1 Perturbed Orthonormality Condition

1.1 Matrix Parameterization

We have the general orthonormality condition, subject to perturbation, as:

$$S_{pq} = \langle p|q\rangle = \delta_{pq} \tag{1}$$

$$\sum_{\mu\nu} C^*_{\mu p} S_{\mu\nu} C_{\nu q} = \delta_{pq} \tag{2}$$

Parameterization of the MO coefficients:

$$\mathbf{C}(\lambda) = \mathbf{C}(0)\mathbf{U}(\lambda) \tag{3}$$

$$C_{\mu p}(\lambda) = \sum_r C_{\mu r}(0) U_{rp}(\lambda) \tag{4}$$

in which $\mathbf{U}(\lambda)$ is the solution to the CPHF equations.

Now the orthonormality condition using this parameterization:

$$\sum_{\mu\nu} \Big(\sum_r U^*_{rp}(\lambda) C^*_{\mu r}(0)\Big) S_{\mu\nu}(\lambda) \Big(\sum_s C_{\nu s}(0) U_{sq}(\lambda)\Big) = \delta_{pq} \tag{5}$$

Introducing the transformed overlap matrix:

$$\mathcal{S}_{pq}(\lambda) = \sum_{\mu\nu} C^*_{\mu p}(0) S_{\mu\nu}(\lambda) C_{\nu q}(0) \tag{6}$$

we have:

$$\sum_{rs} U^*_{rp}(\lambda) \mathcal{S}_{rs}(\lambda) U_{sq}(\lambda) = \delta_{pq} \tag{7}$$

differentiating both sides of the equation gives:

$$\sum_{rs} \frac{\mathrm{d}U^*_{rp}(\lambda)}{\mathrm{d}\lambda} \mathcal{S}_{rs}(\lambda) U_{sq}(\lambda) + \sum_{rs} U^*_{rp}(\lambda) \frac{\mathrm{d}\mathcal{S}_{rs}(\lambda)}{\mathrm{d}\lambda} U_{sq}(\lambda) + \sum_{rs} U^*_{rp}(\lambda) \mathcal{S}_{rs}(\lambda) \frac{\mathrm{d}U_{sq}(\lambda)}{\mathrm{d}\lambda} = 0 \tag{8}$$

Noting that $\boldsymbol{\mathcal{S}}(0) = \mathbf{I}$ because the unperturbed spin-orbitals are orthonormal, and it is trivial that $\mathbf{U}(0) = \mathbf{I}$.
Therefore evaluating the derivative at $\lambda = 0$, and denoting $A^\lambda = \left(\frac{\mathrm{d}A}{\mathrm{d}\lambda}\right)\big|_{\lambda=0}$ results in:

$$\sum_{rs} (U^\lambda_{rp})^* \delta_{rs}\delta_{sq} + \sum_{rs} \delta_{rp} \mathcal{S}^\lambda_{rs} \delta_{sq} + \sum_{rs} \delta_{rp}\delta_{rs} U^\lambda_{sq} = 0 \tag{9}$$

contracting the Kronecker delta tensors we get the perturbed orthonormality condition:

$$(U^\lambda_{qp})^* + \mathcal{S}^\lambda_{pq} + U^\lambda_{pq} = 0 \tag{10}$$

1.2 Exponential Parameterization

$$\begin{aligned}\mathbf{C}(\lambda) =& \mathbf{C}(0)\mathbf{U}(\lambda) && (11)\\ \mathbf{U}(\lambda) =& \boldsymbol{\mathcal{S}}^{-\frac{1}{2}}(\lambda)\exp[-\boldsymbol{\kappa}(\lambda)] && (12)\\ \boldsymbol{\mathcal{S}}(\lambda) =& \mathbf{C}^{\dagger}(0)\mathbf{S}^{\mathrm{AO}}(\lambda)\mathbf{C}(0) && (13)\end{aligned}$$

In this way:

$$\begin{aligned}\mathbf{S}(\lambda) =& \mathbf{C}^{\dagger}(\lambda)\mathbf{S}^{\mathrm{AO}}(\lambda)\mathbf{C}(\lambda)\\ =& \mathbf{U}^{\dagger}(\lambda)\mathbf{C}^{\dagger}(0)\mathbf{S}^{\mathrm{AO}}(\lambda)\mathbf{C}(0)\mathbf{U}(\lambda)\\ =& \mathbf{U}^{\dagger}(\lambda)\boldsymbol{\mathcal{S}}(\lambda)\mathbf{U}(\lambda)\\ =& \exp[-\boldsymbol{\kappa}(\lambda)]^{\dagger}\boldsymbol{\mathcal{S}}^{-\frac{1}{2}\dagger}(\lambda)\boldsymbol{\mathcal{S}}(\lambda)\boldsymbol{\mathcal{S}}^{-\frac{1}{2}}(\lambda)\exp[-\boldsymbol{\kappa}(\lambda)]\\ =& \exp[\boldsymbol{\kappa}(\lambda)]\exp[-\boldsymbol{\kappa}(\lambda)]\\ =& \mathbf{I} && (14)\end{aligned}$$

the orthonormality is ensured by $\boldsymbol{\mathcal{S}}^{-\frac{1}{2}\dagger}\boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{S}}^{-\frac{1}{2}} = \mathbf{I}$, as S^{AO} thus $\boldsymbol{\mathcal{S}}$ matrix is Hermitian. The perturbed orthonormality condition is then trivial:

$$\mathbf{S}^{\lambda} = \frac{\partial \mathbf{S}(\lambda)}{\partial \lambda}\Big|_{\lambda=0} = \mathbf{0} \tag{15}$$

With this new parameterization, the orthonormality condition and perturbed orthonormality condition come naturally, hence do not need to be included explicitly in the Lagrangian.

The derivative of $\boldsymbol{\mathcal{S}}^{-\frac{1}{2}}(\lambda)$ (needed for $\mathbf{U}^{\lambda}$ appearing in second derivative) could be solved by taking derivative on this equation:

$$\boldsymbol{\mathcal{S}}^{-\frac{1}{2}\dagger}\boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{S}}^{-\frac{1}{2}} = \mathbf{I} \tag{16}$$

i.e. (note that $\boldsymbol{\mathcal{S}}^{\dagger} = \boldsymbol{\mathcal{S}}$ and $\boldsymbol{\mathcal{S}}(\lambda = 0) = \mathbf{I}$):

$$\begin{aligned}\mathbf{0} =& \frac{\partial \boldsymbol{\mathcal{S}}^{-\frac{1}{2}}}{\partial \lambda}\boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{S}}^{-\frac{1}{2}}\Big|_{\lambda=0} + \boldsymbol{\mathcal{S}}^{-\frac{1}{2}}\frac{\partial \boldsymbol{\mathcal{S}}}{\partial \lambda}\boldsymbol{\mathcal{S}}^{-\frac{1}{2}}\Big|_{\lambda=0} + \boldsymbol{\mathcal{S}}^{-\frac{1}{2}}\boldsymbol{\mathcal{S}}\frac{\partial \boldsymbol{\mathcal{S}}^{-\frac{1}{2}}}{\partial \lambda}\Big|_{\lambda=0}\\ =& 2\left(\boldsymbol{\mathcal{S}}^{-\frac{1}{2}}\right)^{\lambda} + \boldsymbol{\mathcal{S}}^{\lambda} && (17)\end{aligned}$$

$$\Leftrightarrow \left(\boldsymbol{\mathcal{S}}^{-\frac{1}{2}}\right)^{\lambda} = -\frac{1}{2}\boldsymbol{\mathcal{S}}^{\lambda} \tag{18}$$