made use of the relation $\boldsymbol{\Phi}_k\cdot\mathbf{t}_k = T_k\Phi$. This defines the $(m + 1)_{\Lambda}$ correction to CCS$\cdots m$, generating a hierarchy of perturbatively-corrected CC methods: CCSD(T)$_{\Lambda}$, CCSDT(Q)$_{\Lambda}$, CCSDTQ(P)$_{\Lambda}$, etc.

Example 7.5. *The* (T) *correction.* Since the leading contributions to Λ_1 and Λ_2 are given by

$$^{(1)}\lambda_{ij}^{ab} = \langle\Phi|V_{\rm c}R_0|\Phi_{ij}^{ab}\rangle = {}^{(1)}t_{ab}^{ij*} \qquad\qquad ^{(2)}\lambda_i^a = \langle\Phi|\Lambda_2^{(1)}V_{\rm c}R_0|\Phi_i^a\rangle = {}^{(2)}t_a^{i*}\ , \qquad (7.33)$$

the approximation $\Lambda_1 \approx T_1^{\dagger}, \Lambda_2 \approx T_2^{\dagger}$ will capture the leading terms in the (T)$_{\Lambda}$ correction plus many higher-order ones. The resulting (T) *correction* augments the [T] correction with an order 5^+ contribution, $\langle\Phi|T_1^{\dagger}V_{\rm c}T_3^{[2]}|\Phi\rangle$. Though less rigorous than the (T)$_{\Lambda}$ correction, (T) is computationally advantageous in that it avoids solution of the Λ equations.

$$E_{\rm (T)} = E_{\rm [T]} + \langle\Phi|T_1^{\dagger}V_{\rm c}T_3^{[2]}|\Phi\rangle = \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (7.34)$$

The CCSD(T) method was originally justified on different grounds, but its truly remarkable performance is difficult to understand using ordinary perturbation theory. Several other fifth-order terms could just as easily be included, and there is no strong *a priori* reason to expect this one to be as magical as it is. Löwdin partitioning suggests that $E_{\rm (T)}$ is the leading correction to the error incurred by truncating the configuration space at doubles.

---

[8]Note that I am dropping the subscript e on the Hamiltonian and energy here to avoid confusion with e.
[9]Note that we are assuming $m \geq 2$.