On the Computation of Excited States with MCSCF Methods

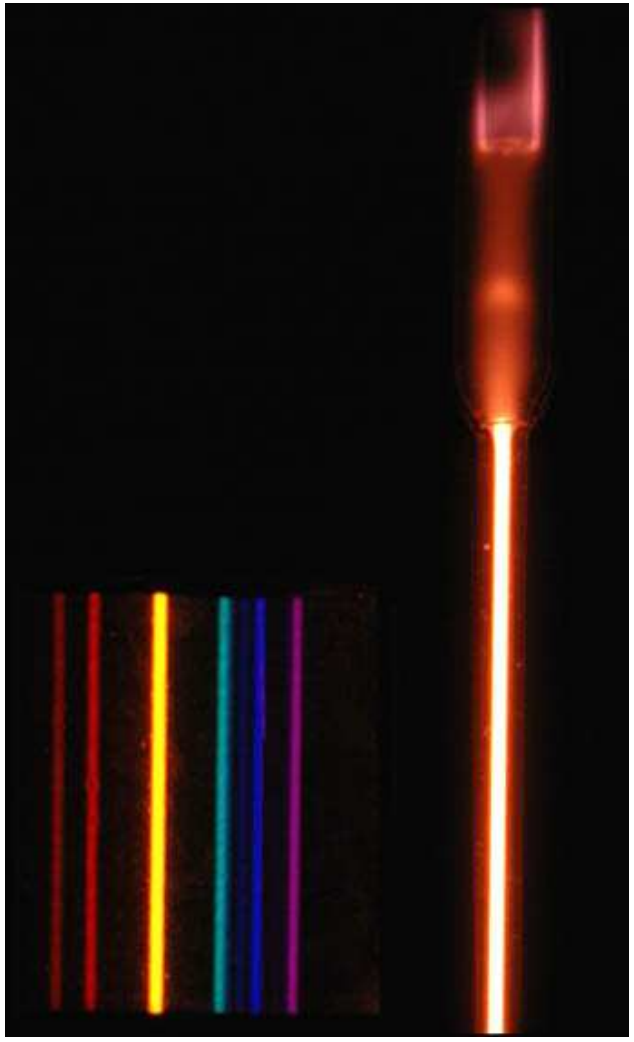
Mathieu LEWIN

Mathieu.Lewin@math.cnrs.fr

CNRS & University of Cergy-Pontoise

collaboration with É. Cancès & H. Galicher (CERMICS, ENPC)

Introduction



Computation of Excited States: very important in practice. Hartree-Fock or Density Functional Theory type methods do not seem to be adapted to this issue.

Multiconfiguration (MC) methods: method of choice for the computation of excited states. However, problems are sometimes encountered in practice.

Right: a helium spectral tube excited by means of a 5000 Volt transformer. Left: the spectral lines through a 600 line/mm diffraction grating. *http://hyperphysics.phy-astr.gsu.edu/*

Some practical difficulties

“Convergence of the orbital optimization procedure is normally good for CASSCF type wavefunctions, but problems can occur in calculations of excited states, especially when several states are close in energy."
(instructions for the MOLCAS software)

“MCSCF is by no means the sort of ‘black box’ that HF these days, so LOOK VERY CAREFULLY AT YOUR RESULTS."
(instructions for the GAMESS software)

This talk

Main goal:

- try to give a mathematical explanation / interpretation of the practical difficulties;
- propose some solutions ?

References:

- [ML] M.L. Solutions of the Multiconfiguration Equations in Quantum Chemistry, *Arch. Rat. Mech. Anal.* **171** (2004), no. 1, 83–114.
- [CGL] E. Cancès, H. Galicher & M.L. Computing Electronic Structures: a New Multiconfiguration Approach for Excited States. *J. Comput. Phys.*, in press.

I. Definition of the model

The time-independent Schrödinger model

We consider a molecule containing

• M **nuclei**, treated as pointwise charged particules. Charges Z_m clamped at positions $R_m \in \mathbb{R}^3$, $m = 1, ..., M$. Define

$$V(x) = \sum_{m=1}^{M} \frac{-Z_m}{|x - R_m|}, \qquad Z = \sum_{m=1}^{M} Z_m \ ;$$

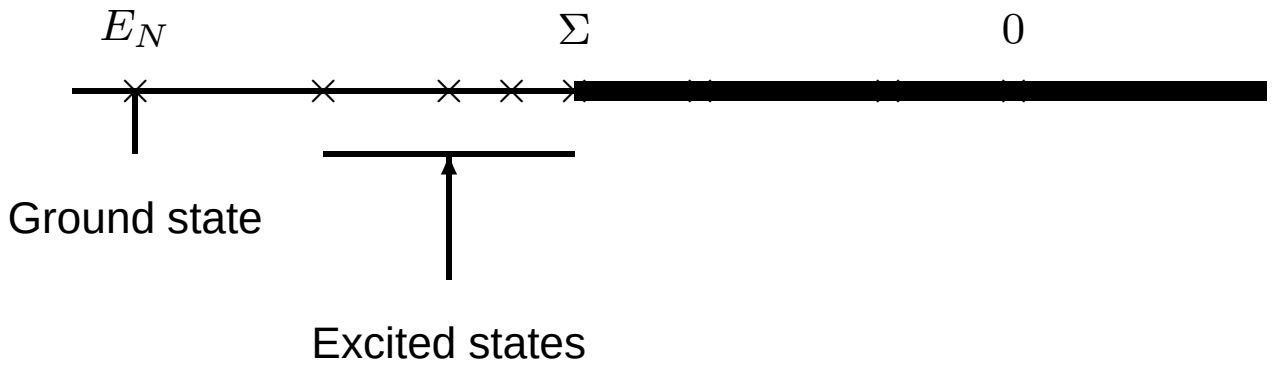
• N **non-relativistic electrons**, whose state is a normalized antisymmetric wavefunction $\Psi \in \bigwedge_{i=1}^{N} L^2(\mathbb{R}^3)$. Energy: $\mathcal{E}(\Psi) = \langle \Psi, H_N \Psi \rangle$ where H_N is the N-body quantum Hamiltonian:

$$H_N = \sum_{i=1}^{N} \left(-\frac{\Delta_{x_i}}{2} + V(x_i) \right) + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|}.$$

(real-valued wavefunctions with no spin, atomic units)

The spectrum of the N-body Hamiltonian

When $Z = \sum_{m=1}^{M} Z_m \geq N$, $\sigma(H_N) = \{E_N = \lambda_0 \leq \lambda_1 \leq \cdots\} \cup [\Sigma; +\infty)$,



(Zhislin, 1960 & Zhislin-Sigalov, 1965). Moreover $\Sigma < 0$ when $N \geq 2$.

A ground or excited state is a solution of the time independent Schrödinger equation: $\boxed{H_N \Psi = \lambda_k \Psi}$

- If $N > Z$, there is a **finite number** of eigenvalues below Σ;
- there is **no** eigenvalue if $N >> Z$ (Sigal, 1982...);
- HVZ Theorem (Hunziker, Van Winter, Zhislin, 1960) : $\Sigma = E_{N-1}$.

The MC approximation

N-body problem not tractable as such $\rightarrow$ approximate models.

Starting point: let (ψ_i) be an orth. basis of $L^2(\mathbb{R}^3)$. Then $(\psi_{i_1} \wedge \cdots \wedge \psi_{i_N})_{i_1<\cdots<i_N}$ is an orth. basis of $\bigwedge_1^N L^2(\mathbb{R}^3)$, where

$$\psi_1 \wedge \cdots \wedge \psi_N(x_1, ..., x_N) = \frac{1}{\sqrt{N!}} \det(\psi_i(x_j)). \quad \text{(Slater determinant)}$$

This means that each normalized $\Psi \in \bigwedge_1^N L^2(\mathbb{R}^3)$ can be written

$$\Psi = \sum_{1\leq i_1<\cdots<i_N} c_{i_1,...,i_N} \psi_{i_1} \wedge \cdots \wedge \psi_{i_N} \quad \text{with} \quad \sum_{1\leq i_1<\cdots<i_N} |c_{i_1,...,i_N}|^2 = 1.$$

MC Approximation:

$$\Psi = \sum_{1\leq i_1<\cdots<i_N\leq K} c_{i_1,...,i_N} \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N}, \qquad K \geq N.$$

Two unknowns

$$\Psi_{(c,\Phi)} = \sum_{1\leq i_1<\cdots<i_N\leq K} c_{i_1,\ldots,i_N}\varphi_{i_1}\wedge\cdots\wedge\varphi_{i_N} = \sum_I c_I\Phi_I, \qquad K\geq N.$$

Unknowns:

- $c=(c_{i_1,\ldots,i_N})$ (lexicog. order) with $\sum_{i_1<\cdots<i_N}|c_{i_1\ldots i_N}|^2=1$;
- $\Phi=(\varphi_1,\ldots,\varphi_K)^T\in(H^1(\mathbb{R}^3))^K$ with $\int_{\mathbb{R}^3}\varphi_i\varphi_j=\delta_{ij}$.

MC energy:

$$\boxed{\mathcal{E}_N^K(c,\Phi)=\left\langle\Psi_{(c,\Phi)},H_N\Psi_{(c,\Phi)}\right\rangle}$$

where $\Psi_{(c,\Phi)}$ is the previous wavefunction; it is defined on the manifold

$$\mathcal{M}_N^K=\left\{(c,\Phi)\in S^{\binom{K}{N}-1}\times(H^1(\mathbb{R}^3))^K,\ \int_{\mathbb{R}^3}\varphi_i\varphi_j=\delta_{ij}\right\}.$$

A class of nonlinear models

- $\mathcal{E}_N^K(c,\Phi)$ is still quadratic with respect to c:

$$\mathcal{E}_N^K(c,\Phi) = \sum_{I,J} c_I c_J \langle H_N \Phi_I, \Phi_J \rangle = \langle H_\Phi c, c \rangle$$

where $\Phi_I = \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N}$ if $I = \{i_1 < \cdots < i_N\}$, and H_Φ is the matrix of H_N when it is restricted to the $\binom{K}{N}$-dimensional space $V_\Phi = \mathrm{span}_{I=\{i_1<\cdots<i_N\}}(\Phi_I)$.

- $\mathcal{E}_N^K(c,\Phi)$is not quadratic with respect to the (φ_i)'s:

$$\begin{aligned}\mathcal{E}_N^K(c,\Phi) \quad = \quad & \sum_{i,j} \frac{\gamma_{ij}}{2} \int_{\mathbb{R}^3} (\nabla\varphi_i \cdot \nabla\varphi_j + V\varphi_i\varphi_j) \text{ (quadratic)} \\ & + \sum_{i,j,k,l} W_{ijkl} \iint_{\mathbb{R}^6} \frac{\varphi_i(x)\varphi_j(x)\varphi_k(y)\varphi_l(y)}{|x-y|} dx\, dy \text{ (quartic)}\end{aligned}$$

where γ_{ij} and W_{ijkl} only depend on c.

Nonlinear equations

Equations: a system of K coupled nonlinear elliptic PDEs (for the φ_i's) + a simple eigenvalue problem (for c).

$$\begin{cases} n_i(c)\left(-\frac{\Delta}{2}+V\right)\varphi_i + \sum_j W_{i,j}^{(c,\Phi)}\varphi_j = \sum_j \lambda_{ij}\varphi_j \\ H_\Phi\, c = \beta\, c \end{cases}$$

The operator $W_i^{(c,\Phi)}$ depends on (c,Φ) in a complicated manner.

Rmk: mathematically, it is not obvious at all to prove existence of solutions to this type of equations !
(since we work in an infinite dimensional space)
Possible *loss of compactness* at infinity (this will occur if $N \gg Z$).

Practical methods

In practice, not all the $\binom{K}{N}$ determinants (Φ_I) are considered ! One has to restrict to wavefunctions that have less Slater determinants.

Main method: Complete Active Space Method (CASSCF), see, e.g. Roos '87 *Adv. Chem. Phys.*.
The N electrons are separated into (usually $k \in [2;10]$)

- $N-k$ inactive (or core) electrons (treated by a HF-type method);
- k active (or valence) electrons which are supposed to mostly contribute to the correlation energy (treated by a "full" MC method).

$$\Psi = \varphi_1 \wedge \cdots \varphi_{N-k} \wedge \left(\sum_{N-k+1 \leq i_{N-k+1} < i_N \leq K} \varphi_{i_{N-k+1}} \wedge \cdots \wedge \varphi_{i_N} \right)$$

Symmetry properties are also frequently enforced.

Minimization

Theorem 1. (Friesecke '04, [ML] '05). *Let be $Z > N - 1$.*
For any $K \geq N$, $\mathcal{E}_N^K$ possesses a minimizer (c_0^K, Φ_0^K) on $\mathcal{M}_N^K$.
Moreover, when $K \to \infty$, the associated wavefunction Ψ_K converges (up to a subsequence) to a ground state $\bar{\Psi}$ of H_N, solution of the time-independent Schrödinger equation $H_N \bar{\Psi} = \lambda_0 \bar{\Psi}$.

- $K = N$: **Hartree-Fock** model. Existence of a minimizer (Lieb-Simon '77, Lions '87). Specific algorithms which are very efficient (see, e.g. Cancès-Le Bris 1998).
- Existence of a minimizer proved in a special subcase of $K = N + 2$ by Le Bris in 1994.
- In [ML], existence of a minimizer for practical methods (CASSCF).
- Numerically: Newton-type method (Werner-Meyer-Knowles, Yeager-Jørgensen et al, etc), '80s.

II. Nonlinear excited states

A definition of nonlinear excited states ?

The definition of nonlinear excited states is problematic: due to the nonlinearity, the functional $\mathcal{E}_N^K$ has many spurious stationary points !

For the d$^{\text{th}}$ excited state:

- *(1st order condition)* it should be a critical point of $\mathcal{E}_N^K$ (=solution of the MC equations) ;
- *(2nd order condition)* its **total** second derivative should have at most d negative eigenvalues (Morse index $\leq d$) see, e.g. Jørgensen-Olsen-Yeager, *J. Chem. Phys.* **75** (1981).
  This in particular implies that c should be **at most** the $(d+1)^{\text{th}}$ eigenvalue of H_Φ;
- *(large K behaviour)* its energy λ_d^K should satisfy $\lambda_d^K \geq \lambda_d$ and $\lambda_d^K \to \lambda_d$ as $K \to \infty$. The associated wavefunction Ψ_d^K should converge to the true excited state $\bar{\Psi}_d$, solution of $H_N \bar{\Psi}_d = \lambda_d \bar{\Psi}_d$.

The natural definition introduced in the 80s

"The MCSCF energy results from minimizing the appropriate eigenvalue of the hamiltonian matrix with respect to orbital variations" (Shepard '87)

For any fixed Φ, denote by $\mu_d^K(\Phi)$, $d = 0...\binom{K}{N} - 1$, the eigenvalues of the hamiltonian matrix H_Φ. By the usual Rayleigh-Ritz formula, one has

$$\mu_d^K(\Phi) \geq \lambda_d \qquad \text{(Hylleraas-Undheim-MacDonald Thm)}$$

This leads to (Löwdin '59, Werner-Meyer '80, Werner-Knowles '84, Shepard '87):

$$\boxed{\mu_d^K := \inf_\Phi \mu_d^K(\Phi).}$$

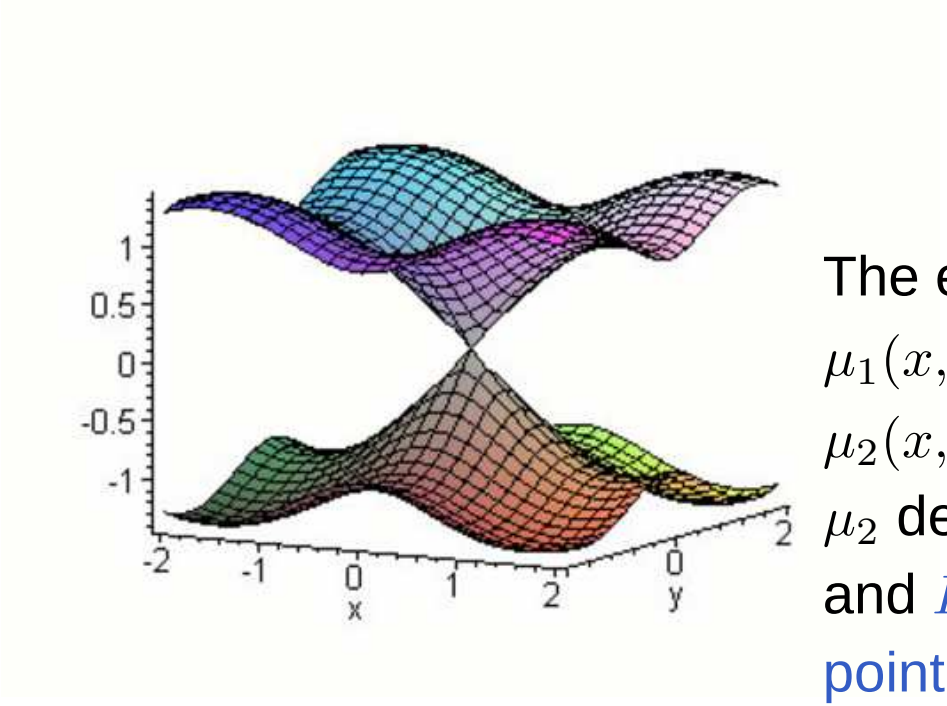
Easy to prove: $\mu_d^K \searrow \lambda_d$ as $K \to \infty$.

But: optimization of the eigenvalue of a matrix depending on a parameter = very ill-posed mathematically !

There could be no critical point at the level μ_d^K

The minimization of the eigenvalue of a matrix depending on a parameter $A(\mathbf{x})$ does not necessarily furnish a critical point of the associated energy $(\mathbf{x}, v) \mapsto \langle A(\mathbf{x})v, v\rangle$. In our case $\mathbf{x} \leftrightarrow \Phi$ and $v \leftrightarrow c$.

Example (Rellich): $E(\mathbf{x}, v) = \langle A(\mathbf{x})v, v\rangle$ with $\mathbf{x} = (x, y) \in \mathbb{R}^2$, $v \in S^1$ and



$$A(x, y) = \begin{pmatrix} -x & y \\ y & x \end{pmatrix}.$$

The eigenvalues of $A(x, y)$ are $\mu_1(x, y) = -\sqrt{x^2 + y^2}$ and $\mu_2(x, y) = \sqrt{x^2 + y^2}$.

μ_2 degenerates at its minimum $\mathbf{x} = 0$ and E does not have any critical point of the form $(0, v) \in \mathbb{R}^2 \times S^1$.

Computational issues

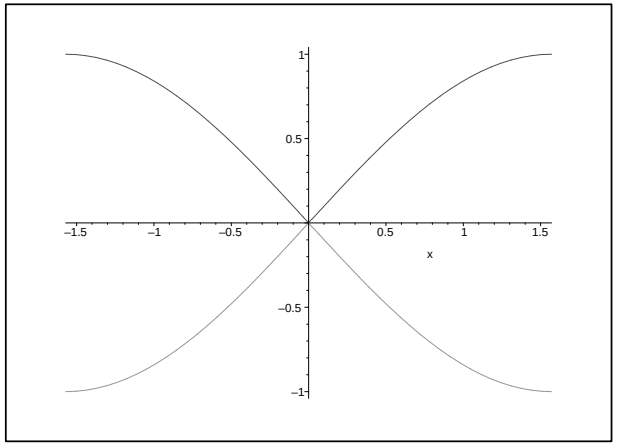
At present, there does *not* exist any general and efficient algorithm for eigenvalue minimization, except in very particular cases !

Naive algorithm ("two step" method):

1) Φ being fixed, take c' = $(d+1)^{\text{th}}$ eigenvector of H_Φ;

2) c' being fixed, find Φ' = minimizer of $\tilde{\Phi} \mapsto \mathcal{E}_N^K(c', \tilde{\Phi})$ and go back to 1).

$\longrightarrow$ does not always converge and can oscillate (root flipping).



Example: take $E(x,v) = \langle A(x)v, v\rangle$ with $(x,v) \in \mathbb{R} \times S^1$ and

$$A(x) = \begin{pmatrix} -\sin(x) & 0 \\ 0 & \sin(x) \end{pmatrix}.$$

$\longrightarrow$ oscillations between $x = -\pi/2$ and $x = \pi/2$, whereas $\mu_2 = 0$.

In practice

Root flipping issues are frequently encountered in MCSCF calculations.

In many Quantum Chemistry programs: an improved version of the previous algorithm (Cheung-Elbert-Ruedenberg ’79, Werner-Meyer ’80, Werner-Knowles ’85, Shepard ’87).

- Symmetry considerations are used to avoid degeneracies;
- one can optimize an average of different eigenvalues (Docken-Hinze ’72, Werner-Meyer ’81).

DALTON team: excited state = stationary state whose 2nd derivative has the right number of negative eigenvalues.

$\longrightarrow$ *one step* Newton-like method which is well-behaved.

But states with an energy $< \lambda_d$ are sometimes obtained !
(H.J.Å. Jensen, private communication)

Abstract definition of the excited states

Theorem 2. [ML] *Let be $Z > N-1$. For any $K \geq N$, there exists a sequence $(c_d, \Phi_d)_{d=0}^{\binom{K}{N}-1}$ of critical points (with a Morse index at most d) of $\mathcal{E}_N^K$ on $\mathcal{M}_N^K$ which are such that, denoting $\lambda_d^K = \mathcal{E}_N^K(c_d, \Phi_d)$,*

$$\lambda_d \leq \lambda_d^K \leq \mu_d^K, \quad \text{and} \quad \lim_{K\to\infty} \lambda_d^K = \lambda_d.$$

Sch.	$K = +\infty$	λ_0	λ_1	$\cdots$	$\cdots$	$\cdots$	λ_d	$\cdots$	$\cdots$
MC	$\vdots$	$\uparrow$	$\uparrow$				$\uparrow$		
	K	λ_0^K	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\lambda^K_{\binom{K}{N}-1}$	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$				
	$K = N+2$	λ_0^{N+2}	λ_1^{N+2}	$\cdots$	$\lambda^{N+2}_{\binom{N+2}{N}-1}$				
HF	$K = N$	λ_0^N							

Construction of the excited states

In general: λ_d^K = a complicated nonlinear min-max method.

Case of the first excited state energy λ_1^K: define

$$\Gamma_{(c,\Phi)} = \left\{\gamma \in \mathcal{C}^0([0;1], \mathcal{M}_N^K) \mid \gamma(0) = (c,\Phi),\ \gamma(1) = (-c,\Phi)\right\}$$

then

$$\boxed{\lambda_1^K = \inf_{(c,\Phi)\in\mathcal{M}_N^K} \left(\inf_{\gamma\in\Gamma_{(c,\Phi)}} \sup_{t\in[0;1]} \mathcal{E}_N^K(\gamma(t)) \right)}$$

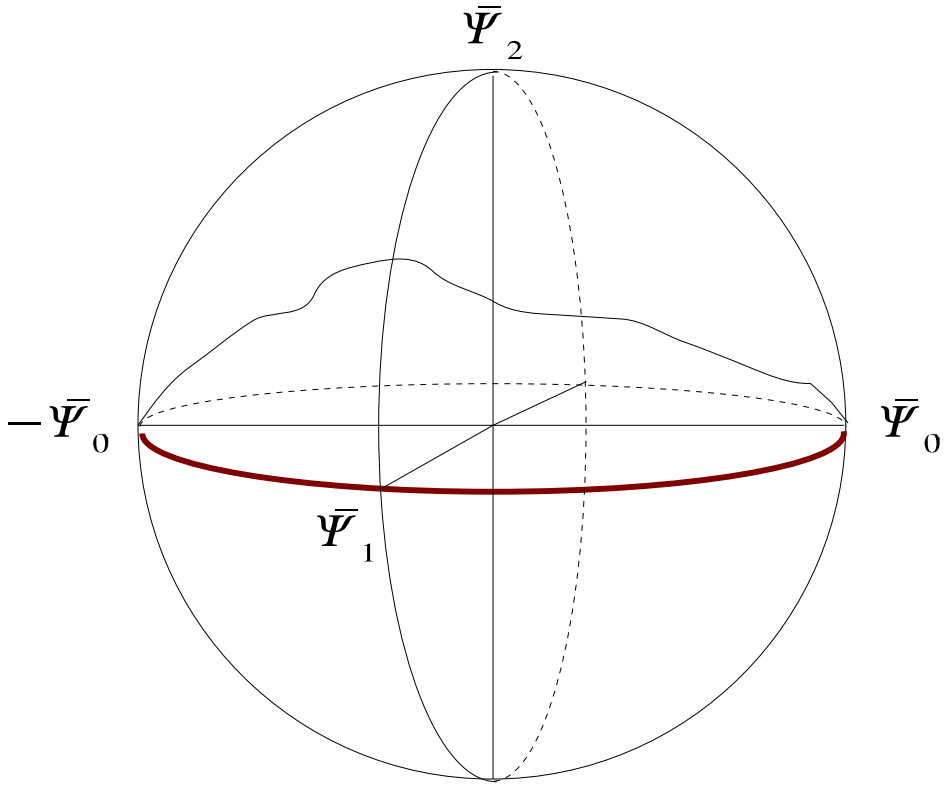
Conjecture: the first inf is attained for a *global* minimizer $(\bar{c}, \bar{\Phi})$ (at least if $K \gg 1$).

$\Rightarrow$ the first excited state is obtained as a mountain pass point (=transition state) between $(\bar{c}, \bar{\Phi})$ and $(-\bar{c}, \bar{\Phi})$!

Interpretation of λ_1^K

λ_1^K = mountain pass energy between $(\bar{c}, \bar{\Phi})$ and $(-\bar{c}, \bar{\Phi})$.

For the linear Schrödinger case: the true λ_1 can also be obtained as a mountain pass point between $\bar{\Psi}_0$ and $-\bar{\Psi}_0$.



A new algorithm

New algorithm for the first excited state [CGL]:

1. find $(\bar{c}, \bar{\Phi})$, a global minimizer of $\mathcal{E}_N^K$ (Newton-like method);
2. find an initial path γ_0 linking $(\bar{c}, \bar{\Phi})$ and $(-\bar{c}, \bar{\Phi})$. For instance $\gamma_0(t) = (c(t), \bar{\Phi})$ with $c(t) = \cos(\pi t)\bar{c} + \sin(\pi t)\bar{c}'$ where $\bar{c}'$ is the second eigenvector of the hamiltonian matrix $H_{\bar{\Phi}}$.
   (can then be perturbed randomly);
3. deform γ_0 to solve the mountain pass problem between $(\bar{c}, \bar{\Phi})$ and $(-\bar{c}, \bar{\Phi})$;
4. when the highest point on the path has a small-enough derivative, switch to a Newton algorithm.

This algorithm is guaranteed to converge to a critical point; it cannot oscillate between states.

But the optimization of paths is quite slow (first order method).

III. A simple test case: computation of the first excited state of two-electron systems

The singlet state of the H_2 molecule

H_2: two nuclei at $(-R/2, 0, 0)$ and $(R/2, 0, 0)$.

Singlet state: the antisymmetry is in the spin variable,

$$\Psi(x, \sigma; y, \sigma') = \psi(x, y)|\alpha\beta\rangle(\sigma, \sigma')$$

where ψ is symmetric, $\psi(x, y) = \psi(y, x)$.
The MCSCF approximation is applied to $\psi(x, y)$ similarly to the antisymmetric case.
The N-body Hamiltonian H_N only acts on the space wavefunction $\psi(x, y)$.

Symmetry considerations: H_N commutes with the operator τ defined as $(\tau\psi)(x, y) = \psi(-x, -y) \Longrightarrow H_N$ and τ can be diagonalized simultaneously ! τ has two eigenspaces:

$\Sigma_g := \{\psi \mid \psi(-x, -y) = \psi(x, y)\}$ and $\Sigma_u := \{\psi \mid \psi(-x, -y) = -\psi(x, y)\}$.

Symmetry

It is “known" in Quantum Chemistry that

- the ground state is in Σ_g
- the first excited state is in Σ_u.

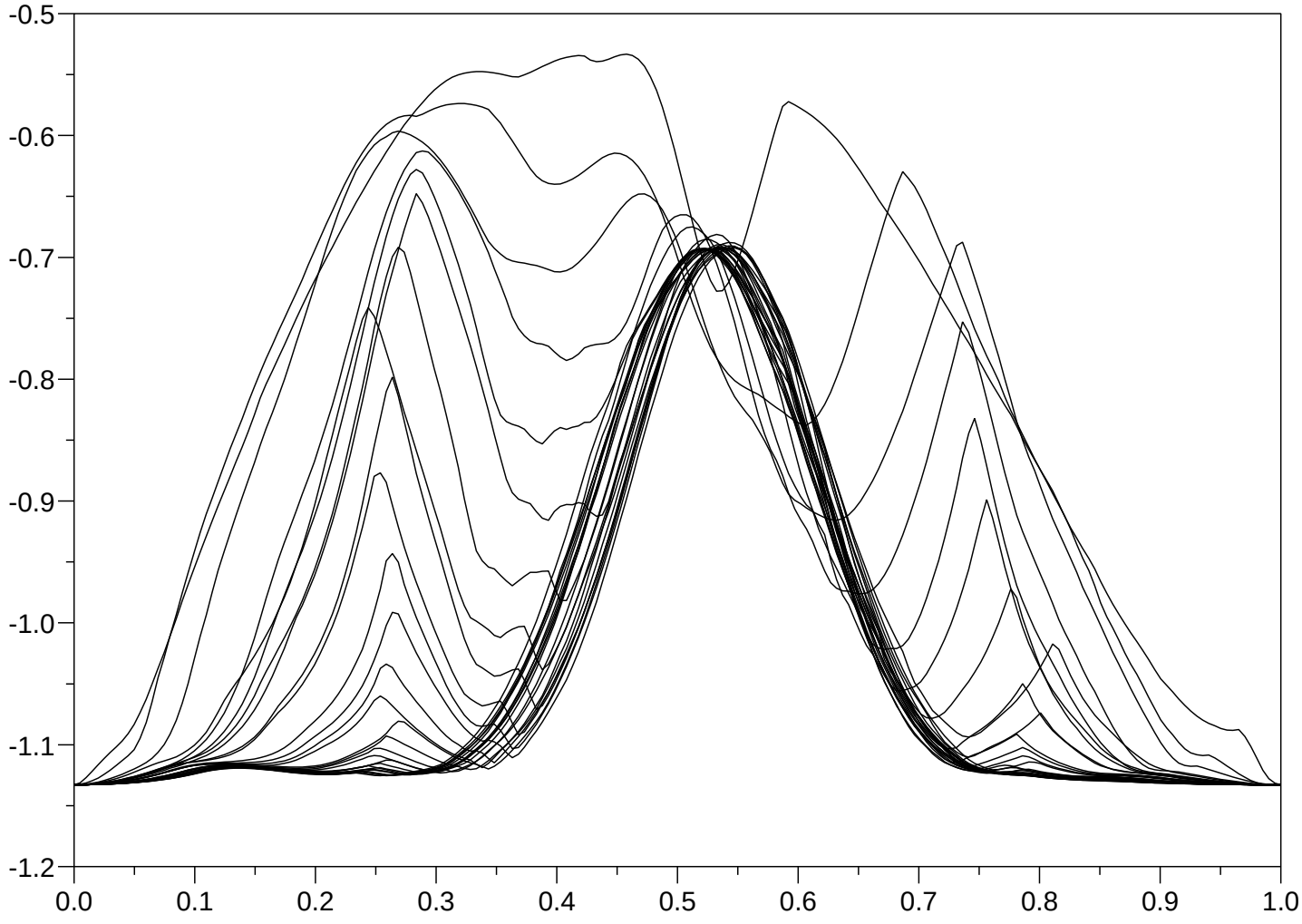
In practice, the first excited state is obtained as a minimum of the Σ_u symmetry.

$\rightarrow$ To illustrate the possible difficulties of the MCSCF method, we shall **not** impose any symmetry on the wavefunction.

McCourt-Iver, *J. Comp. Chem.* 1987: they argued that in this case the definition μ_1^K does *not* provide the true first excited state.

Rmk: shows that the computation for complicated molecules can be very tricky !

First singlet excited state of H_2 ($R = 1$ Å)



Form of c for the ground and excited states

Basis size (CC-PVDZ): $N_b = 10$. Number of determinants: $N_{\rm det} = 4$.

$$c_{\rm min} = \begin{bmatrix} 0.9860929 \\ -0.1564182 \\ -0.0548179 \\ -0.0122131 \end{bmatrix} \qquad c_{\rm 1ex} = \begin{bmatrix} -0.7086355 \\ 0.7051798 \\ 0.0166917 \\ 0.0166917 \end{bmatrix}$$

(H_2 molecule with an interatomic distance $R = 1$ Å)

- the ground state can be suitably approximated by a Hartree-Fock state for $R = 1$ Å ;
- for the first excited state, MCSCF methods are really unavoidable !

Symmetry of the first nonlinear excited state

We found that:

- the first nonlinear excited state $\Psi_{1\text{ex}}$ is *very close* to have the symmetry Σ_u;
  $\rightarrow$ *symmetry breaking* due to the nonlinearity !
- $c_{1\text{ex}}$ is the *first* eigenvector of the hamiltonian matrix $H_{\Phi_{1\text{ex}}}$!
  (McCourt-Iver *J. Comp. Chem.* 1987)
  $\rightarrow$ the Morse index is "carried on" by the orbitals and not c !
- one has $\lambda_1^K < \mu_1^K$;
  the stationary state obtained by μ_1^K is a spurious solution in the Σ_g symmetry subspace;
  $\rightarrow$ *non validity* of the μ_1^K definition if no symmetry conditions are imposed !

Rmk: no degeneracy problem in this case.

The PES of the H_2 molecule