Optimisation of distance in the Grassmannian to an external wave function

Yuri Alexandre Aoto

History:

24 mar 2019 - Start

1 Notation

$\mathcal{H}$	The electronic Hamiltonian
n_α	The number of alpha electrons
n_β	The number of beta electrons
n	The total number of electrons
	$n = n_\alpha + n_\beta$
$\mathcal{V}$	The space of spatial one-electron wave functions
	(The full space of one-electron wave functions is
	assumed to be a direct sum of two such spaces,
	the first associated to α-spin and the second
	to β-spin electrons. A subscript might be used,
	but the space is the same: $\mathcal{V}_\alpha = \mathcal{V}_\beta = \mathcal{V}$)
N	The dimension of $\mathcal{V}$: $N = \dim \mathcal{V}$
i,j,k,l	orbital indices for occupied orbitals
	(the context should make clear the reference determinant)
a,b,c,d	orbital indices for virtual orbitals
	(the context should make clear the reference determinant)
$\sigma(p)$	spin (α or β) of orbital with index p
$\mathscr{H}$	The complete Hilbert space of quantum states
	(within finite basis set approximation)
	$\mathscr{H} = \Lambda^n(\mathcal{V}_\alpha \oplus \mathcal{V}_\beta)$
Ψ	Elements of $\mathscr{H}$
Φ	Determinantal elements of $\mathscr{H}$
ψ	Elements of $\mathcal{V}$

2 Introduction

Let

$$|\Psi_{\mathrm{FCI}}\rangle = \sum_I c_I \, |\Phi_I\rangle \tag{1}$$

be a n-electron normalised Full-CI like wave function represented in the orbital basis

$$|\phi_I\rangle = \phi_{I_1} \wedge \phi_{I_2} \wedge \cdots \wedge \phi_{I_n} \, , \tag{2}$$

where $\{\phi_p\}$ is an orthonormal basis for the space of one-electron wave functions and I is an ordered multi-index. We want to find $|\Phi\rangle \in Gr$ such that $|\left\langle \Psi_{\mathrm{FCI}} \middle| \Phi \right\rangle|$ is maximum, where Gr is the image of the Grassmannian in the space of the n-electron wave functions. Recall that

$$D(\psi_1, \psi_2) = \sqrt{2}\sqrt{1 - |\left\langle \psi_1 \middle| \psi_2 \right\rangle|} \tag{3}$$

is a metric in $\mathbb{P}\mathscr{H}$.[] Also, Gr is the set of all elements of $\mathscr{H}$ that can be written as a single Slater determinant (a decomposable element) for some orthonormal basis of $\mathcal{V}$.

3 Minimisation with respect to FCI

To find $|\Phi\rangle$, it is equivalent to find the orbitals (namely, a basis of $\mathcal{V}$) such that

$$|\Psi_{\rm FCI}\rangle = C_0\,|\Phi\rangle + \sum_{\substack{I\\ I\neq\{1,2,\ldots,n\}}} c_I\,|\Phi_I\rangle\ , \tag{4}$$

and $|C_0|$ is maximum (over all possible coefficients in all possible basis), since $|C_0| = |\left\langle\Psi_{\rm FCI}\middle|\Phi\right\rangle|$.

3.1 Parametrisation by orbital rotations

We parametrise Gr by the orbital rotations as[]

$$|\Phi\rangle = e^{-\hat{K}}\,|\Phi_0\rangle\ , \tag{5}$$

where

$$\begin{aligned}\hat{K} &= \sum_{i,a} K_i^a(a_a^\dagger a_i - a_i^\dagger a_a) &(6)\\ &= \sum_{i,a} K_i^a(a_i^a - a_a^i) &(7)\end{aligned}$$

This parametrisation comes from the most general

$$\hat{K} = \sum_{p,q} K_q^p a_p^\dagger a_q\,, \tag{8}$$

but using that K_q^p is anti-symmetric (so that $e^{-\hat{K}}$ is orthogonal), and excluding rotations within the occupied or virtual spaces of $|\Phi_0\rangle$, that are redundant (do not alter the Slater determinant with $|\Phi_0\rangle$). For $\hat{K}=0$ it is clear that $|\Phi\rangle = |\Phi_0\rangle$.

Let

$$f(K_i^a) = \left|\left\langle\Psi_{\rm FCI}\middle|e^{-\hat{K}}\middle|\Phi_0\right\rangle\right|\ , \tag{9}$$

where the argument K_i^a represent all the $n_\alpha(N-n_\alpha)+n_\beta(N-n_\beta)$ elements. Note that rotations that mix α and β orbitals. We will also assume that $\left\langle\Psi_{\rm FCI}\middle|\Phi_0\right\rangle > 0$ and this remains true for all steps of our optimisation. If $\left\langle\Psi_{\rm FCI}\middle|\Phi_0\right\rangle < 0$ we of course can change the phase of the wave function and if $\left\langle\Psi_{\rm FCI}\middle|\Phi_0\right\rangle = 0$ for the first or any step of the optimisation, we likely started with a very poor initial guess.

3.2 Jacobian and Hessian

We want to maximise f and we need its Jacobian and Hessian. The expressions at $\hat{K}=0$ are given below. For the derivation of the expressions, see that hand notes.

$$\frac{\partial f(\hat{K}=0)}{\partial K_i^a} = (-1)^{n_{\sigma(i)}-i+1}C_i^a \tag{10}$$

$$\frac{\partial^2 f(\hat{K}=0)}{\partial K_i^a \partial K_j^b} = \begin{cases} -C_0 & (i=j, a=b)\\ 0 & (i\neq j, a=b)\\ 0 & (i=j, a\neq b)\\ (-1)^{n_{\sigma(i)}+n_{\sigma(j)}-i-j}C_{ij}^{ab} & (\sigma(i)\neq\sigma(j))\\ (-1)^{i+j+1}C_{ij}^{ab} & (\sigma(i)=\sigma(j), i<j, a<b)\\ (-1)^{i+j}C_{ij}^{ab} & (\sigma(i)=\sigma(j), i<j, a>b)\end{cases} \tag{11}$$

In these equations, C_i^a and C_{ij}^{ab} are the CI coefficients of the single and double excited determinants in the (normalised) wave function $|\Psi_{\mathrm{FCI}}\rangle$. The canonical order of the orbitals is assumed to be "first all α, then all β".

3.3 Transformation of the wave function

In the optimisation process, $|\Phi\rangle$ varies and we would need the Jacobian and the Hessian at $\hat{K} \neq 0$. The expressions are much more complicated and we avoid this by making a full transformation of $|\Psi_{\mathrm{FCI}}\rangle$ to the new orbital basis. Let $U = e^{-\hat{K}}$ be the matrix that transform the orbital basis:

$$\phi_p = \sum_q \phi'_q U_{qp} \,. \tag{12}$$

Given the coefficients C_I of the expansion in the first basis, we want to know the coefficients C'_I such that

$$|\Psi_{\mathrm{FCI}}\rangle = \sum_I c_I \,|\Phi_I\rangle = \sum_I c'_I \,|\Phi'_I\rangle \ . \tag{13}$$

These are given by:

$$C'_I = \sum_J C_J \det(U_{IJ}) \,, \tag{14}$$

where U_{IJ} is the minor of the matrix U with the entries in the rows and columns given by the multi-indices I and J.

Such transformation is the most time consuming step.

3.4 Newton-Raphson step

Starting from a orbital basis $\{\phi_p\}$ such that the first determinant (that is, with the first n_α α orbitals and first n_β β occupied) is $|\Phi_0\rangle$, we calculate the Jacobian $\mathbf{J}$ and the Hessian $\mathbf{H}$ as shown above. The Newton step (in the space of the K_i^a parameters) is

$$\mathbf{z} = -\mathbf{H}^{-1}\mathbf{J} \,. \tag{15}$$

From this vector, the operator $\hat{K}$ is constructed and the orbital transformation matrix is given by

$$U = e^{-\hat{K}} \,. \tag{16}$$

This is done for the α and the β orbitals and the wave function $|\Psi_{\mathrm{FCI}}\rangle$ is transformed to the new orbital basis, by equation 14. This proceeds until convergence.