

8 Orbital relaxation

Remark 8.1. *Orbital relaxation.* According to the Thouless theorem (appendix A), the effect of the singles CC operator is to transform the orbitals of the reference determinant into a new set $\{\tilde{\psi}_i\}$ by “mixing in” some of the virtual orbitals.

$$\Psi_{CC} = \exp(T_2 + T_3 + \dots)\tilde{\Phi} \quad \tilde{\Phi} \equiv \exp(T_1)\Phi = \frac{1}{\sqrt{n!}} \det(\tilde{\psi}_1 \dots \tilde{\psi}_n) \quad \tilde{\psi}_i = \psi_i + \sum_a \psi_a t_a^i \quad (8.1)$$

This can be thought of as “relaxing” the orbitals in the presence of electron correlation. The size of this *orbital relaxation effect* can be monitored as the root mean square difference from the reference orbitals, which is known as the \mathcal{T}_1 *diagnostic*.

$$\mathcal{T}_1 \equiv \sqrt{\frac{1}{n} \sum_{i=1}^n \|\tilde{\psi}_i - \psi_i\|^2} = \frac{\|\mathbf{t}_1\|}{\sqrt{n}} \quad (8.2)$$

Significant orbital relaxation generally indicates that the reference determinant forms a poor approximation to the wavefunction, which can lead to large errors for low-order truncated methods like CCSD or CCSD(T). In closed-shell systems, significant orbital relaxation is usually associated with an inherent *multireference character*, which means that no single determinant dominates the wavefunction with any choice of orbitals. Empirically, $\mathcal{T}_1 \geq 0.02$ is considered large in this context. In open-shell systems, mean-field methods like Hartree-Fock theory are often deficient even for non-multireference systems. In this case, orbital relaxation effects can generally be cured by choosing a new determinant which is optimized in the presence of dynamical¹ electron correlation.

Remark 8.2. *Brueckner and orbital-optimized methods.* The two most common ways of defining an ideal reference determinant for the correlated wavefunction are the *best overlap criterion* and the *best energy criterion*.²

$$\Phi_B = \arg \max_{\Phi} \langle \Phi | \Psi \rangle \quad \Phi_O = \arg \min_{\Phi} \langle \Psi | H | \Psi \rangle, \quad \Psi = \Omega \Phi \quad (8.3)$$

The *best overlap* or *Brueckner determinant*, Φ_B , has maximum overlap with the wavefunction. The *best energy* or *orbital-optimized determinant*, Φ_O , yields the lowest energy for a given Ansatz. The spin-orbital sets from which these determinants are constructed will be denoted $\{\psi_p\}_B$ and $\{\psi_p\}_O$.

Remark 8.3. Appendix B shows that non-redundant transformations of the orbitals in Φ can be parametrized as

$$\Phi(\mathbf{x}) = \exp(X - X^\dagger)\Phi \quad (8.4)$$

where X has the form of a T_1 operator. This parametrization can be substituted into equation 8.3 to derive the explicit conditions satisfied by Brueckner orbitals and energy-optimized orbitals.³

$$\{\psi_p\}_B : \left. \frac{\partial}{\partial x_a^{i*}} \langle \Phi(\mathbf{x}) | \Psi \rangle \right|_{\mathbf{x}=\mathbf{0}} = \langle \Phi_i^a | \Psi \rangle \stackrel{!}{=} 0 \quad \{\psi_p\}_O : \left. \frac{\partial}{\partial x_a^{i*}} \langle \Psi(\mathbf{x}) | H | \Psi(\mathbf{x}) \rangle \right|_{\mathbf{x}=\mathbf{0}} = \langle \Psi | [a_a^i, H] | \Psi \rangle \stackrel{!}{=} 0$$

¹As opposed to mean-field.

²See https://en.wikipedia.org/wiki/Arg_max for the notation used here.

³ $|\Psi(\mathbf{x})\rangle = \exp(X - X^\dagger)|\Psi\rangle$ applies the orbital rotation to all of the determinants in the wavefunction expansion.

A The Thouless theorem

Notation A.1. Let ψ be a row vector of orthonormal spin-orbitals, $(\psi)_p = \psi_p$, which can be split into occupied and virtual blocks as $\psi = [\psi_o \ \psi_v]$. Other sets of spin-orbitals are related to this one by a transformation $|\psi'\rangle = |\psi\rangle \mathbf{U}$ which is unitary if the primed orbitals are also orthonormal. Let Φ' be the *transformed reference determinant*, constructed from the first n orbitals in the transformed space. Then the occupied and virtual orbitals of the transformed space are given by

$$|\psi'_o\rangle = |\psi_o\rangle \mathbf{U}_{oo} + |\psi_v\rangle \mathbf{U}_{vo} \quad |\psi'_v\rangle = |\psi_o\rangle \mathbf{U}_{ov} + |\psi_v\rangle \mathbf{U}_{vv} \quad (\text{A.1})$$

in terms of the occupied and virtual blocks of the transformation. This kind of unitary transformation of the spin-orbital basis is sometimes referred to as an *orbital rotation*.

Theorem A.1. The Thouless theorem.

1. The function $e^{T_1} \Phi$ is an intermediately normalized determinant $\frac{1}{\sqrt{n!}} \det(\tilde{\psi}_1 \cdots \tilde{\psi}_n)$ with orbitals $\tilde{\psi}_i = \psi_i + \sum_a \psi_a t_a^i$.

Proof: Intermediate normalization follows from $\langle \Phi | e^{T_1} \Phi \rangle = 1$. This function has the form of a determinant

$$e^{T_1} |\Phi\rangle = e^{\sum_a t_a^1 a_1^a + \cdots + \sum_a t_a^n a_n^a} a_1^\dagger \cdots a_n^\dagger |\text{vac}\rangle = \tilde{a}_1^\dagger \cdots \tilde{a}_n^\dagger |\text{vac}\rangle = |\tilde{\Phi}\rangle \quad \tilde{a}_i^\dagger \equiv \exp(\sum_a t_a^i a_i^a) a_i^\dagger$$

since $\sum_a t_a^i a_i^a$ commutes with all creation operators except a_i^\dagger . The transformed orbitals are given by

$$|\tilde{\psi}_i\rangle = \tilde{a}_i^\dagger |\text{vac}\rangle = \exp(\sum_a t_a^i a_a^\dagger a_i) a_i^\dagger |\text{vac}\rangle = (1 + \sum_a t_a^i a_a^\dagger a_i) a_i^\dagger |\text{vac}\rangle = |\psi_i\rangle + \sum_a t_a^i |\psi_a\rangle$$

using $a_i^2 = 0$ and $a_i a_i^\dagger |\text{vac}\rangle = |\text{vac}\rangle$.

2. Any intermediately normalized determinant $\tilde{\Phi} = \frac{1}{\sqrt{n!}} \det(\tilde{\psi}_1 \cdots \tilde{\psi}_n)$ can be written as $e^{T_1} \Phi$.

Proof: Intermediate normalization is only possible if $\tilde{\Phi}$ has non-zero overlap with the reference determinant. Therefore, $\tilde{\Phi}$ can be written as $\Phi' / \langle \Phi | \Phi' \rangle$ where Φ' is a Slater determinant. The normalization factor is given by

$$\langle \Phi | \Phi' \rangle = \frac{1}{n!} \sum_{\pi, \sigma} \varepsilon_\pi \varepsilon_\sigma \langle \psi_{\pi(1)} | \psi'_{\sigma(1)} \rangle \cdots \langle \psi_{\pi(n)} | \psi'_{\sigma(n)} \rangle = \sum_{\sigma} \varepsilon_\sigma \langle \psi_1 | \psi'_{\sigma(1)} \rangle \cdots \langle \psi_n | \psi'_{\sigma(n)} \rangle = \det(\mathbf{U}_{oo})$$

following notation A.1. Therefore, $\tilde{\Phi} = \Phi' / \det(\mathbf{U}_{oo}) = \Phi' \det(\mathbf{U}_{oo}^{-1})$ and the rows of $\tilde{\Phi}$ are given by

$$|\tilde{\psi}_o\rangle = |\psi'_o\rangle \mathbf{U}_{oo}^{-1} = |\psi_o\rangle + |\psi_v\rangle \mathbf{U}_{vo} \mathbf{U}_{oo}^{-1}$$

where we have expanded $|\psi'_o\rangle$ according to eq A.1. The columns of this equation are $\tilde{\psi}_i = \psi_i + \sum_a \psi_a (\mathbf{U}_{vo} \mathbf{U}_{oo}^{-1})_{ai}$. Referring back to the first proposition, this shows that $\tilde{\Phi} = e^{T_1} \Phi$ with $t_a^i = (\mathbf{U}_{vo} \mathbf{U}_{oo}^{-1})_{ai}$.

B Orbital rotations

Definition B.1. Normal matrix. A square matrix satisfying $\mathbf{N}^\dagger \mathbf{N} = \mathbf{N} \mathbf{N}^\dagger$ is termed *normal*. Several important kinds of matrices meet this criterion: *Hermitian matrices*, $\mathbf{H}^\dagger = \mathbf{H}$; *anti-Hermitian matrices*, $\mathbf{A}^\dagger = -\mathbf{A}$; and *unitary matrices*, $\mathbf{U}^\dagger = \mathbf{U}^{-1}$. Note that Hermitian and anti-Hermitian matrices can always be written as $\mathbf{X} + \mathbf{X}^\dagger$ and $\mathbf{X} - \mathbf{X}^\dagger$.

Remark B.1. The Spectral Theorem⁴ says that every normal matrix can be diagonalized by a unitary transformation. A direct corollary⁵ is that the eigenvalues of Hermitian, anti-Hermitian, and unitary matrices can be written as follows.

$$h^* = h \implies h = \phi \quad a^* = -a \implies a = i\phi \quad u^* = u^{-1} \implies u = e^{i\phi} \quad \phi \in \mathbb{R} \quad (\text{B.1})$$

In words, Hermitian eigenvalues are real, anti-Hermitian eigenvalues are pure imaginary, and unitary eigenvalues lie on the unit circle. Note that unitary eigenvalues have the form $u = \exp(a)$ where a is an anti-Hermitian eigenvalue. This implies that any unitary matrix \mathbf{U} can be written as $\exp(\mathbf{A})$, where \mathbf{A} is anti-Hermitian.

Remark B.2. According to def B.1 and rmk B.1, unitary transformations of the spin-orbitals can be parametrized as

$$\psi'_p = \sum_q \psi_q (\exp(\mathbf{X} - \mathbf{X}^\dagger))_{qp} \quad (\text{B.2})$$

in terms a square matrix \mathbf{X} . The anti-Hermitian form of this parametrization leads to redundancies. In particular, notice that $\mathbf{X} = [z \delta_{pq}]$ generates the same transformation as $\mathbf{X}^\dagger = [-z^* \delta_{qp}]$. These redundancies are eliminated by setting the upper or lower triangle of \mathbf{X} to zero. For most single-reference methods, transformations within the occupied and virtual blocks are also redundant because they don't change the energy.⁶ Setting all but one of the off-diagonal blocks to zero yields a non-redundant parametrization. Here, we will choose $\mathbf{X}_{\text{vo}} = [x_a^i]$.

Proposition B.1. The identity $\exp(G) a_p^\dagger \exp(-G) = \sum_q a_q^\dagger (\exp(\mathbf{G}))_{qp}$ holds for any $G = \sum_{pq} (\mathbf{G})_{pq} a_p^\dagger a_q$.

Proof: This follows from $[G, \cdot]^m (a_p^\dagger) = \sum_q a_q^\dagger (\mathbf{G}^m)_{qp}$, which we will prove by induction. For $m = 0$ the statement is trivially true. If we assume it holds for m , then the following shows that it also holds for $m + 1$,⁷

$$[G, \cdot]^{m+1} (a_p^\dagger) = \sum_q [G, a_q^\dagger] (\mathbf{G}^m)_{qp} = \sum_{qr} a_r^\dagger (\mathbf{G})_{rq} (\mathbf{G}^m)_{qp} = \sum_r a_r^\dagger (\mathbf{G}^{m+1})_{rp}$$

which completes the induction. Substituting this result into the Hausdorff expansion of $\exp(G) a_p^\dagger \exp(-G)$ recognizing the Taylor expansion of $\exp(\mathbf{G})$ completes the proof.

Remark B.3. The identity $a_p^\dagger |\text{vac}\rangle = |\psi_p\rangle$ implies that creation operators transform like orbitals. Therefore, the creation operator corresponding to ψ'_p in equation B.2 is given by $a_p'^\dagger = \sum_q a_q^\dagger (\exp(\mathbf{X} - \mathbf{X}^\dagger))_{qp}$. By prop B.1, this is equivalent to

$$a_p'^\dagger = \exp(X - X^\dagger) a_p^\dagger \exp(X^\dagger - X) \quad X \equiv \sum_{ai} x_a^i a_i^a \quad (\text{B.3})$$

where we have eliminated the redundant parameters. Substituting this into $|\Phi'_{(p_1 \dots p_n)}\rangle = a_{p_1}'^\dagger \dots a_{p_n}'^\dagger |\text{vac}\rangle$ gives a convenient expression for determinants of the transformed orbitals.⁸

$$|\Phi'_{(p_1 \dots p_n)}\rangle = \exp(X - X^\dagger) |\Phi_{(p_1 \dots p_n)}\rangle \quad \langle \Phi'_{(p_1 \dots p_n)} | = \langle \Phi_{(p_1 \dots p_n)} | \exp(X^\dagger - X) \quad (\text{B.4})$$

⁴See https://en.wikipedia.org/wiki/Spectral_theorem

⁵Since there exists a basis in which \mathbf{N} is diagonal, statements about \mathbf{N} translate into statements about its eigenvalues.

⁶This is what allows us to diagonalize the Fock matrix in canonical Hartree-Fock theory.

⁷The second equality here follows from expanding G and using $[a_p^\dagger a_s, a_q^\dagger] = :a_r^\dagger \overline{a_s} a_q^\dagger: = a_r^\dagger \delta_{sq}$.

⁸This follows from $\exp(X^\dagger - X) \exp(X - X^\dagger) = 1$ and $\exp(X^\dagger - X) |\text{vac}\rangle = 0$.