Notes on Gradients

B.Mussard
https://mussard.github.io

May 7, 2019

Contents

We will follow derivations from the Yamaguchi, Osamura, Goddard and Schaefer (YOGS) book, and from the Sherrill notes (SHERRILL) and equations will refer to those documents. Everywhere, $mnopqrs$ are any orbitals, $ijkl$ are occupied (in HF) and $abcd$ are unoccupied (in HF). We use Einstein notations.

1 Energy and derivative expressions

The CASCI energy is written:

$$E^{\text{CASCI}} = Q_{pq}^{\text{CASCI}} h_{pq} + G_{pqrs}^{\text{CASCI}} v_{pqrs}, \tag{1}$$

YOGS - 6.8 - p84

where Q_{pq}^{CASCI} and G_{pqrs}^{CASCI} are the one- and two-RDMs respectively and h_{pq} and v_{pqrs} are the corresponding tensors. The RDMs only depend on the structure of the wavefunction, *i.e.* on the CI coefficients. The tensors depend on the orbital coefficients and are the source of complexity here. The first-order derivative of the energy with respect to an arbitrary *external* parameter is given by:

$$\frac{\mathrm{d}E^{\mathrm{CASCI}}}{\mathrm{d}\alpha} = Q_{pq}^{\mathrm{CASCI}} \frac{\partial h_{pq}}{\partial \alpha} + G_{pqrs}^{\mathrm{CASCI}} \frac{\partial v_{pqrs}}{\partial \alpha}, \tag{2}$$

YOGS - 6.17 - p85

that is to say:

$$\frac{\mathrm{d}E}{\mathrm{d}\alpha} = Q_{pq}^{\mathrm{CASCI}} h_{pq}^{\alpha} + G_{pqrs}^{\mathrm{CASCI}} v_{pqrs}^{\alpha} + 2U_{pq}^{\alpha} X_{pq}^{\mathrm{CASCI}}, \tag{3}$$

YOGS - 6.23 - p86

where the two first terms are the intrinsic derivative coming from the change in the AO basis, and the last term captures the change in the MO coefficients. The quantities h_{pq}^{α}, v_{pqrs}^{α} (and later S_{pq}^{α}) are not defined here and are available in any quantum chemistry package. The elements X_{pq}^{CASCI} are given by

$$X_{pq}^{\mathrm{CASCI}} = Q_{qm}^{\mathrm{CASCI}} h_{pm} + 2G_{qmno}^{\mathrm{CASCI}} v_{pmno}, \tag{4}$$

YOGS - 6.24 - p87

and we need to solve the CPHF equations for U_{pq}^{α}.

2 CPHF equation

The CPHF equation to be solved for U_{pq}^{α} are derived by taking the derivative of the variational condition of the HF equation, $F_{ia}^{\mathrm{HF}} = 0$. They are derived by differentiating all the Fock elements:

$$F_{pq}^{\mathrm{HF}} = h_{pq} + 2v_{pqii} - v_{piqi}, \tag{5}$$

YOGS - 10.1 - p128

yielding:

$$\frac{\mathrm{d}F_{pq}^{\mathrm{HF}}}{\mathrm{d}\alpha} = 0 \iff \Delta\epsilon_{pq}^{\mathrm{HF}} U_{pq}^{\alpha} - A_{pq,bj}^{\mathrm{HF}} U_{bj}^{\alpha} = B_{pq}^{\mathrm{HF},\alpha}, \tag{6}$$

YOGS - 10.5,10.6,10.9,10.12,10.18,10.21,10.22 - p131

with:

$$\Delta\epsilon_{pq}^{\mathrm{HF}} = (\epsilon_q - \epsilon_p) \qquad B_{pq}^{\mathrm{HF},\alpha} = F_{pq}^{\mathrm{HF},\alpha} - S_{pq}^{\alpha}\epsilon_q - S_{ij}^{\alpha}(2v_{pqij} - v_{piqj})$$

$$A_{pq,rs}^{\mathrm{HF}} = 4v_{pqrs} - v_{prqs} - v_{psqr} \qquad F_{pq}^{\mathrm{HF},\alpha} = h_{pq}^{\alpha} + 2v_{pqii}^{\alpha} - v_{piqi}^{\alpha}$$

YOGS - 10.8,10.23,10.7 - pp129,131

The elements of U^α_{pq} outside of the occupied-virtual block are not well-defined, hence the CPHF equation are first solved in the occupied-virtual block (see the section on z-vector equations) and then the non-independant pair *in the sense of HF* are defined as (from Eq.(6)):

$$\begin{aligned} U^\alpha_{kl} &= \frac{1}{\Delta\epsilon^{\mathrm{HF}}_{kl}}(A^{\mathrm{HF}}_{kl,bj}U^\alpha_{bj} + B^{\mathrm{HF},\alpha}_{kl}) \\ U^\alpha_{cd} &= \frac{1}{\Delta\epsilon^{\mathrm{HF}}_{cd}}(A^{\mathrm{HF}}_{cd,bj}U^\alpha_{bj} + B^{\mathrm{HF},\alpha}_{cd}) \end{aligned} \tag{7}$$

YOGS - 10.26 - p132

3 Working with U^α and z-vector equation

Note that in the following, we'll drop the "HF" and "CASCI" super-script for clarity: remember that X_{pq} contains CASCI data and $A_{pq,rs}$ and $B^\alpha_{pq,rs}$ contain HF data. In Eq.(3), there is a summation over all pair of orbitals. It can be worked as:

$$2\sum_{pq} U^\alpha_{pq}X_{pq} = 2\sum_{p>q} U^\alpha_{pq}\Delta X_{pq} - \sum_{pq} S^\alpha_{pq}\bar{X}_{pq}, \tag{8}$$

SHERRILL - 42 - p8

which only uses the derivative of the orthogonality condition of the MOs, $U^\alpha_{pq} + U^\alpha_{qp} + S^\alpha_{pq} = 0$, and which defines:

$$\Delta X_{pq} = X_{pq} - X_{qp} \quad \text{and} \quad \bar{X}_{pq} = \begin{cases} X_{pq} & p \le q \\ X_{qp} & p > q \end{cases} \tag{9}$$

SHERRILL - 41 - p8

We can work further into this expression with:

$$2\sum_{p>q} U^\alpha_{pq}\Delta X_{pq} = 2\sum_{j>k} U^\alpha_{jk}\Delta X_{jk} + 2\sum_{ai} U^\alpha_{ai}\Delta X_{ai} + 2\sum_{b>c} U^\alpha_{bc}\Delta X_{bc}, \tag{10}$$

SHERRILL - 54 - p10

where U^α_{jk} and U^α_{bc} are given by Eq.(7), so that:

$$2\sum_{p>q} U^\alpha_{pq}\Delta X_{pq} = 2\sum_{j>k} \frac{\Delta X_{jk}}{\Delta\epsilon_{jk}}B^\alpha_{jk} + 2\sum_{ai} U^\alpha_{ai}\Delta X'_{ai} + 2\sum_{b>c} \frac{\Delta X_{bc}}{\Delta\epsilon_{bc}}B^\alpha_{bc}, \tag{11}$$

SHERRILL - 56 - p11

where:

$$\Delta X'_{ai} = \Delta X_{ai} + \sum_{j>k} \frac{\Delta X_{jk}}{\Delta\epsilon_{jk}}A_{jk,ai} + \sum_{b>c} \frac{\Delta X_{bc}}{\Delta\epsilon_{bc}}A_{bc,ai}, \tag{12}$$

SHERRILL - 57 - p11

and where the construction of $2U^{\alpha}_{ai}\Delta X'_{ai}$ in Eq.(11) requires to solve Eq.(6) "number of α" times for all U^{α}_{ai}. It is replaced by the *z-vector* equation to be solved once, as follows:

$$\begin{aligned} & 2U^{\alpha}_{ai}\Delta X'_{ai} \quad \text{where} \quad A'_{ai,bj}U^{\alpha}_{bj} = B^{\alpha}_{ai} && (13)\\ \Longleftrightarrow\ & 2Z_{ai}B^{\alpha}_{ai} \quad \text{where} \quad A'^{\mathrm{T}}_{ai,bj}Z_{bj} = \Delta X'_{ai}, && (14)\end{aligned}$$

YOGS-p356/7 - and - SHERRILL-p11/2

where of course we have:

$$A'_{ai,bj} = \delta_{ab}\delta_{ij}(\epsilon_i - \epsilon_a) - A_{ai,bj} \tag{15}$$

SHERRILL - 50 - p10

This leads to Eq.(11) being contracted into:

$$2\sum_{p>q} U^{\alpha}_{pq}\Delta X_{pq} = 2\sum_{p>q} Z_{pq}B^{\alpha}_{pq} \tag{16}$$

SHERRILL - 65 - p12

where:

$$\begin{aligned} & A^{\mathrm{T}}_{ai,bj}Z_{bj} = \Delta X'_{ai} \\ & Z_{jk} = \frac{\Delta X_{jk}}{\Delta\epsilon_{jk}} \\ & Z_{bc} = \frac{\Delta X_{bc}}{\Delta\epsilon_{bc}} \end{aligned} \tag{17}$$

SHERRILL - 66-67-68 - p12

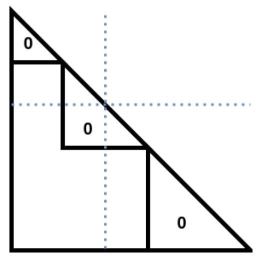
4 Note on the structure of U^{α} and X

In the derivation, the term that holds the focus of attention are the ones encountered in Eqs.(11) and (12), where ΔX_{pq} holds CASCI information and U^{α}_{pq} stems from changes in the MO, *i.e.* will carry HF denominations. In particular, the structure of X_{pq} will be separated in core, active and virtual blocks, while U^{α}_{pq} is separated in occupied and virtual parts.

About U^{α} : the "ov" block is defined by solving the CPHF equations. The diagonal blocks are not well-defined and are chosen to be constructed as indicated in Eq.(7).

$U^\alpha =$ 

About ΔX : by construction, the matrix whose elements are X_{pq} is symmetric when pq are non-independant pairs *for CASCI*. That is not the non-independant pairs *for HF*, which would be "oo" and "vv", but it is rather "cc", "aa" and "vv". Hence ΔX_{pq} is zero in those blocks.

$\Delta X =$ 

Hence, most of the summations seen in Section 3 run over a lot of pairs of orbitals that are non-independant for CASCI, but those terms will not contribute because then $\Delta X_{pq} = 0$. For example, the following term seen in the derivation could be reduced to run over the independant pair *for CASCI*:

$$2 \sum_{j>k} U^\alpha_{jk} \Delta X_{jk} \rightarrow 2 \sum_{j>k}^{\text{CASCI-IP}} U^\alpha_{jk} \Delta X_{jk}, \tag{18}$$

both results being equal. These restrictions could be used to implement a code that would be more conservative in term of memory, but are not necessary to the derivations because of the $\Delta X_{pq} = 0$ property. In these notes, those restrictions are never marked.

A An interesting link

This is an effort to show a link between X_{pq} and the derivative with respect to changes in the orbital coefficients. This strays out of the path of the book and of the notes. The energy depends explicitly on the external parameter, and implicitly through the dependance of the energy parameter to the external parameter.

Hence, I would of course write the derivative of the energy $E^{\text{CASCI}}[\alpha; c[\alpha], \kappa[\alpha]]$ as:

$$\frac{\mathrm{d}E^{\text{CASCI}}}{\mathrm{d}\alpha} = \frac{\partial E^{\text{CASCI}}}{\partial \alpha} + \frac{\partial E^{\text{CASCI}}}{\partial c}\frac{\partial c}{\partial \alpha} + \frac{\partial E^{\text{CASCI}}}{\partial \kappa}\frac{\partial \kappa}{\partial \alpha} \tag{19}$$

where the first term is the two first terms of Eq.(3), the second term is zero because the CASCI energy is variational with respect to the CI coefficients and where in the last term I believe $\frac{\partial E^{\text{CASCI}}}{\partial \kappa}$ is simply $2X_{pq}^{\text{CASCI}}$ and $\frac{\partial \kappa}{\partial \alpha}$ is U_{pq}^{α}. Indeed, let's remember that the CASCI energy is:

$$E^{\text{CASCI}} = Q_{mn}^{\text{CASCI}} h_{mn} + G_{mnor}^{\text{CASCI}} v_{mnor} \tag{20}$$

and let's write the orbital coefficients as (this is where I'm not quite confortable):

$$C_{m\mu} = C_{m\mu} + \kappa_{mo} C_{o\mu} \tag{21}$$

then we have:

$$h_{mn} = (C_{m\mu} + \kappa_{mo} C_{o\mu}) h_{\mu\nu} (C_{\nu n}^{\dagger} + C_{\nu r}^{\dagger} \kappa_{rn}^{\dagger}) \tag{22}$$

and:

$$\frac{\partial h_{mn}}{\partial \kappa_{pq}} = \delta_{mp}\delta_{oq} h_{on} + \delta_{np}\delta_{rq} h_{mr} \tag{23}$$

so that:

$$\sum_{mn} Q_{mn}^{\text{CASCI}} \frac{\partial h_{mn}}{\partial \kappa_{pq}} = \sum_{n} Q_{pn}^{\text{CASCI}} h_{qn} + \sum_{m} Q_{mp}^{\text{CASCI}} h_{mq} = 2\sum_{m} Q_{mp}^{\text{CASCI}} h_{mq} \tag{24}$$

The two-index quantity h_{pq} takes a factor 2, and similarly the four-index quantity v_{pqrs} will take a factor 4, so that:

$$\frac{1}{2}\frac{\partial E^{\text{CASCI}}}{\partial \kappa_{pq}} = Q_{mp}^{\text{CASCI}} h_{mq} + 2G_{mnop}^{\text{CASCI}} v_{mnoq} \tag{25}$$

which is X_{pq}^{CASCI}.

B Connexion to Lagrangian framwork

The CASCI energy is not variational with respect to the MO coefficient parameter. The Lagrangian framework would define:

$$\mathcal{L}[\kappa, c, \lambda] = E^{\text{CASCI}}[\kappa, c] + \lambda \frac{\partial E^{\text{HF}}[\kappa]}{\partial \kappa}, \tag{26}$$

and make it variational with respect to all its parameters:

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \frac{\partial E^{\mathrm{HF}}}{\partial \kappa} = 0 \tag{27}$$

$$\frac{\partial \mathcal{L}}{\partial c} = \frac{\partial E^{\mathrm{CASCI}}}{\partial c} = 0 \tag{28}$$

$$\frac{\partial \mathcal{L}}{\partial \kappa} = \frac{\partial E^{\mathrm{CASCI}}}{\partial \kappa} + \lambda \frac{\partial^2 E^{\mathrm{HF}}}{\partial \kappa \partial \kappa} = 0 \tag{29}$$

where the two first equations are realized by design of the HF and CASCI methods (ensuring that the Lagrangian is not defining a functional of another theory), and where the last equation is solved for λ to make the Lagrangian variational. Under these conditions, the first order derivative of the Lagrangian is merely given by its intrinsic derivative:

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\alpha} = \frac{\partial \mathcal{L}}{\partial \alpha} = Q^{\mathcal{L}}_{pq} h^{\alpha}_{pq} + G^{\mathcal{L}}_{pqrs} v^{\alpha}_{pqrs}, \tag{30}$$

where $Q^{\mathcal{L}}_{pq}$ and $G^{\mathcal{L}}_{pq,rs}$ are one- and two-RDMs that can be defined as $\frac{\partial \mathcal{L}}{\partial h_{pq}}$ and $\frac{\partial \mathcal{L}}{\partial v_{pqrs}}$, *i.e* from Eq.(26):

$$\mathcal{L} = (Q^{\mathrm{CASCI}}_{pq} h_{pq} + G^{\mathrm{CASCI}}_{pqrs} v_{pqrs}) + \lambda_{mn}(Q^{\mathrm{HF}}_{mq} h_{nq} + 2G^{\mathrm{HF}}_{mqrs} v_{nqrs}) \tag{31}$$

$$= (Q^{\mathrm{CASCI}}_{pq} + \lambda_{mp} Q^{\mathrm{HF}}_{mq}) h_{pq} + (G^{\mathrm{CASCI}}_{pqrs} + 2\lambda_{mp} G^{\mathrm{HF}}_{mqrs}) v_{pqrs} \tag{32}$$

yielding:

$$Q^{\mathcal{L}}_{pq} = Q^{\mathrm{CASCI}}_{pq} + \lambda_{mp} Q^{\mathrm{HF}}_{mq} \tag{33}$$

$$G^{\mathcal{L}}_{pqrs} = G^{\mathrm{CASCI}}_{pqrs} + 2\lambda_{mp} G^{\mathrm{HF}}_{mqrs} \tag{34}$$

Let's look at what Eq.(29) yields:

(35)

(36)

(37)