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 $\{\widetilde{\psi}_i\}$ by mixing in some of the virtual orbitals.

$$\Psi_{\text{CC}} = \exp(T_2 + T_3 + \cdots)\widetilde{\Phi} \qquad \widetilde{\Phi} \equiv \exp(T_1)\Phi = \tfrac{1}{\sqrt{n!}}\det(\widetilde{\psi}_1 \cdots \widetilde{\psi}_n) \qquad \widetilde{\psi}_i = \psi_i + \sum_a \psi_a t_a^i \qquad (8.1)$$

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*.

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

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 for any choice of orbitals. Empirically, $\mathcal{T}_1 \geq 0.02$ is considered large for closed-shell species. 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 dynamical[1] electron correlation.

Remark 8.2. *Brueckner orbitals and optimized orbitals.* 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]

$$\{\psi_p\}_{\text{B}} = \underset{\{\psi_p\}}{\arg\max}\,\|\langle\Phi|\Psi\rangle\|^2 \qquad\qquad \{\psi_p\}_{\text{O}} = \underset{\{\psi_p\}}{\arg\min}\,\langle\Psi|H|\Psi\rangle \qquad (8.3)$$

The *best overlap* or *Brueckner orbitals* yield a reference determinant, Φ_{B}, that has maximum overlap with the wavefunction. The *best energy* or *optimized orbitals* yield the lowest energy expectation value within a given Ansatz. Note that these two conditions are in general mutually exclusive, so one must choose one or the other.

Definition 8.1. *Orbital invariance.* An Ansatz is usually termed *orbital invariant* if it preserves the unitary invariances of its reference function. For single-reference methods, this means that the total energy is unchanged by orbital rotations within the occupied and virtual spaces. In this context, the occupied and virtual blocks of $\mathbf{X}$ in equation B.3 are redundant in the sense that they produce transformations which do not change the energy.

Remark 8.3. The non-redundant orbital rotations (appendix B) for an orbital-invariant Ansatz can be parametrized as

$$\Theta(\mathbf{x}) = U(\mathbf{x})\Theta \qquad\qquad U(\mathbf{x}) = \exp(X - X^\dagger) \qquad\qquad X = \sum_{ia} x_a^i a_i^a \qquad (8.4)$$

where Θ is an arbitrary Fock state. Brueckner orbitals and optimized orbitals can then be determined from the following.[3]

$$\frac{\partial}{\partial x_a^{i*}}\|\langle\Phi(\mathbf{x})|\Psi\rangle\|^2\Big|_{\mathbf{x}=0} = \langle\Phi|\,a_a^i\,|\Psi\rangle\langle\Psi|\Phi\rangle \qquad\qquad \frac{\partial}{\partial x_a^{i*}}\langle\Psi(\mathbf{x})|H|\Psi(\mathbf{x})\rangle\Big|_{\mathbf{x}=0} = \langle\Psi|a_a^i H|\Psi\rangle - \langle\Psi|Ha_a^i|\Psi\rangle \qquad (8.5)$$

In the overlap derivative, we have allowed the orbitals of the determinant to vary while holding the wavefunction constant. In the energy derivative, we have allowed the wavefunction to vary while holding the orbitals of the Hamiltonian constant.[4] The best overlap and best energy orbitals make these derivatives vanish, which leads to a new set of conditions

$$\{\psi_p\}_{\text{B}} \;:\; \langle\Phi_i^a|\Psi\rangle \overset{!}{=} 0 \qquad\qquad \{\psi_p\}_{\text{O}} \;:\; \langle\Psi|[a_a^i, H]|\Psi\rangle \overset{!}{=} 0 \qquad (8.6)$$

which are equivalent to equation 8.3. These conditions can be used to iteratively determine best overlap or best energy orbitals for any wavefunction Ansatz. Since orbital rotations generate singly-excited determinants, one generally excludes singles from the wavefunction parametrization in methods like *orbital-optimized coupled cluster* (OCC) theory.[5] For Brueckner methods, the best overlap condition is equivalent to determining the orbitals which make the singles coefficients vanish entirely, which means $T_1 = 0$ for a converged *Brueckner coupled-cluster* (BCC) wavefunction.

---

[1] As opposed to mean-field.

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

[3] Note that $a_a^i|\Phi\rangle = 0$, which eliminates the second term in the overlap derivative.

[4] In each case we could have made the opposite choice, but note that we do not want to transform *everything*. Transforming everything corresponds to a Fock space isomorphism rather than a transformation of orbitals used to construct the wavefunction. By definition, an isomorphism would leave these matrix elements unchanged: $\langle\Phi(\mathbf{x})|\Psi(\mathbf{x})\rangle = \langle\Phi|\Psi\rangle$ and $\langle\Psi(\mathbf{x})|H(\mathbf{x})|\Psi(\mathbf{x})\rangle = \langle\Psi|H|\Psi\rangle$.

[5] Caveat: This is actually slightly problematic for TCC because it turns out that OCCD$\cdots m$ is not equivalent to full CI for an m-electron system. This is a quirk of the way amplitudes are determined in TCC and is not an issue when the wavefunction parametrization is determined from a stationarity condition, such as for orbital-optimized configuration interaction.

Remark 8.4. One method of determining Brueckner orbitals is to set the non-redundant orbital rotation parameters equal to the T_1 amplitudes each iteration. To see how this works, consider the Taylor expansion of $|\psi(\mathbf{x})\rangle = |\psi\rangle \exp(\mathbf{X} - \mathbf{X}^\dagger)$.[6]

$$|\psi(\mathbf{x})\rangle = |\psi_\mathrm{v}\rangle \mathbf{X}_\mathrm{vo} - |\psi_\mathrm{o}\rangle \mathbf{X}_\mathrm{vo}^\dagger + \mathcal{O}(\mathbf{x}^2) \tag{8.7}$$

The columns of this equation are as follows

$$\psi_i(\mathbf{x}) = \psi_i + \sum_a \psi_a \, x_a^i + \mathcal{O}(\mathbf{x}^2) \qquad\qquad \psi_a(\mathbf{x}) = \psi_a - \sum_i \psi_i \, x_a^{i*} + \mathcal{O}(\mathbf{x}^2) \tag{8.8}$$

which implies $\Phi(\mathbf{x})|_{x_a^i = t_a^i} \approx \exp(T_1)\Phi$ by the Thouless theorem. Repeatedly absorbing this factor into the transformed determinant causes the singles amplitudes to vanish at convergence, so that equation 8.6 is satisfied.

Remark 8.5. *BCC algorithm.* The discussion of rmk 8.4 leads to the following algorithm for BCC methods.

1. Update non-singles amplitudes: $t_{ab\cdots}^{ij\cdots} = (\mathcal{E}_{ab\cdots}^{ij\cdots})^{-1} \langle \Phi_{ij\cdots}^{ab\cdots} | V_\mathrm{c} \exp(T) | \Phi \rangle_\mathrm{C}$.

2. Update singles amplitudes: $t_a^i = (\mathcal{E}_a^i)^{-1} \langle \Phi_i^a | V_\mathrm{c} \exp(T) | \Phi \rangle_\mathrm{C}$.

3. Build an orbital rotation matrix from the singles amplitudes: $\mathbf{U} = \exp(\mathbf{X} - \mathbf{X}^\dagger)$ where $\mathbf{X}_\mathrm{vo} = [t_a^i]$.

4. Rotate the spin-orbital coefficients:[7] $\mathbf{C} \leftarrow \mathbf{C}\mathbf{U}$.

5. Transform the one- and two-electon integrals to the spin-orbital basis using the new coefficient matrix.

6. Unless $\|\mathbf{t}_1\| \approx 0$ and the non-singles amplitudes are converged, return to step 1.

Note that the amplitude equations in steps 1 and 2 here are non-canonical CCSD$\cdots m$ equations.

Remark 8.6. *The Newton-Raphson step for orbital rotations.* Optimized orbitals can be determined from

$$E(\mathbf{x}) - E(\mathbf{0}) = \mathbf{x}^\dagger \mathbf{w} + \tfrac{1}{2} \mathbf{x}^\dagger \mathbf{A} \mathbf{x} + \mathcal{O}(\mathbf{x}^3) \qquad (\mathbf{w})_{ia} \equiv \left.\frac{\partial E}{\partial x_a^{i*}}\right|_{\mathbf{x}=\mathbf{0}} \qquad (\mathbf{A})_{ia,jb} \equiv \left.\frac{\partial^2 E}{\partial x_a^{i*} \partial x_b^j}\right|_{\mathbf{x}=\mathbf{0}} \tag{8.9}$$

which is a Taylor expansion for the total energy with respect to orbital rotations, $\mathbf{w}$ being the *orbital gradient* and $\mathbf{A}$ the *orbital Hessian* evaluated at the current orbitals. These are given by

$$(\mathbf{w})_{ia} = \langle \Psi | [a_a^i, H] | \Psi \rangle \qquad\qquad (\mathbf{A})_{ia,jb} = \langle \Psi | [[a_a^i, H], a_j^b] | \Psi \rangle \tag{8.10}$$

where the gradient was already derived in equation 8.5 and the Hessian can be derived similarly. The step

$$\mathbf{x} = -\mathbf{A}^{-1} \mathbf{w} \tag{8.11}$$

rotates the orbitals to make the gradient vanish, assuming $E(\mathbf{x})$ is quadratic.[8] This known as the *Newton-Raphson step*. Note that the expectation values in equation 8.10 will require us to solve the lambda equations in order to determine $\langle \Psi |$.

Remark 8.7. *OCC algorithm.* The discussion of rmk 8.6 leads to the following algorithm for OCC methods.

1. Update non-singles amplitudes: $t_{ab\cdots}^{ij\cdots} = (\mathcal{E}_{ab\cdots}^{ij\cdots})^{-1} \langle \Phi_{ij\cdots}^{ab\cdots} | V_\mathrm{c} \exp(T) | \Phi \rangle_\mathrm{C}$.

2. Update the Lagrange multipliers: $\lambda_{ij\cdots}^{ab\cdots} = (\mathcal{E}_{ab\cdots}^{ij\cdots})^{-1} \langle \Phi | (1 + \Lambda) V_\mathrm{c} \exp(T) | \Phi_{ij\cdots}^{ab\cdots} \rangle_\mathrm{C}$.

3. Compute the Newton-Raphson step: $x_a^i = (-\mathbf{A}^{-1} \mathbf{w})_{ia}$

4. Build the Newton-Raphson orbital rotation matrix: $\mathbf{U} = \exp(\mathbf{X} - \mathbf{X}^\dagger)$ where $\mathbf{X}_\mathrm{vo} = [x_a^i]$.

5. Rotate the spin-orbital coefficients: $\mathbf{C} \leftarrow \mathbf{C}\mathbf{U}$.

6. Transform the one- and two-electron integrals to the spin-orbital basis using the new coefficient matrix.

7. Unless $\|\mathbf{w}\| \approx 0$ and the non-singles amplitudes are converged, return to step 1.

Note that the amplitude and lambda equations in steps 1 and 2 here are non-canonical CCD$\cdots m$ equations.

---

[6]This is equation B.2. Note that $|\psi\rangle \mathbf{X} = |\psi_\mathrm{v}\rangle \mathbf{X}_\mathrm{vo}$ and $|\psi\rangle \mathbf{X}^\dagger = |\psi_\mathrm{o}\rangle \mathbf{X}_\mathrm{vo}^\dagger$ since the other blocks of $\mathbf{X}$ are zero.

[7]If χ is a row vector of basis functions, the spin-orbitals are obtained from this coefficient matrix as $\psi = \chi \mathbf{C}$.

[8]To see this, solve for $\frac{\partial E(\mathbf{x})}{\partial \mathbf{x}} \overset{!}{=} \mathbf{0}$ in equation 8.9.

Definition 8.2. *Density matrices.* The energy expectation value of the Hamiltonian, $H = h_p^q a_q^p + \frac{1}{4}\bar{g}_{rs}^{pq} a_{pq}^{rs}$, can be expressed in terms of the *one-particle* and *two-particle density matrices* of the wavefunction

$$E = \langle\Psi|H|\Psi\rangle = h_p^q \gamma_q^p + \tfrac{1}{4}\bar{g}_{pq}^{rs}\gamma_{rs}^{pq} \qquad \gamma_q^p = \langle\Psi|a_q^p|\Psi\rangle \qquad \gamma_{rs}^{pq} = \langle\Psi|a_{rs}^{pq}|\Psi\rangle \qquad (8.12)$$

which generalize the one-particle reference density matrix that we have already encountered. In traditional coupled-cluster theory this expectation value is the CC Lagrangian, whose density matrices are the *CC response density matrices*.

$$\mathcal{L} = \langle\Phi|(1+\Lambda)H\exp(T)|\Phi\rangle_\mathrm{C} = h_p^q\gamma_q^p + \tfrac{1}{4}\bar{g}_{pq}^{rs}\gamma_{rs}^{pq} \qquad \begin{aligned}\gamma_q^p &= \langle\Phi|(1+\Lambda)\,a_q^p\exp(T)|\Phi\rangle_\mathrm{C} \\ \gamma_{rs}^{pq} &= \langle\Phi|(1+\Lambda)\,a_{rs}^{pq}\exp(T)|\Phi\rangle_\mathrm{C}\end{aligned} \qquad (8.13)$$

For Hermitian methods, $\gamma_q^p = \gamma_p^{q*}$ and $\gamma_{rs}^{pq} = \gamma_{pq}^{rs*}$. This is not the case for TCC.

Remark 8.8. Evaluating the commutators of equation 8.10 using Wick's theorem[9]

$$[a_a^i, a_q^p] = \,:a_a^i a_q^{p^\bullet}: - :a_{q^\bullet}^p a_a^{i^\bullet}: = \delta_q^p\,a_a^i - \delta_i^q\,a_a^p \qquad (8.14)$$

$$[a_a^i, a_{rs}^{pq}] = P^{(p/q)}:a_a^i a_r^{p^\bullet}{}_s^q: - P^{(r/s)}:a_{r^\bullet}^p{}_s^q a_a^{i^\bullet}: = P^{(p/q)}\delta_a^p a_{rs}^{iq} - P_{(r/s)}\delta_r^i a_{as}^{pq} \qquad (8.15)$$

yields an expression for the orbital gradient in terms of density matrices.

$$(\mathbf{w})_{ia} = \langle\Psi|[a_a^i, H]|\Psi\rangle = (\mathbf{F}-\mathbf{F}^\dagger)_a^i \qquad (\mathbf{F})_p^q \equiv h_p^r\gamma_r^q + \tfrac{1}{2}\bar{g}_{pr}^{st}\gamma_{st}^{qr} \qquad (8.16)$$

The intermediate $\mathbf{F}$ is sometimes called a *generalized Fock matrix*.

Remark 8.9. The orbital Hessian can also be expanded in terms of one- and two-particle density matrices, but this expression is rather complicated. A more common approach is to use its zeroth order approximation.

$$(\mathbf{A})_{ia,jb} \approx \langle\Psi|[[a_a^i, H], a_j^b]|\Psi\rangle^{(0)} = \langle\Phi|[[a_a^i, H_0], a_j^b]|\Phi\rangle = \langle\Phi_i^a|H_0|\Phi_j^b\rangle = \mathcal{E}_j^b\delta_j^i\delta_a^b \qquad (8.17)$$

This approximation has the additional advantage of being diagonal and hence trivial to invert. This leads to the following formula for the Newton-Raphson step in terms of orbital energies, integrals, and density matrices.

$$(-\mathbf{A}^{-1}\,\mathbf{w})_{ia} = \sum_{jb}(-\mathbf{A}^{-1})_{ia,jb}(\mathbf{w})_{jb} \approx -\sum_{jb}\frac{\delta_j^i\delta_a^b}{\mathcal{E}_j^b}(\mathbf{F}-\mathbf{F}^\dagger)_b^j = \frac{(\mathbf{F}-\mathbf{F}^\dagger)_a^i}{\mathcal{E}_a^i} \qquad (8.18)$$

**Remark 8.10.** Expanding excitation operators of equation 8.12 using Wick's theorem gives

$$\gamma_q^p = \tilde{\gamma}_q^p + \mathring{\gamma}_q^p \qquad \gamma_{rs}^{pq} = \tilde{\gamma}_{rs}^{pq} + P_{(r/s)}^{(p/q)}\tilde{\gamma}_r^p\mathring{\gamma}_s^q + P_{(r/s)}\mathring{\gamma}_r^p\mathring{\gamma}_s^q \qquad \begin{aligned}\tilde{\gamma}_q^p &= \langle\Psi|\tilde{a}_q^p|\Psi\rangle \\ \tilde{\gamma}_{rs}^{pq} &= \langle\Psi|\tilde{a}_{rs}^{pq}|\Psi\rangle\end{aligned} \qquad (8.19)$$

where $\mathring{\gamma}_q^p \equiv \langle\Phi|a_q^p|\Phi\rangle$ is the density matrix of the reference determinant, whose non-vanishing elements are $\mathring{\gamma}_j^i = \delta_j^i$ as usual as long as Φ is constructed from the current spin-orbital basis. For CCD, the correlation contributions are given by

$$\tilde{\gamma}_q^p = \langle\Phi|(1+\Lambda_2)\,\tilde{a}_q^p\exp(T_2)|\Phi\rangle_\mathrm{C} \qquad \tilde{\gamma}_{rs}^{pq} = \langle\Phi|(1+\Lambda_2)\,\tilde{a}_{rs}^{pq}\exp(T_2)|\Phi\rangle_\mathrm{C} \qquad (8.20)$$

which can be evaluated diagrammatically.
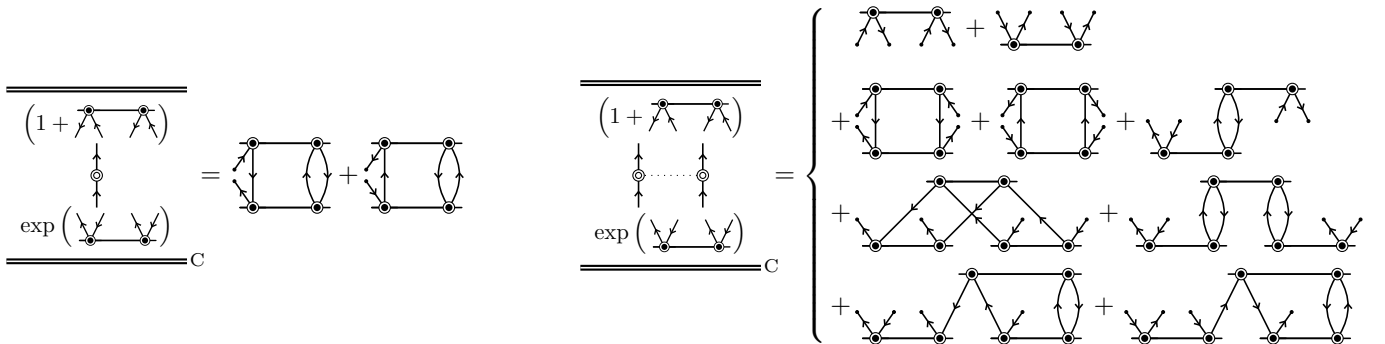
Separating these diagrams into unique blocks yields the following.

$$\tilde{\gamma}_b^a = \tfrac{1}{2}\lambda_{ij}^{ac}t_{bc}^{ij} \qquad \tilde{\gamma}_j^i = -\tfrac{1}{2}\lambda_{jk}^{ab}t_{ab}^{ik} \qquad \tilde{\gamma}_{cd}^{ab} = \tfrac{1}{2}\lambda_{ij}^{ab}t_{cd}^{ij} \qquad \tilde{\gamma}_{kl}^{ij} = \tfrac{1}{2}\lambda_{kl}^{ab}t_{ab}^{ij} \qquad \tilde{\gamma}_{ib}^{aj} = \lambda_{ik}^{ac}t_{cb}^{kj} \qquad (8.21)$$

$$\tilde{\gamma}_{ij}^{ab} = \lambda_{ij}^{ab} \qquad \tilde{\gamma}_{ab}^{ij} = t_{ab}^{ij} + (\tfrac{1}{2})^2\,t_{ab}^{kl}\lambda_{kl}^{cd}t_{cd}^{ij} + \tfrac{1}{2}P_{(a/b)}^{(i/j)}t_{ac}^{ik}\lambda_{kl}^{cd}t_{cb}^{lj} - \tfrac{1}{2}P^{(i/j)}t_{ab}^{ik}\lambda_{kl}^{cd}t_{cd}^{jl} - \tfrac{1}{2}P_{(a/b)}t_{ac}^{ij}\lambda_{kl}^{cd}t_{bd}^{kl} \qquad (8.22)$$

---

[9]Note that here we are doing Wick's theorem with respect to the physical vacuum.

A   The Thouless theorem

Notation A.1. Let $\boldsymbol{\psi} = [\psi_p]$ be a row vector of orthonormal spin-orbitals, composed of occupied and virtual blocks, $\boldsymbol{\psi} = [\boldsymbol{\psi}_{\mathrm{o}}\ \boldsymbol{\psi}_{\mathrm{v}}]$, with respect to a reference determinant, Φ. Other spin-orbital bases relate to this one via $|\boldsymbol{\psi}'\rangle = |\boldsymbol{\psi}\rangle\,\mathbf{U}$, which is a unitary transformation if the primed orbitals are orthonormal. Let Φ' be the *transformed reference determinant*, constructed from the first n orbitals in $\boldsymbol{\psi}'$. Then the occupied and virtual orbitals of the transformed space are given by

$$|\boldsymbol{\psi}'_{\mathrm{o}}\rangle = |\boldsymbol{\psi}_{\mathrm{o}}\rangle\,\mathbf{U}_{\mathrm{oo}} + |\boldsymbol{\psi}_{\mathrm{v}}\rangle\,\mathbf{U}_{\mathrm{vo}} \qquad\qquad |\boldsymbol{\psi}'_{\mathrm{v}}\rangle = |\boldsymbol{\psi}_{\mathrm{o}}\rangle\,\mathbf{U}_{\mathrm{ov}} + |\boldsymbol{\psi}_{\mathrm{v}}\rangle\,\mathbf{U}_{\mathrm{vv}} \qquad\text{(A.1)}$$

in terms of the occupied and virtual blocks of $\mathbf{U}$. This kind of transformation is sometimes called an *orbital rotation*.

Theorem A.1. *The Thouless theorem.*

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

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

   $$e^{T_1}|\Phi\rangle = e^{\sum_a t_a^1 a_1^a + \cdots + \sum_a t_a^n a_n^a}\,a_1^\dagger\cdots a_n^\dagger|\mathrm{vac}\rangle = \widetilde{a}_1^\dagger\cdots\widetilde{a}_n^\dagger|\mathrm{vac}\rangle = |\widetilde{\Phi}\rangle \qquad \widetilde{a}_i^\dagger \equiv \exp(\textstyle\sum_a t_a^i a_i^a)\,a_i^\dagger$$

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

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

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

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

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

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

   Therefore, $\widetilde{\Phi} = \Phi'/\det(\mathbf{U}_{\mathrm{oo}}) = \Phi'\det(\mathbf{U}_{\mathrm{oo}}^{-1})$ and the rows of $\widetilde{\Phi}$ are given by the following vector[10]

   $$|\widetilde{\boldsymbol{\psi}}_{\mathrm{o}}\rangle = |\boldsymbol{\psi}'_{\mathrm{o}}\rangle\,\mathbf{U}_{\mathrm{oo}}^{-1} = |\boldsymbol{\psi}_{\mathrm{o}}\rangle + |\boldsymbol{\psi}_{\mathrm{v}}\rangle\,\mathbf{U}_{\mathrm{vo}}\mathbf{U}_{\mathrm{oo}}^{-1}$$

   with elements $\widetilde{\psi}_i = \psi_i + \sum_a \psi_a\,(\mathbf{U}_{\mathrm{vo}}\mathbf{U}_{\mathrm{oo}}^{-1})_{ai}$. Referring back to part one, $\widetilde{\Phi} = e^{T_1}\Phi$ with $t_a^i = (\mathbf{U}_{\mathrm{vo}}\mathbf{U}_{\mathrm{oo}}^{-1})_{ai}$.

---

[10]The second equality follows from expanding $|\boldsymbol{\psi}'_{\mathrm{o}}\rangle$ according to eq A.1.

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 theorem[11] for normal matrices says that $\mathbf{N} = \mathbf{V}\widetilde{\mathbf{N}}\mathbf{V}^\dagger$ where $\mathbf{V}$ is unitary and $\widetilde{\mathbf{N}}$ is diagonal. A direct corollary[12] is that the eigenvalues of Hermitian, anti-Hermitian, and unitary matrices can be written as follows.

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

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

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

in terms of a square matrix $\mathbf{X}$. The form of this parametrization leads to redundancies. In particular, notice that $(\mathbf{X})_{pq} \equiv z\,\delta_{pp'}\delta_{qq'}$ generates the same transformation as $(\mathbf{X})_{pq} \equiv -z^*\,\delta_{qp'}\delta_{pq'}$. These redundancies are eliminated by setting the upper or lower triangle of $\mathbf{X}$ to zero. The creation operators for these orbitals are given by $a'^\dagger_p = \sum_q a^\dagger_q (\exp(\mathbf{X}-\mathbf{X}^\dagger))_{qp}$.

Proposition B.1. *The identity* $\exp(G)\,a^\dagger_p\exp(-G) = \sum_q a^\dagger_q\,(\exp(\mathbf{G}))_{qp}$ *holds for any* $G = \sum_{pq}(\mathbf{G})_{pq}\,a^\dagger_p a_q$.

Proof: This follows from $[G,\cdot]^m(a^\dagger_p) = \sum_q a^\dagger_q(\mathbf{G}^m)_{qp}$, 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 $m+1$,[13]

$$[G,\cdot]^{m+1}(a^\dagger_p) = \sum_q [G, a^\dagger_q]\,(\mathbf{G}^m)_{qp} = \sum_{qr} a^\dagger_r(\mathbf{G})_{rq}(\mathbf{G}^m)_{qp} = \sum_r a^\dagger_r(\mathbf{G}^{m+1})_{rp}$$

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

Remark B.3. Given rmk B.2 and prop B.1, the transformation of particle-hole operators can be expressed as

$$\begin{aligned} a'^\dagger_p &= \exp(X - X^\dagger)a^\dagger_p \exp(X^\dagger - X) \\ a'_p &= \exp(X - X^\dagger)a_p \exp(X^\dagger - X) \end{aligned} \qquad\qquad X = \sum_{p>q}(\mathbf{X})_{pq}a^\dagger_p a_q \qquad (\text{B.3})$$

where the annihilation operator transformation is simply the adjoint of the one for creation operators.

Remark B.4. If Θ' is obtained by replacing all of the orbitals in the basis expansion of $\Theta \in \mathcal{F}$ with primed orbitals, then

$$\Theta' = \exp(X - X^\dagger)\Theta \qquad (\text{B.4})$$

which follows from substituting equation B.3 into $a'^\dagger_{p_1}\cdots a'^\dagger_{p_n}|\text{vac}\rangle$ to prove that $|\Phi'_{(p_1\cdots p_n)}\rangle = \exp(X - X^\dagger)|\Phi_{(p_1\cdots p_n)}\rangle$[14] for any basis state.

---

[11]See https://en.wikipedia.org/wiki/Spectral_theorem
[12]Since there exists a basis in which $\mathbf{N}$ is diagonal, statements about $\mathbf{N}$ translate into statements about its eigenvalues.
[13]The second equality here follows from expanding G and using $[a^\dagger_r a_s, a^\dagger_q] = {:}a^\dagger_r \overline{a_s a^\dagger_q}{:} = a^\dagger_r\,\delta_{sq}$.
[14]Note that $\exp(X^\dagger - X) = (\exp(X - X^\dagger))^{-1}$ and $\exp(X^\dagger - X)|\text{vac}\rangle = 0$.