Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 212 (2006) 73–98
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Computing electronic structures: A new multiconfiguration approach for excited states

Éric Cancès [a,b,*], Hervé Galicher [a], Mathieu Lewin [a,b]

[a] *CERMICS, École Nationals des Ponts et Chaussées, 6&8 Avenue Blaise Pascal, Cité Descartes, 77455 Marne-La-Vallée Cedex 2, France*
[b] *INRIA Rocquencourt, MICMAC Project, Domaine de Voluceau, B.P. 105, 78 150 Le Chesnay, France*

Received 16 September 2004; received in revised form 23 June 2005; accepted 27 June 2005
Available online 10 August 2005

Abstract

We present a new method for the computation of electronic excited states of molecular systems. This method is based upon a recent theoretical definition of multiconfiguration excited states [due to one of us, see M. Lewin, Solutions of the multiconfiguration equations in quantum chemistry, Arch. Rat. Mech. Anal. 171 (2004) 83–114]. Our algorithm, dedicated to the computation of the first excited state, always converges to a stationary state of the multiconfiguration model, which can be interpreted as an approximate excited state of the molecule.

The definition of this approximate excited state is variational. An interesting feature is that it satisfies a non-linear Hylleraas–Undheim–MacDonald type principle: the energy of the approximate excited state is an upper bound to the true excited state energy of the N-body Hamiltonian.

To compute the first excited state, one has to deform paths on a manifold, like this is usually done in the search for transition states between reactants and products on potential energy surfaces. We propose here a general method for the deformation of paths which could also be useful in other settings.

We also compare our method to other approaches used in Quantum Chemistry and give some explanation of the unsatisfactory behaviours which are sometimes observed when using the latters.

Numerical results for the special case of two-electron systems are provided: we compute the first singlet excited state potential energy surface of the H_2 molecule.
© 2005 Elsevier Inc. All rights reserved.

PACS: 31.25.Jf; 31.50.Df; 31.25.Nj; 02.60.Cb

Keywords: Multiconfiguration method; Excited state; Time-independent Schrödinger equation; Quantum chemistry; Mountain pass method; Minimax principle; Hartree–Fock theory; Configuration–Interaction method

* Corresponding author. Tel.: +33 1 64 15 35 69; fax: +33 1 64 15 35 86.
E-mail addresses: cances@cermics.enpc.fr (É. Cancès), galicher@cermics.enpc.fr (H. Galicher), lewin@cermic.enpc.fr (M. Lewin).

0021-9991/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2005.06.015

Electronic excited states play an essential role in various phenomena of high interest, such as photo-induced chemical reactions, femtosecond spectroscopy, or laser control of molecular processes. Whereas most of the currently used electronic structure models, notably the Hartree–Fock and the Kohn–Sham models, are rigorously founded and quite successful in the description of ground states, their approach to excited states is questionable [1]. The method which seems to be best-adapted to this issue is to date the multiconfiguration self-consistent field (denoted by MCSCF in the following) method [2–4]; loosely speaking, this approach leads to variational models which fill the gap between the mean-field Hartree–Fock and the N-body Schrödinger models [3]. However, the definition of what actually is an excited state for a non-linear theory such as MCSCF is still unclear; it is indeed observed that non-linear electronic structure models have a lot of spurious critical points that cannot be interpreted as approximations of excited states. In other words, solving the equations of the model is clearly not sufficient to obtain a state which really approximates some excited state. In addition, even if we leave aside the above mentioned difficulty, the practical calculation of MCSCF critical points is difficult and some numerical algorithms available to date do not always converge. Even if they converge, the interpretation of the obtained solution is not always clear. For all these reasons, the computation of electronic excited states remains one of the main challenges of modern Quantum Chemistry.

In [5], it is emphasized that those difficulties are likely to stem from the currently used definitions of MCSCF excited states that are not correct, for they do not fully take into account the non-linearity of the model. The purpose of [5] was to provide a more rigorous definition of MCSCF excited states. Our goal in this paper is to show that this theoretical definition can actually be used in practice, at least for the computation of the *first* excited state.

The paper is organized as follows. In Section 1, we introduce the MCSCF description of electronic structures. In Section 2, we present the new definition of MCSCF excited states and compare it to other definitions currently used in Computational Chemistry. Finally, in Section 3, we describe in details our new algorithm and present numerical results for the case of two-electron systems.

1. MCSCF approximation of the time-independent Schrödinger equation

In this section, we recall some classical properties of the N-body time-independent Schrödinger equation, and briefly present the MCSCF approximation. We refer the reader to [4–8] for more details.

Let us consider a molecular system consisting of N electrons, and of M nuclei of positive charges $z_1,\ldots,z_M$. The nuclei are supposed to be correctly described by a classical model and are represented by pointwise charges clamped at positions $\bar{x}_1,\ldots,\bar{x}_M$ ($\bar{x}_m \in \mathbb{R}^3$ for $1 \leqslant m \leqslant M$). This is the so-called Born–Oppenheimer approximation [9]. The electrons are described by the N-body quantum Hamiltonian (written in atomic units, see, e.g. [8])

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

which acts on normalized electronic wavefunctions $\Psi(x_i,\ldots,x_N) \in L^2_a((\mathbb{R}^3)^N)$, $\|\Psi\|_{L^2} = 1$. The subscript a indicates that, due to the fermionic nature of the electrons, one solely considers wavefunctions Ψ which are antisymmetric under permutations of variables

$$\forall \sigma \in S_N, \quad \Psi(x_1,\ldots,x_N) = \varepsilon(\sigma)\Psi\left(x_{\sigma(1)},\ldots,x_{\sigma(N)}\right)$$

almost everywhere. Here and below, S_N denotes the set of the permutations of the indices $\{1,\ldots,N\}$ and $\varepsilon(\sigma)$ the signature of the permutation σ. Finally, V is the electrostatic potential generated by the nuclei

$$V(x) = -\sum_{m=1}^{M} \frac{z_m}{|x - \bar{x}_m|}.$$

In what follows, we denote by $Z = \sum_{m=1}^{M} z_m$ the total nuclear charge which is an integer as we work in atomic units.

For the sake of clarity, we do not take the spin into account in the first two sections of the article, but the following arguments can be straightforwardly adapted to the case of spin-dependent wavefunctions. The spin will be reintroduced in Section 3, in which numerical examples on real molecular systems will be provided.

The operator H_N is self-adjoint in $L^2_a((\mathbb{R}^3)^N)$, with domain $H^2_a((\mathbb{R}^3)^N)$ and form domain $H^1_a((\mathbb{R}^3)^N)$. When $Z > N - 1$ (an assumption that we will make throughout this article), it is known [10] that its spectrum $\sigma(H_N)$ has the form

$$\sigma(H_N) = \{E_N = \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n \leqslant \cdots\} \cup [\Sigma; +\infty),$$

where $(\lambda_i)_{i \geqslant 1}$ are eigenvalues strictly below and which converge to Σ, the bottom of the essential spectrum. The *N-body ground state energy* is the lowest eigenvalue of H_N also defined by

$$E_N = \inf \left\{ \langle \Psi, H_N \Psi \rangle, \qquad \Psi \in H^1_a(\mathbb{R}^{3N}), \quad \|\Psi\|_{L^2_a(\mathbb{R}^{3N})} = 1 \right\}. \tag{2}$$

The eigenfunctions corresponding to the $\lambda_i > E_N$ are called *excited states*. Both the ground states and the excited states obviously solve the *time-independent Schrödinger equation*

$$H_N \Psi = \lambda_i \Psi. \tag{3}$$

Recall that the excited state energies λ_d, $d \geqslant 1$, can be obtained by the Rayleigh–Ritz principle

$$\lambda_d = \inf_{\dim(W)=d} \max_{\substack{\Psi \in W, \\ \|\Psi\|_{L^2}=1}} \langle \Psi, H\Psi \rangle, \tag{4}$$

where the first infimum is taken over all d-dimensional subspaces W of the domain of H_N.

The Schrödinger equation is a model of extremely high accuracy, except for heavy atoms for which core electrons are relativistic. For systems involving a few (say today six or seven) electrons, a direct Galerkin discretization of problem (3) is possible; such a technique is referred to as *Full CI* in Computational Chemistry. For larger systems, this direct approach is out of reach, due to the excessive dimension of the space $\mathbb{R}^{3N}$ on which the wavefunctions are defined, and problem (3) must then be approximated. To date, the most commonly used approximations are the Hartree–Fock model (see, e.g. [11]) on the one hand, and the Kohn–Sham model (see, e.g. [12,13]) on the other hand. Both of them have been designed for the calculation of ground states and are not really adapted to the calculation of excited states. On the contrary, the MCSCF approximation can be applied to both ground and excited state calculations.

The MCSCF method is based on the following remark:

$$L^2_a((\mathbb{R}^3)^N) = \bigwedge_{n=1}^{N} L^2(\mathbb{R}^3),$$

an equality which can be explicited in the following way. Consider an orthonormal basis $(\varphi_i)_{1 \leqslant i < +\infty}$ of $L^2(\mathbb{R}^3)$. It is well known that the sequence $(\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_N})_{1 \leqslant i_k < \infty}$ forms an orthonormal basis of $L^2((\mathbb{R}^3)^N) = \otimes_{n=1}^{N} L^2(\mathbb{R}^3)$, where by definition

$$(\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_N})(x_1, \ldots, x_N) = \varphi_{i_1}(x_1) \cdots \varphi_{i_N}(x_N).$$

An orthonormal basis of the subspace $L^2_a((\mathbb{R}^3)^N)$ of $L^2((\mathbb{R}^3)^N)$ can then be obtained by simply considering the antisymmetrized products $(\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N})_{1 \leqslant i_1 < \cdots < i_N < +\infty}$, where $\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N}$ denotes the so-called Slater determinant of the φ_{i_k}'s:

$$(\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N})(x_1, \ldots, x_N) = \frac{1}{\sqrt{N!}} \sum_{\sigma \in \mathcal{S}_N} \varepsilon(\sigma) \varphi_{i_1}(x_{\sigma(1)}) \cdots \varphi_{i_N}(x_{\sigma(N)}) = \frac{1}{\sqrt{N!}} \det(\varphi_{i_k}(x_l))_{k,l}.$$

In other words, every antisymmetric wavefunction Ψ is an infinite linear combination of such Slater determinants

$$\Psi = \sum_{1 \leqslant i_1 < \cdots < i_N < +\infty} c_{i_1 \ldots i_N} \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N},$$

the sum being convergent in $L^2_a((\mathbb{R}^3)N)$. Remark that $\|\Psi\|_{L^2} = 1$ is then equivalent to the condition $\sum_{i_1 < \cdots < i_N} |c_{i_1 \ldots i_N}|^2 = 1$.

An integer $K \geqslant N$ being fixed, we now consider the subset of $L^2_a((\mathbb{R}^3)^N)$ consisting of the wavefunctions Ψ which are *finite* linear combinations of the $\binom{K}{N}$ Slater determinants constructed from a set of K orthonormal functions $(\varphi_1, \ldots, \varphi_K)$ of $L^2(\mathbb{R}^3)$, i.e.,

$$\Psi = \sum_{1 \leqslant i_1 < \cdots < i_N \leqslant K} c_{i_1 \ldots i_N} \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N}. \tag{5}$$

The MCSCF approach is a variational method for approximating (3) in which both the coefficients $c_{i_1 \ldots i_N}$ and the functions $(\varphi_1, \ldots, \varphi_K)$ are variational parameters. Let us mention incidently that the MCSCF method differs from the Configuration–Interaction (CI) method [14], for in the latter, only the coefficients $c_{i_1 \ldots i_N}$ are variational parameters (in a CI calculation, the functions $(\varphi_1, \ldots, \varphi_K)$ are issued from a previous Hartree–Fock or Kohn–Sham calculation and are kept fixed). When there is no ambiguity, we shall use the following notation:

$$\Psi = \sum_{I \subset \{1, \ldots, K\}, |I| = N} c_I \Phi_I,$$

where $\Psi_I = \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N}$, when $I = \{i_1 < \cdots < i_N\}$.

Following our purpose to describe the MCSCF approach, we introduce the manifold

$$\mathcal{M}_N^K = \left\{ (c, \Phi) \in \mathbb{R}^{\binom{K}{N}} \times \left(H^1(\mathbb{R}^3)\right)^K, \quad \sum_{i_1 < \cdots < i_N} |c_{i_1 \ldots i_N}|^2 = 1, \quad \int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij} \right\}, \tag{6}$$

where we have used the notation

$$c = (c_{i_1 \cdots i_N}) \in \mathbb{R}^{\binom{K}{N}}, \quad \Phi = (\varphi_1, \ldots, \varphi_K) \in H^1(\mathbb{R}^3)^K$$

(we arrange the $c_{i_1 \ldots i_N}$ in a column vector c using for instance the lexicographical order). Let us note that the functions $(\varphi_1, \ldots, \varphi_K)$ are now requested to have a H^1 regularity, in order to ensure that the MCSCF energy is well defined. The MCSCF energy functional that we denote here by $\mathcal{E}_N^K$, is defined by the formula

$$\mathcal{E}_N^K(c, \Phi) = \langle \Psi_{(c,\Phi)}, H_N \Psi_{(c,\Phi)} \rangle, \quad \Psi_{(c,\Phi)} = \sum_{1 \leqslant i_1 < \cdots < i_N \leqslant K} c_{i_1 \ldots i_N} \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N} \tag{7}$$

and the MCSCF ground state energy then reads

$$E_N^K = \inf_{\mathcal{M}_N^K} \mathcal{E}_N^K. \tag{8}$$

An explicit expression of the non-linear functional $\mathcal{E}_N^K$ can be found in [5, Eq. (6)].

Let us point out that, whereas the Schrödinger energy functional $\Psi \mapsto \langle \Psi, H_N \Psi \rangle$ is quadratic, the MCSCF energy functional is not. Consequently, the MCSCF equations, namely the first order stationarity conditions for the critical points of $\mathscr{E}_N^K$ on the manifold $\mathscr{M}_N^K$, will be non-linear. More precisely, $\mathscr{E}_N^K$ is not quadratic with respect to the orbitals φ_i's, but it is indeed quadratic with respect to the c_I's since

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

where (recall that $\Phi_I = \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N}$, when $I = \{i_1 < \cdots < i_N\}$)

$$(H_\Phi)_{IJ} = \langle \Phi_I, H_N \Phi_J \rangle. \tag{9}$$

In other words, H_Φ is the $\binom{K}{N} \times \binom{K}{N}$ matrix of the quadratic form associated with H_N when it is restricted to the $\binom{K}{N}$-dimensional space $V_\Phi = \mathrm{Span}(\Phi_I)$. It can be seen that the *MCSCF equations* take the following general form [6,5]:

$$\begin{cases} \gamma_i \left(-\frac{\Delta}{2} + V\right) \varphi_i + \sum\limits_{1 \leqslant j,k,l \leqslant K} b_{ijkl} \left((\varphi_j \varphi_k) * \frac{1}{|x|} \right) \varphi_l = \sum\limits_{j=1}^{K} \lambda_{ij} \varphi_j, \quad 1 \leqslant i \leqslant K, \\ H_\Phi c = \beta c, \end{cases} \tag{10}$$

where the b_{ijkl} are real numbers which can be expressed in terms of c. The first line of (10) is in fact a system of K non-linear coupled partial differential equations accounting for the stationarity conditions with respect to Φ; the symmetric matrix (λ_{ij}) is the Lagrange multiplier matrix associated with the orthonormality constraints on Φ. The numbers γ_i are called the *occupation numbers* and satisfy $0 \leqslant \gamma_i \leqslant 1$ (see [5] for details). A compact form of the first equations of (10) is given in [5]. The second equation is a simple eigenvalue problem and conveys the stationarity condition with respect to c.

When $K = N$, Ψ is a single Slater determinant and one recovers the celebrated Hartree–Fock approximation [11,15,16]. In this case, (10) can be written in a simpler way, using the invariance by rotation of the orbitals. The difference between the Hartree–Fock and the exact (non-relativistic) ground state energy

$$E^{\mathrm{corr}} = E_N^N - E_N$$

is called the *correlation energy* [3], for it originates from correlations between the positions of individual electrons, which are averaged out by the mean-field Hartree–Fock scheme. Estimating the correlation energy is essential for reliably calculating many of the properties of molecules [7,1], in particular in situations where the Hartree–Fock method fails. Since

$$\lim_{k \to +\infty} E_N^K = E_N,$$

the MCSCF method is a method of choice for computing the correlation energy.

Mathematically, it is known that a minimizer of (8) exists, and that the associated wavefunction converges to the ground state of H_N as K goes to infinity [17,6,5]. A minimizer of (8) is usually numerically computed by a Newton-like algorithm, sometimes improved by a trust-region method [18–20,4,21–26]. For the Hartree–Fock model, efficient numerical methods based on combinations of fixed-point and optimization strategies are available [8]. Unfortunately, such algorithms are specifically designed for solving the Hartree–Fock problem and seem to be difficult to adapt to the more general MCSCF setting.

Remark that in (5), all the Slater determinants that can be built with the functions φ_i are taken into account. Most often, this cannot be done in practice for $\binom{K}{N}$ is too large a number. It is then necessary to resort to an additional approximation consisting in dividing the electrons into two groups, the *inactive* electrons that are supposed to be correctly described by a Hartree–Fock type model, and the *active* electrons that mostly contribute to the correlation energy, and in using the MCSCF methodology for the active

electrons only. This is the so-called Complete Active Space Self-Consistent Field (CASSCF) approach [27]. All what we shall mention here can be straightforwardly adapted to the CASSCF setting. In particular, the first excited state of a CASSCF model can be computed using a slightly modified version of the numerical algorithm presented in Section 2.3.

2. On the definition of MCSCF excited states

Numerical investigations show that the MCSCF energy $\mathscr{E}_N^K$ possesses a lot of critical points on the manifold $\mathscr{M}_N^K$, and it is not obvious to characterize, among all these critical points, those which can be regarded as approximate excited states. As explained below, the following three properties can be considered as necessary conditions for $(c,\Phi) \in \mathscr{M}_N^K$ being an approximate jth excited state:

(a) *First order condition:* (c,Φ) is a critical point of the energy functional $\mathscr{E}_N^K$, that is to say a solution of the MCSCF equations (10).
(b) *Second order condition:* (c,Φ) has a Morse index[1] equal to j or, saying differently, its *total* hessian matrix (with respect to both c and Φ) has j negative eigenvalues.
(c) *Non-linear Hylleraas–Undheim–MacDonald type theorem:* the energy of (c,Φ), denoted as λ_{j+1}^K satisfies

$$\lambda_{j+1}^K \geqslant \lambda_{j+1} \quad \text{and} \quad \lim_{k\to\infty} \lambda_{j+1}^K = \lambda_{j+1}, \tag{11}$$

where we recall that λ_{j+1} is the $(j+1)$th eigenvalue of the N-body Hamiltonian H_N.

The first condition (a) is very natural and simply means that the approximate excited state is a stationary state of the model. The second condition (b) is also natural; it has been proposed and studied in Quantum Chemistry in [28,23]. Since H_Φ defined in (9) is the second derivative of the functional $\mathscr{E}_N^K$ *with respect to c only*, a consequence of (b) and (10) is that c is at most the $(j+1)$th eigenvector of the Hamiltonian matrix H_Φ. In other words, β in (10) is at most the $(j+1)$th eigenvalue of H_Φ, but may correspond to a lower eigenvalue [28,23].

The third condition (c) is a generalization of the result claiming that, in the linear case, the eigenvalue of a quadratic form restricted to a subspace is greater than the corresponding true eigenvalue on the whole space. This result, obvious consequence of the Rayleigh–Ritz principle (4), is usually called the Hylleraas–Undheim–MacDonald (HUM) theorem [29,30] in Quantum Chemistry.

We are aware of two different definitions of approximate excited states in Quantum Chemistry. They are discussed and compared for instance in [28,31,24]. We will show in the following section that they do not always provide solutions satisfying the conditions (a)–(b)–(c). The mathematical definition of [5] which we shall use to construct our new method will then be presented in Section 2.2. This definition allows to construct specific MCSCF states satisfying the three conditions (a)–(b)–(c).

2.1. Two definitions currently used in Quantum Chemistry

2.1.1. The standard definition of MCSCF excited states

Let us start with the standard definition which is still mostly used today, and is based on the idea that condition (c) should be enforced. To this end, we denote by $\mu_d^K(\Phi),\ d = 1 \ldots \binom{K}{N}$, the eigenvalues of the matrix H_Φ, which has been defined in (9). Let us recall that this is the matrix of the quadratic form

[1] Recall that the Morse index of a critical point is the number of negative eigenvalues of the Hessian matrix.

associated with the N-body Hamiltonian H_N, restricted to the space span $(\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N})$, where $\Phi = (\varphi_1, \ldots, \varphi_K)$ is a fixed set of orbitals. Then, one deduces from the Rayleigh–Ritz principle (4) that

$$\lambda_d \leqslant \mu_d^K(\Phi)$$

for all K and d such that $1 \leqslant d \leqslant \binom{K}{N}$, and any fixed set of orbitals $\Phi \in H^1(\mathbb{R}^3)^K$ such that $\int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij}$. This inequality suggested the mostly used current definition of approximate excited state energies [3,4,19,32,18,22]

$$\mu_d^K = \inf_{\substack{\Phi \in H^1(\mathbb{R}^3)^K \\ \int_{\mathbb{R}^3} \Phi \Phi^{\mathrm{T}} = I_K}} \mu_d^K(\Phi), \tag{12}$$

that is to say, quoting [4], "the MCSCF energy results from minimizing the appropriate eigenvalue of the Hamiltonian matrix with respect to orbital variations" . It can be shown [5] that such a definition indeed satisfies the Hylleraas–Undheim–MacDonald type condition (c), namely

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

This property has been the main argument in favour on the definition (12) in many chemical papers. However, we believe that this commonly admitted definition of MCSCF excited state energies is the source of various difficulties of both practical and theoretical natures, since the two other conditions (a)–(b) are not always satisfied. Reservations of the same kind have been expressed in [28,31,24,37].

Indeed, solving problem (12) amounts to minimizing the dth eigenvalue of a matrix depending on a set of parameters Φ. This is known to be a challenging task and no completely satisfactory numerical method dedicated to solving such problems is available to date, except for very special cases (for instance when the matrix linearly depends on the parameters, see, e.g. [33]). In fact, we shall see in the following paragraph that the algorithms which are currently implemented in the Quantum Chemistry simulation packages using definition (12) [19,34,20,32,18,4] are not fully adapted to this issue.

Serious difficulties can occur when optimizing $\mu_d^K(\Phi)$, due to a possible loss of differentiability of this function in case of degeneracies. As an illustration, let us simply mention a celebrated example due to Rellich and reported in [35]: consider the family of 2×2 matrices $(A(x,y))_{(x,y) \in \mathbb{R}^2}$ defined by

$$A(x,y) = \begin{pmatrix} -\sin x & \sin y \\ \sin y & \sin x \end{pmatrix} \tag{13}$$

with eigenvalues

$$\lambda_1(x,y) = -\sqrt{\sin^2(x) + \sin^2(y)} \quad \text{and} \quad \lambda_2(x,y) = \sqrt{\sin^2(x) + \sin^2(y)}.$$

The second eigenvalue λ_2 degenerates and is not differentiable at its minimum $(x_0, y_0) = (0,0)$. Moreover, it is easily seen that there exists no critical point of the form $(0,0,v) \in \mathbb{R}^2 \times S^1$, of the associated energy $(x,y,v) \in \mathbb{R}^2 \times S^1 \mapsto \langle A(x,y)v, v \rangle$.

Coming back to our main context, this exactly means that the conditions (a)–(b) are not necessarily fulfilled when the definition (12) is used: it could be that there does not exist any stationary state with energy μ_d^K when a degeneracy occurs.

Let us now make some comments on the numerical methods used to solve (12). Following [4,19,32,18] and loosely speaking, the general form of the numerical algorithms currently used to calculate the $(d-1)$th excited state can be summarized as follows:

1. start with some (c, Φ) obtained for instance from a previous Hartree–Fock or Configuration–Interaction calculation;
2. compute the matrix H_Φ;
3. find c' as the dth eigenvector of this matrix;
4. this c' being fixed, minimize the energy with respect to Φ to obtain a new Φ';
5. replace (c, Φ) by (c', Φ') and return to step 2.

The main difficulty with this so-called *two-step method* is that the energy is not necessarily decreasing during the computation; it can in fact oscillate, as this can be easily seen when this algorithm is applied to the following toy problem [5]: find the first excited state for the energy functional

$$\tilde{\mathscr{E}}(c, \Phi) = c^{\mathrm{T}} \begin{pmatrix} -\sin\Phi & 0 \\ 0 & \sin\Phi \end{pmatrix} c$$

with $c \in S^1$ and $\Phi \in]-\pi,\pi[$ (an oscillation between $\Phi = -\pi/2$ and $\Phi = \pi/2$ is obtained).

This phenomenon is called *root flipping* in Quantum Chemistry. It is well known and observed in practice in MCSCF calculations [36,34,32,28,31,24]. Many solutions have been proposed to avoid this drawback. First, the computation is always done in a restricted set by adding special requirements like space symmetry, which need to be intuited before starting the optimization. This choice seems to deeply influence the behaviour of the algorithm. This is not satisfactory from a numerical point of view: a method should not depend on the space on which the optimization is made and such a restriction should only be used as a tool to speed up the convergence or simplify the computation. Then, instead of (12), the average of different eigenvalues μ_d^K's is often considered [36,34]. Although this allows to avoid oscillations in many cases, this obviously results in a loss of accuracy and cannot be considered as a general solution for the problems described above concerning the definition (12).

Even when the definition (12) provides a critical point and no root-flipping is observed, the interpretation of the so-obtained N-body wavefunction is unclear. For the special case of two-electron systems (H_2 molecule and Helium-like atoms), our numerical results, presented in Section 3.3, seem to show that the state which is obtained by solving (12) without imposing any space symmetry is *not* an appropriate approximation of the first excited state. Indeed, it does not have the right space symmetry properties (in practice, the space symmetry is *imposed* during the computation to obtain the right approximate excited state). This issue has been raised for the first time in [37]. It means that (12) cannot be considered as a relevant definition in general: imposing that c is a specific eigenvector of H_Φ, may lead to unphysical results.

2.1.2. The DALTON method

The issues raised by the definition (12) have already been described and studied in details in [28,23,24,31] by the team of the DALTON software [38]. They proposed a different definition of excited states which consists in just imposing that conditions (a)–(b) hold, neglecting (c). In their approach, a Newton-like method is used to compute the state under consideration, imposing that the Hessian matrix has exactly $d-1$ negative eigenvalues [28,23,24,31,25,39,26]. This algorithm is extremely well-behaved and efficient. This is a *one step method* in which the orbitals φ_i's and c_I coefficients are optimized simultaneously. It does not suffer from root-flipping and always provides a critical point with the right Morse index.

However, one might ask why any critical point having the right Morse index should be interpreted as an approximate excited state, that is to say a state which is close to the true eigenfunctions of the N-body operator. In our simulations for the first singlet excited state of the H_2 molecule (see Section 3.3), we obtain two critical points with Morse index one. One of them turns out to be a good approximation of the first excited state, whereas the other one is a spurious state which has no physical interpretation (in fact, we believe that it is the solution of (12)). Moreover, it apparently sometimes happens that the state obtained by DALTON

does *not* satisfy the condition (c): it can have an energy which is *below* the correct one.[2] This is not surprising: the functional $\mathscr{E}_N^K$ has a lot of critical points of a given Morse index, and there is no reason for them to have an energy above the eigenvalues of H_N. In [28], an additional condition expressed in terms of the linear response is added but it is rarely checked on the solution found out by the algorithm.

As a conclusion, none of these two methods provides a state which for sure satisfies (a)–(b)–(c). We shall now present the result of [5] which does provide such a solution. Indeed, one might say that it provides the variational interpretation which is missing in the DALTON method.

2.2. A new definition of MCSCF excited states

In this section, we present the new definition of MCSCF excited states introduced in [5]. Let be

$$\mathscr{B}^{d-1} = \left\{ f(S^{d-1}) | f \in C^0(S^{d-1}, \mathscr{M}_N^K), \quad f(x) = (c, \Phi) \Rightarrow f(-x) = (-c, \Phi) \right\}. \tag{14}$$

The definition of an excited state energy used in [5] is

$$\lambda_d^K = \inf_{B \in \mathscr{B}^{d-1}} \sup_{(c,\Phi) \in B} \mathscr{E}_N^K(c, \Phi) \tag{15}$$

and the following result has been established

Theorem 1 (Existence of MCSCF excited states [5]). *Assume* $Z > N - 1$ *and* $1 \leqslant d \leqslant \binom{K}{N}$. *Then there exists a critical point* (c_d, Φ_d) *of the energy* $\mathscr{E}_N^K$ *on* $\mathscr{M}_N^K$, *with a Morse index lower than or equal to* $d - 1$, *and which satisfies* $\mathscr{E}_N^K(c_d, \Phi_d) = \lambda_d^K$. *Moreover,* λ_d^K *satisfies*

$$\lambda_d \leqslant \lambda_d^K \leqslant \mu_d^K \tag{16}$$

and therefore

$$\lim_{k \to \infty} \lambda_d^K = \lambda_d. \tag{17}$$

This result shows that contrarily to what occurs with the definition (12), one always obtains with (15) a critical point which is a solution of the MCSCF equations (10) and satisfies the conditions (a)–(b)–(c). In particular, one knows that the so-obtained energy λ_d^K is always an upper bound of the true energy. But (16) is a *non-linear* version of the celebrated Hylleraas–Undheim–MacDonald theorem which does not depend on the index of $c = (c_I)$ as an eigenvector the Hamiltonian matrix H_Φ. This is the total Hessian matrix which has the right Morse index, and not necessarily the one taking the variations with respect to c only.

Recall that μ_d^K has been defined in (12). We have no general criterion to decide whether the strict inequality $\lambda_d^K < \mu_d^K$ holds or not. Generally speaking, one can guess that it holds in practice when, due to a problem of degeneracy, no critical point exists at the level μ_d^K (λ_d^K is always a critical value by Theorem 1). More specifically, we believe that it holds for the special case of the H_2 molecule when the two-body Hamiltonian is not restricted to a particular symmetry, as suggested by our computations presented in Section 3.3 and claimed in [37]. For the Rellich example defined above (13), a simple calculation indeed shows that, with obvious notations, $-1 = \lambda_2 < \mu_2 = 0$.

Remark. In a non-interacting system, i.e., when the interaction term

$$\sum_{1 \leqslant i < j \leqslant N} \frac{1}{|x_i - x_j|}$$

[2] H.J.Aa. Jensen, private communication.

in the expression (1) of H_N is turned off, one can see that for any $d \leqslant \binom{K}{N}$, $\lambda_d^K = \lambda_d$ (i.e., the MCSCF and the Schrödinger excited state energies coincide). Moreover, the critical point which is found in Theorem 1 is precisely in this case the dth eigenfunction of the N-body Hamiltonian.

2.3. A new method for the computation of the first excited state

Let us emphasize that the definition (15) is valid for all the excited states for which $1 \leqslant d \leqslant \binom{K}{N}$. However, it has probably not much practical interest for large d due to its complicated formulation (when $d > 2$, one has to deform surfaces of dimension $d-1$). For the first excited state $d = 2$, it can indeed be transformed into a totally new and simple computational method that we present in this section.

Let us first clarify the structure of the set $\mathscr{B}^1$ defined in (14). Using the parametrization $t \in [0;\ 2] \to (\cos(\pi t), \sin(\pi t))$ of S^1, we see that a function satisfying the conditions of (14) can be written $t \in [0;2] \mapsto (c(t), \Phi(t)) \in \mathscr{M}_N^K$ with $c(1+t) = -c(t)$ and $\Phi(1+t) = \Phi(t)$. Since $\mathscr{E}_N^K$ is even with regards to c which means

$$\mathscr{E}_N^K(-c, \Phi) = \mathscr{E}_N^K(c, \Phi),$$

we obtain

$$\sup_{t\in[0;2]} \mathscr{E}_N^K(c(t), \Phi(t)) = \sup_{t\in[0;1]} \mathscr{E}_N^K(c(t), \Phi(t)).$$

Therefore, we can rewrite (15) as

$$\lambda_2^K = \inf_{(c,\Phi)\in\mathscr{M}_N^K} \left\{ \inf_{\gamma\in\Gamma_{(c,\Phi)}} \sup_{t\in[0;1]} \mathscr{E}_N^K(\gamma(t)) \right\}, \tag{18}$$

where

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

Notice that the inf–sup problem which is in brackets in (18) is a mountain-pass problem (between (c, Φ) and $(-c, \Phi)$), similar to those encountered in molecular simulation in the search for transition states between reactants and products on potential energy surfaces [40,41]. To compute the term in brackets, one thus has to deform paths, as this is usually done in the latter setting.

We now *conjecture* that when K is large enough, a global minimizer of the MCSCF energy $(\bar{c}, \bar{\Phi})$ is also a minimizer of the outer minimization in (18). Therefore, we are able to simplify the resolution of problem (18) as follows: we clamp both ends of the trial paths at $(\bar{c}, \bar{\Phi})$ and $(-\bar{c}, \bar{\Phi})$, respectively, and solve the mountain pass problem

$$\lambda_2^K = \inf_{\substack{\gamma\in C^0\left([0;1],\mathscr{M}_N^K\right) \\ \gamma(0)=(\bar{c},\bar{\Phi}),\gamma(1)=(-\bar{c},\bar{\Phi})}} \sup_{t\in[0;1]} \mathscr{E}_N^K(\gamma(t)). \tag{19}$$

We have no proof of the equality (19) but it will be fulfilled provided there exists a path linking $(\bar{c}, \bar{\Phi})$ and an actual minimizer of the outer minimization of (18), along which the energy does not exceed λ_2^K. It is indeed likely to be the case, at least for K large enough. Notice that (19) mimics a well-known formula which allows, in the linear case, to obtain the second eigenfunction Ψ_2 of H_N, as a mountain pass point between Ψ_1 and $-\Psi_1$, where $H_N\Psi_1 = \lambda_1\Psi_1$.

In practice, solving a mountain-pass problem is rather demanding in terms of CPU time since one has to deform paths. Therefore, we shall choose a not too tight convergence criteria to stop the path optimization step. The state of highest energy on the final path is then used as initial guess in a Newton-like procedure to

solve (10). Our algorithm to compute the first excited state can therefore be summarized as follows (details will be given in the following section):

1. use a Newton-like method to compute a ground state $(\bar{c}, \bar{\Phi})$ of the MCSCF energy;
2. find an initial continuous path γ_0 satisfying $\gamma_0(0) = (\bar{c}, \bar{\Phi})$ and $\gamma_0(1) = (-\bar{c}, \bar{\Phi})$;
3. deform γ_0 to decrease the highest energy along the path, keeping the end points at $(\bar{c}, \bar{\Phi})$ and $(-\bar{c}, \bar{\Phi})$;
4. when the highest point on the deformed path has a small enough gradient, switch to a Newton-like method to converge to the closest critical point.

We have found many algorithms in the literature for the optimization of paths (often applied to the simulation of chemical reactions on potential energy surfaces) [40,42–48,41,49–51], some of them being quite peculiar in our opinion. The method that we propose below for the deformation of paths, and which seems to give good results on our problem, is of general concern and could therefore also be useful for some other problems.

2.4. Solving the mountain pass problem: a method of deformation of paths

Let us first point out that solving a mountain pass problem is by no means equivalent to finding a saddle point somewhere "between" two minima. The example of the search of the first excited state of the H_2 molecule (Section 3.3.1) is an illustration of this statement. In this example indeed, the optimal path obtained with our algorithm contains two saddle points of different energies; an algorithm of saddle point localization could converge toward the one of lower energy, and thus underestimate the mountain pass energy.

The best way for properly solving a mountain pass problem is in fact to deform paths. A mathematical study of an algorithm of this type can be found in [52,53]. Our method has been inspired by the one described in these references, but it is not identical (see below). In this section, we present it in the following abstract setting: solve the mountain pass problem on the energy surface defined by the functional $\mathscr{E}$ on the Riemann manifold $\mathscr{M}$ between the two points M_0 and M_0' of $\mathscr{M}$, or in other words, find a minimizer of

$$\inf_{\substack{\gamma \in C^0([0;1],\mathscr{M}) \\ \gamma(0)=M_0, \gamma(1)=M_0'}} \max_{t\in[0;1]} \mathscr{E}(\gamma(t)).$$

Like in [42,44,45,50,51], the main idea is to sample a given path linking M_0 and M_0' with a sequence of points $M_0, M_1, \ldots, M_{N+1}$ of $\mathscr{M}$, such that $M_{N+1} = M_0'$. During the optimization process, the number N of points used to represent the current path is not necessarily fixed. In our method, we associate with each sequence $(t_k, M_k)_{0\leqslant k\leqslant N+1}$ where $0 = t_0 < t_1 < \cdots < t_N < t_{N+1} = 1$ are real numbers, a uniquely defined continuous path $\gamma : [0;1] \to \mathscr{M}$ which satisfies $\gamma(t_k) = M_k$. This is done by selecting once and for all, a convenient interpolation scheme. A possible choice is to take for $\gamma(t)$ some piecewise geodesic curve on the manifold $\mathscr{M}$. Simplest interpolation schemes can also be chosen, for in practice, M_k and M_{k+1} will be close together. In some cases, spline-type interpolation functions can also be used.

A sequence $(t_k, M_k)_{0\leqslant k\leqslant N+1}$ being given, one can use the gradient field of the functional $\mathscr{E}$ to deform the associated continuous path. A naive approach consists in simply moving each M_k in the direction opposite to the gradient with a step-length α_k [54]. Remark that since the new point M_k' has to lay on the (curved) manifold $\mathscr{M}$, one has to make precise the statement "in the direction opposite to the gradient". The most intrinsic rule is to move M_k on the geodesic curve which spurts out from M_k in the direction opposite to the gradient [55]. A simpler alternative is first to move M_k in the tangent space, then to project the so-obtained point on the manifold (we shall use this method in our problem).

When this naive procedure is iterated, each point M_k falls down in one of the valleys of the function. In [42,44,45], it is suggested to circumvent this problem by linking the points $(M_k)_{0\leqslant k\leqslant N+1}$ with strings or

elastic bands. In some situation, one can also project the gradient in some direction which is fixed or not [56]. The convergence to the saddle point then depends deeply on the new parameters introduced (strength of the elastic bands, direction of the projection, etc.).

We use a simpler but apparently more efficient solution, similar to ideas of [52,53,57–60]. It consists in first computing the path γ' associated with (t_k, M'_k), and then finding new points (t'_k, M''_k) which are better distributed in some sense on the (uniquely defined) continuous path γ'. We have observed that for stability reasons, the points need to be redistributed after each minimization step. We use in addition the following rule: the larger the difference between the maximum of the energy on γ and $\mathscr{E}(M_k)$, the smallest the step-length α_k in the direction opposite to the gradient. This simple trick helps in preventing the points M_k from falling down in the valleys (see Fig. 1).

We have applied the above method to several test cases (notably to the ones described in [49]) and we have observed a convergence to the saddle point in all the cases, when the number N of points is large enough. We have also checked on these test cases that switching to a Newton-like method once the mountain-pass algorithm has found a state close enough to the saddle point, is an efficient strategy. In the following section, we apply this method to the calculation of the first MCSCF excited state.

3. Computation of the first excited state of two-electrons systems

3.1. The MCSCF approximation for two-electron systems

3.1.1. Spin symmetry: singlet and triplet states

In order to be able to simulate real molecular systems, we now need to reintroduce the spin variables. As the N-body Hamiltonian H_N and the spin operators S^2 and S_z (see, e.g. [61]) commute, it is convenient to search for eigenfunctions of H_N that also are eigenfunctions of S^2 and S_z. For two-electron systems, the situation is particularly simple. There are only two types of wavefunctions which are eigenfunctions of both S^2 and S_z, namely the so-called singlet and triplet states.

A singlet state is a wavefunction of the form

$$\Psi^{\mathrm{s}}(x,\sigma;y,\sigma') = \psi(x,y)|\alpha\beta\rangle(\sigma,\sigma'),$$

where $\psi(x,y)$ *is symmetric* in $L^2(\mathbb{R}^3\times\mathbb{R}^3)$, i.e., such that $\psi(y,x)=\psi(x,y)$. The antisymmetry is carried by the spin function $|\alpha\beta\rangle(\sigma,\sigma')$, which is defined for $(\sigma,\sigma')\in\{|\uparrow\rangle,|\downarrow\rangle\}\times\{|\uparrow\rangle,|\downarrow\rangle\}$ by

$$|\alpha\beta\rangle(\sigma,\sigma') = \frac{1}{\sqrt{2}}(\alpha(\sigma)\beta(\sigma') - \beta(\sigma)\alpha(\sigma')),$$

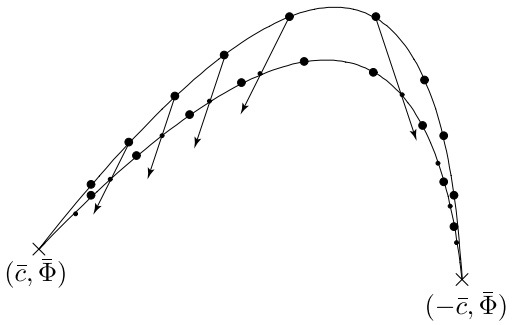


Fig. 1. The deformation method.

where

$$\alpha(|\uparrow\rangle) = 1, \quad \alpha(|\downarrow\rangle) = 0, \quad \beta(|\uparrow\rangle) = 0, \quad \beta(|\downarrow\rangle) = 1.$$

A triplet state takes the form

$$\Psi^{\mathrm{t}}(x,\sigma;y,\sigma') = \psi(x,y)\alpha(\sigma)\alpha(\sigma'),$$

where $\psi(x,y)$ is *antisymmetric* in $L^2(\mathbb{R}^3 \times \mathbb{R}^3)$, i.e., $\psi(y,x) = -\psi(x,y)$ (the spin function $\alpha(\sigma)\alpha(\sigma')$ is symmetric and the antisymmetry is carried by the function of the space variables).

For two-electron systems, the MCSCF wavefunctions thus read

$$\psi = \sum_{1 \leqslant i,j \leqslant K} c_{ij}\varphi_i \otimes \varphi_j, \tag{20}$$

where the $K \times K$ matrix $C = (c_{ij})$ is symmetric for singlet states and antisymmetric for triplet states. The condition $\|\psi\|_{L^2} = 1$ also reads $\|C\| = 1$ where $\|C\| = \mathrm{tr}(CC^{\mathrm{T}})^{1/2}$.

3.1.2. The MCSCF model in finite dimension

For numerical simulations, one most often resorts to a Galerkin approximation: each φ_i is expanded on a finite basis $(\chi_\mu)_{1 \leqslant \mu \leqslant N_b}$ of $H^1(\mathbb{R}^3)$ functions specially designed for electronic structure calculations, the so-called atomic orbitals. This approximation is referred to as the Linear Combination of Atomic Orbitals (LCAO) approximation in the Computational Chemistry literature (see, e.g. [62]). Let S be the matrix defined by

$$S_{\mu\nu} = \int_{\mathbb{R}^3} \chi_\mu \chi_\nu$$

and $\varphi = (\varphi_{\mu i})$ be the $N_b \times K$ coordinate matrix of the functions $(\varphi_i)_{1 \leqslant i \leqslant K}$ in the basis $(\chi_\mu)_{1 \leqslant \mu \leqslant N_b}$. The condition $\int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij}$ also reads $\varphi^{\mathrm{T}} S \varphi = I_K$ (of course, N_b must be chosen greater or equal to K) and the energy of a state Ψ^{s} or Ψ^{t}, as a function of C and φ, has the following expression:

$$\mathscr{E}(C,\varphi) = 2\mathrm{tr}(C^{\mathrm{T}}\varphi^{\mathrm{T}} h \varphi C) + \mathrm{tr}\Big((\varphi C \varphi^{\mathrm{T}})^{\mathrm{T}} \mathscr{W}(\varphi C \varphi^{\mathrm{T}})\Big), \tag{21}$$

where h is the $N_b \times N_b$ matrix defined by

$$h_{\mu\nu} = \frac{1}{2}\int_{\mathbb{R}^3} \nabla\chi_\mu \cdot \nabla\chi_\nu + \int_{\mathbb{R}^3} V \chi_\mu \chi_\nu$$

and where $\mathscr{W}$ is the linear map associated with the tensor W defined by

$$W_{\mu\nu\kappa\lambda} = \int_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{\chi_\mu(x)\chi_\nu(y)\chi_\kappa(x)\chi_\lambda(y)}{|x-y|}\,\mathrm{d}x\,\mathrm{d}y,$$

i.e., for any $N_b \times N_b$ matrix X

$$[\mathscr{W}(X)]_{\mu\nu} = \sum_{\kappa,\lambda=1}^{N_b} W_{\mu\nu\kappa\lambda} X_{\kappa\lambda}. \tag{22}$$

Remark that expression (21) is valid for both singlet and triplet states, but that the matrix C appearing in this formula is symmetric for singlet states and antisymmetric for triplet states.

For the sake of brevity, we now only deal with the singlet state case. The manifold of admissible singlet states is

$$\mathscr{M} = \{(C,\varphi) \in M(K,K) \times M(N_b,K), \quad C^{\mathrm{T}} = C, \quad \mathrm{tr}(C^{\mathrm{T}}C) = 1, \quad \varphi^{\mathrm{T}} S \varphi = I_K\},$$

where $M(K,K')$ denotes the set of $K\times K'$ real matrices. The MCSCF equations (i.e., the stationarity conditions of the MCSCF energy (21) on the manifold $\mathscr{M}$) then take the form:

$$\begin{cases} (\varphi^{\mathrm{T}} h\varphi C + C^{\mathrm{T}}\varphi^{\mathrm{T}} h\varphi) + \varphi^{\mathrm{T}}\mathscr{W}(\varphi C\varphi^{\mathrm{T}})\varphi = \beta C, \\ h\varphi CC^{\mathrm{T}} + \mathscr{W}(\varphi C\varphi^{\mathrm{T}})\varphi C = S\varphi\Lambda, \end{cases} \tag{23}$$

where $\beta\in\mathbb{R}$ and where Λ is a $K\times K$ symmetric matrix.

3.1.3. The reduced model

The problem can be dramatically simplified by using a rotation invariance property [63]. Indeed, using the constraint $\varphi^{\mathrm{T}} S\,\varphi = I_K$, we see that $\mathscr{E}(C,\varphi)$ only depends on $X = \varphi C\varphi^{\mathrm{T}}$

$$\mathscr{E}(C,\varphi) = 2\mathrm{tr}(SX^{\mathrm{T}}hX) + \mathrm{tr}\big(X^{\mathrm{T}}\mathscr{W}(X)\big). \tag{24}$$

Notice now that for any rotation matrix $U\in \mathrm{O}_K(\mathbb{R})$, one has $X = \varphi C\varphi^{\mathrm{T}} = \varphi' C'\varphi'^{\mathrm{T}}$ with $\varphi' = \varphi U$ and $C' = U^{\mathrm{T}}CU$. Since φ' obviously satisfies the constraint $\varphi'^{\mathrm{T}} S\varphi' = I_K$, we see that the energy functional $\mathscr{E}$ is invariant under the action

$$U\cdot(C,\varphi) = (U^{\mathrm{T}}CU, \varphi U) \tag{25}$$

of the orthogonal group $\mathrm{O}_K(\mathbb{R})$. Now, since C is symmetric and real, there exists a $U\in\mathrm{O}_K(\mathbb{R})$ such that $C' = \mathrm{diag}(c_1,\ldots,c_K)$. Up to a rotation of the orbitals, $\varphi' = \varphi U$, this means that *there is no restriction in assuming that the matrix C is diagonal*. When using this reduced model (which is *not* an approximation), the manifold of admissible singlet states reads

$$\mathscr{M}^{\mathrm{red}} = S^{K-1}\times \mathscr{W}_K^{N_b} \tag{26}$$

with

$$\mathscr{W}_K^{N_b} = \{\varphi\in M(N_b,K),\quad \varphi^{\mathrm{T}} S\varphi = I_K\},$$

and the energy functional is given by

$$\mathscr{E}^{\mathrm{red}}(c,\varphi) = \mathscr{E}(C^s(c),\varphi)$$

where $C^s(c) = \mathrm{diag}(c_1,\ldots,c_K)$. Lastly, the MCSCF equations become:

$$\begin{cases} H(\varphi)\cdot c = \beta\cdot c, \\ h\varphi(C^s(c))^2 + \mathscr{W}(\varphi C^s(c)\varphi^{\mathrm{T}})\varphi C^s(c) = S\varphi\Lambda \end{cases} \tag{27}$$

with

$$H(\varphi)_{ij} = 2\varphi_i^{\mathrm{T}} h\varphi_i\delta_{ij} + \varphi_i^{\mathrm{T}}\mathscr{W}(\varphi_j\varphi_j^{\mathrm{T}})\varphi_i.$$

We notice that the use of the reduced model indeed corresponds to choosing a gauge for the invariance under the action of the orthogonal group $\mathrm{O}_K(\mathbb{R})$. We also point out that the discrete permutation group still acts on the reduced manifold $\mathscr{M}^{\mathrm{red}}$, by simply changing the order of the orbitals. A same kind of reduction is done in [63–66] and [5, Appendix]. Triplet states can also be simplified with the same kind of argument since for any antisymmetric matrix C, there exists a rotation matrix U such that $U^{\mathrm{T}}CU = C' = \mathrm{diag}(C_i,\ldots,C_P)$ if $K = 2_p$, and $C' = \mathrm{diag}(C_1,\ldots,C_p,0)$ it $K = 2p + l$, where

$$C_i = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & c_i \\ -c_i & 0 \end{pmatrix}.$$

At this stage, some important comments have to be made. It is crucial to notice that when working with the reduced model, we do not restrict the associated set of two-body wavefunctions, but *we change the geometry*

of the manifold on which the optimization of paths has to be made. For a minimization problem, the restriction to the reduced manifold (26) has no theoretical consequence. But for the new variational method giving the first excited state, Theorem 1 is not valid anymore.

It could be that a path in the non-restricted manifold $\mathscr{M}$ cannot be continuously projected onto the reduced manifold $\mathscr{M}^{\mathrm{red}}$. Saying differently, it could be that it is not possible to choose a gauge continuously along the path. To give a simple example of such a situation when $K = 2$, let be $(c, \varphi) \in \mathscr{M}^{\mathrm{red}}$ such that $c = (1, 0)$ and $\varphi = (\varphi_1, \varphi_2)$ for some orthogonal vectors φ_1 and φ_2 in $\mathscr{W}_K^{N_b}$. We now introduce $(c', \varphi') = U \cdot (c, \varphi)$ with $c' = (0, 1)$ and $\varphi' = (\varphi_2, -\varphi_1)$, U being the permutation matrix

$$U = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

We also introduce $C = C^{\mathrm{s}}(c)$ and $C' = C^{\mathrm{s}}(c')$ the diagonal matrices associated with c and c'. Of course, $(C, \varphi) \in \mathscr{M}$ and $(C', \varphi') \in \mathscr{M}$ represent the same wavefunction $\Psi = \varphi_1 \otimes \varphi_1$ and one can indeed find a continuous path linking the two points of $\mathscr{M}$, and on which the energy is constant: $\mathscr{E}(C(t), \varphi(t)) = \mathscr{E}(C, \varphi) = \mathscr{E}^{\mathrm{red}}(c, \varphi)$. By (25), it suffices to take, for $t \in [0; 1]$,

$$(C(t), \varphi(t)) = U(t) \cdot (C, \varphi) \quad \text{with} \quad U(t) = \begin{pmatrix} \cos(\pi t/2) & -\sin(\pi t/2) \\ \sin(\pi t/2) & \cos(\pi t/2) \end{pmatrix}.$$

However, along this path, $C(t)$ is not necessarily diagonal. Indeed, none of the points on this path (except the end points) belong to $\mathscr{M}^{\mathrm{red}}$. It is even easy to prove that it is *not* possible to link (c,φ) and (c',φ') in the reduced manifold $\mathscr{M}^{\mathrm{red}}$ by a path on which the energy is constant, whereas this is trivial in the original manifold $\mathscr{M}$.

The same problem will appear in any computation using the reduced model: in the reduced manifold $\mathscr{M}^{\mathrm{red}}$, it is impossible to link by a continuous path on which the energy is constant, two MCSCF states differing only by the order of their orbitals. Therefore, when the reduced model is used to solve the mountain pass problem between the two MCSCF ground states $(\bar{c}, \bar{\varphi})$ and $(-\bar{c}, \bar{\varphi})$, it might be impossible to obtain the first excited state as the highest saddle point along the path. This is due to the specific geometry of the reduced manifold $\mathscr{M}^{\mathrm{red}}$. To obtain the first excited state, one might have to *permute the orbitals* of one of the end points of the path. On the other hand, the reduced model allows to significantly decrease the number of variables, a huge advantage from the numerical point of view. In addition, the results obtained within the reduced model are most often easily interpreted. For these reasons, we shall therefore describe our main algorithm within the reduced model framework. The practical difficulties inherent to this choice will be commented in Section 3.3, in which we provide and analyse numerical results for the H_2 molecule.

3.2. Description of the algorithm

We only deal with the reduced model of the singlet state. We shall make use of the following interpolation rule: a discrete path on $\mathscr{M}^{\mathrm{red}} = S^{K-1} \times \mathscr{W}_K^{N_b}$ being given as a finite sequence $(t_k,(c^k, \varphi^k))_{0 \leqslant k \leqslant N+1}$, where:

- $t_0 = 0 < t_1 < \cdots < t_N < t_{N+1} = 1$ are real numbers;
- $(c^k, \varphi^k)_{0 \leqslant k \leqslant N+1}$ are points of $S^{K-1} \times \mathscr{W}_K^{N_b}$ such that $c^k \neq c^{k+1}$ for any $0 \leqslant k \leqslant N$,

we define the associated continuous path $\gamma \in C^0([0, 1], S^{K-1} \times \mathscr{W}_K^{N_b})$ according to

$$\forall t \in [0, 1], \quad \gamma(t) = (c(t), \varphi(t)),$$

where

$$\forall 0 \leqslant k \leqslant N, \quad \forall t \in [t_k, t_{k+1}], \quad c(t) = \cos(\theta_k(t))c^k + \sin(\theta_k(t))\widetilde{c}^{k+1}$$

with

$$\begin{cases} \widetilde{c}^{k+1} = \frac{c^{k+1} - (c^{k+1}, c^k)c^k}{\|c^{k+1} - (c^{k+1}, c^k)c^k\|}, \\ \theta_k(t) = \frac{t - t_k}{t_{k+1} - t_k} \arccos(c^{k+1}, c^k) \end{cases}$$

and

$$\forall 0 \leqslant k \leqslant N, \quad \forall t \in [t_k, t_{k+1}], \quad \varphi(t) = \widetilde{\varphi}(t)[\widetilde{\varphi}(t)^{\mathrm{T}} S \widetilde{\varphi}(t)]^{-1/2}$$

with

$$\widetilde{\varphi}(t) = \varphi^k + \frac{t - t_k}{t_{k+1} - t_k}(\varphi^{k+1} - \varphi^k).$$

We can now describe our algorithm for computing the first excited state of two-electron systems.

Step A: search for a MCSCF ground state $(\bar{c}, \bar{\varphi})$, i.e., solve

$$\inf\{\mathscr{E}^{\mathrm{red}}(c, \varphi), \quad (c, \varphi) \in S^{K-1} \times \mathscr{W}_K^{N_b}\}$$

with the Newton algorithm; a convenient initial guess is the Hartree–Fock ground state, which can itself be obtained by a self-consistent field algorithm [8].

Step B: construction of an initial trial path.

As already mentioned in Section 2.3, we get rid of the outer minimization in (18) and concentrate on solving

$$\lambda_2^{s,r} = \inf_{\substack{\gamma \in C^0([0,1], \mathscr{M}^{\mathrm{red}}) \\ \gamma(0) = (\bar{c}, \bar{\varphi}), \gamma(1) = (-\bar{c}, \bar{\varphi})}} \max_{t \in [0;1]} \mathscr{E}^{\mathrm{red}}(\gamma(t)).$$

Notice that the method presented here can easily be generalized if other end points are chosen for the path (see the comments at the end of Section 3.1.3).

Let $\bar{c}_1$ be the second eigenvector of the Hamiltonian matrix

$$[H(\bar{\varphi})]_{ij} = 2\bar{\varphi}_i^{\mathrm{T}} h \bar{\varphi}_i \delta_{ij} + \bar{\varphi}_i^{\mathrm{T}} \mathscr{W}(\bar{\varphi}_j \bar{\varphi}_j^{\mathrm{T}}) \bar{\varphi}_i \tag{28}$$

(note that $\bar{c}$ is the ground state of $H(\bar{\varphi})$ and that $\bar{c} \cdot \bar{c}_1 = 0$). A possible initial trial path is a path on which the parameter φ is constant, for instance

$$\bar{\gamma}_0(t) = (c(t), \bar{\varphi}) \tag{29}$$

with $c(t) = \cos(t\pi)\bar{c} + \sin(t\pi)\bar{c}_1$, which is indeed the discrete path associated with the sequence

$$(C^0, \varphi^0) = (\bar{c}, \bar{\varphi}), \quad (c^1, \varphi^1) = (\bar{c}_1, \bar{\varphi}), \quad (c^2, \varphi^2) = (-\bar{c}, \bar{\varphi}) \tag{30}$$

and $t_0 = 0$, $t_1 = 1/2$, $t_2 = 1$.

A better initial guess can, however, be obtained by random perturbations of that reference path $\bar{\gamma}_0$. In practice, we randomly choose a collection of N_{sto} states $(\bar{c}'_j, \bar{\varphi}'_j) \in (\mathrm{vect}(\bar{c})^{\perp} \cap S^{K-1}) \times \mathscr{W}_K^{N_b}$ such that for all $j = 1, \ldots, N_{\mathrm{sto}}$

$$\|\bar{c}'_j - \bar{c}_1\| \leqslant \varepsilon \|\bar{c}_1\| \quad \text{and} \quad \|\bar{\varphi}'_j - \bar{\varphi}\| \leqslant \varepsilon \|\bar{\varphi}\|$$

for a small ε, and we consider the N_{sto} continuous paths $\bar{\gamma}'_j(t)$ obtained by taking $(c^1, \varphi^1) = (\bar{c}'_j, \bar{\varphi}'_j)$ in (30). We then select, among the N_{sto} paths $\bar{\gamma}'_j$, the path $\gamma_0(t)$ for which the maximal energy $\max \mathscr{E}^{\mathrm{red}}(\bar{\gamma}'_j([0;1]))$ is

minimum. The above method can obviously be generalized to discrete paths containing more than three points and can also be used to improve the following Step C (path optimization) when necessary.

We then set $m = 0$, $t_k = k/(N+1)$, $\gamma_0^k = (c^{k,0}, \varphi^{k,0}) = \gamma_0(t_k)$ for $1 \leqslant k \leqslant N$ and $E_0^{\min} = \mathscr{E}^{\mathrm{red}}(\bar{c}, \bar{\varphi})$

Step C: path optimization.

For the sake of simplicity, we displace the nodes in the direction opposite to the gradient; for this purpose:

1. we compute for each $1 \leqslant k \leqslant N$, the MCSCF energy at the point $\gamma_m^k = (c^{k,m}, \varphi^{k,m})$
$$E_m^k = \mathscr{E}^{\mathrm{red}}(\gamma_m^k) = (H(\varphi^{k,m})c^{k,m}, c^{k,m})$$
and set
$$E_m^{\max} = \max_{1\leqslant k\leqslant N} E_m^k, \quad i_m^{\max} = \arg\max_{1\leqslant k\leqslant N} E_m^k;$$
2. if $\|E_m^{\max} - E_{m-1}^{\max}\| < \eta$, M times consecutively, we set $(\widetilde{c}, \widetilde{\varphi}) = \gamma_m^{i_m^{\max}}$ and go to Step D (i.e., we switch to a Newton-like algorithm);
3. we project the components $\nabla_c \mathscr{E}^{\mathrm{red}}(\gamma_m^k)$ and $\nabla_\varphi^S \mathscr{E}^{\mathrm{red}}(\gamma_m^k)$ of the gradients of the energy at the points γ_m^k on the tangent spaces of the underlying manifolds:
$$g_m^k = \nabla_c \mathscr{E}^{\mathrm{red}}(c^{k,m}, \varphi^{k,m}) - (\nabla_c \mathscr{E}^{\mathrm{red}}(c^{k,m}, \varphi^{k,m}), c^{k,m})c^{k,m} = 2(H(\varphi^{k,m}) - E_m^k)c^{k,m}$$
$$G_m^k = 4P^S_{\varphi^{k,m}}[S^{-1}(h\varphi^{k,m}C^s(c^{k,m})^2 + \mathscr{W}(\varphi^{k,m}C^s(c^{k,m})(\varphi^{k,m})^{\mathrm{T}})\varphi^{k,m}C^s(c^{k,m}))],$$
where $P^S_{\varphi^{k,m}}$ is the orthogonal projector of $M(N_b,K)$ on the tangent space $T_{\varphi^{k,m}}\mathscr{W}_K^{N_b}$ for the scalar product $\langle\cdot,\cdot\rangle_S$ defined on $M(N_b,K)$ by $\langle A,B\rangle_S = \mathrm{tr}(A^{\mathrm{T}}SB)$, which means
$$\forall \varphi \in \mathscr{W}_K^{N_b}, \quad \forall Z \in M(N_b,K), \quad P_\varphi^S Z = Z - \frac{1}{2}\varphi(\varphi^{\mathrm{T}}SZ + Z^{\mathrm{T}}S\varphi); \tag{31}$$
4. for each $1 \leqslant k \leqslant N$, we search for an optimal step $0 \leqslant \alpha_m^k \leqslant 1$ and set
$$d_m^k = -f_d(\beta_m^k)\alpha_m^k g_m^k, \quad D_m^k = -f_D(\beta_m^k)\alpha_m^k G_m^k$$
with
$$\beta_m^k = \frac{E_m^k - E_0^{\min}}{E_0^{\max} - E_m^{\min}}$$
and where f_d and f_D are well-chosen functions satisfying $f_d(0) = f_D(0) = 0$ and $f_d(1) = f_D(1) = 1$.
5. we displace the nodes γ_m^k along the descent directions:
$$\bar{c}^{k,m+1} = \frac{1}{\|c^{k,m} + d_m^k\|}(c^{k,m} + d_m^k),$$
$$\bar{\varphi}^{k,m+1} = \widetilde{\varphi}[\widetilde{\varphi}^{\mathrm{T}}S\widetilde{\varphi}]^{-1/2} \quad \text{with} \quad \widetilde{\varphi} = \varphi^{k,m} + D_m^k$$
and introduce the continuous path $\gamma_{m+1}(t)$ associated with the sequence $(k/(N+1), (\bar{c}^{k,m+1}, \bar{\varphi}^{k,m+1}))$;
6. we reparametrize the new path γ_{m+1}. For this purpose, we define the length of the discrete path by
$$L = \sum_{k=1}^{N+1} \|\bar{c}^{k,m+1} - \bar{c}^{k-1,m+1}\| + \|\bar{\varphi}^{k,m+1} - \bar{\varphi}^{k-1,m+1}\|$$
and search for an increasing sequence $0 = t_0 < t_1 < \cdots < t_{N'} < 1$ satisfying
$$\forall k = 0, \ldots, N'-1, \quad \left\|\gamma_{m+1}(t_{k+1}) - \gamma_{m+1}(t_k)\right\| \in [L/N, L/N + \varepsilon'],$$

where ε' is a small enough threshold, and with N' as large as possible. We now set $t_{N'+1} = 1$ and define the reparametrized path $\gamma^{\mathrm{rep}}_{m+1}(t)$ associated with the sequence $(t_k, \gamma_{m+1}(t_k))_{k=1\ldots N'+1}$. Finally, since $N' \neq N$ a priori, we set

$$\gamma^k_{m+1} = (c^{k,m+1}, \varphi^{k,m+1}) = \gamma^{\mathrm{rep}}_{m+1}(k/(N+1));$$

7. we set $m = m + 1$ and return to Step C.1.

Step D: use the Newton algorithm, with $(\widetilde{c}, \widetilde{\varphi})$ as initial guess, to solve (27).

3.3. Numerical results: the H_2 molecule

In this section, we present some numerical results concerning the singlet state of the H_2 molecule. We assume that the two protons of the H_2 molecule are located along the x-axis with a distance r, at $(-r/2, 0, 0)$ and $(r/2, 0, 0)$.

In the usual Quantum Chemistry programs, the computation is always restricted to a particular symmetry space. Indeed, the two-body Hamiltonian H commutes with the symmetry operator T defined as $(Tf)(x, y) = f(-x, -y)$, for any symmetric $f \in L^2_s(\mathbb{R}^3 \times \mathbb{R}^3)$. Therefore, H stabilizes the two eigenspaces of T which are:

$$\Sigma_{\mathrm{g}} := \{f \in L^2_s(\mathbb{R}^3 \times \mathbb{R}^3) | f(x, y) = f(-x, -y)\}, \tag{32}$$

$$\Sigma_{\mathrm{u}} := \{f \in L^2_s(\mathbb{R}^3 \times \mathbb{R}^3) | f(x, y) = -f(-x, -y)\}, \tag{33}$$

and one can search for eigenfunctions in these spaces. In the results presented below, we have used the algorithm presented in Section 3.2 without imposing symmetry restrictions. It turns out that the MCSCF ground state does have the correct symmetry Σ_{g}, but that the first excited state is only close to the symmetry Σ_{u}: in general, due to the non-linearity of the MCSCF method, one cannot be sure that the computed states will be in the same linear spaces as the true eigenfunctions of the N-body Hamiltonian. Of course, the algorithm of Section 3.2 can be very easily adapted if one wants to restrict the computation of the first excited state to a particular symmetry.

The results presented in the following sections have been obtained with a Scilab [67] program, interfaced with a few C routines aiming in particular at speeding up the tensor–matrix products (22). Let us mention that the overlap matrices S, the core Hamiltonians h, and the bielectronic integral tensors W have been extracted from Gaussian 98 calculations [68].

We have applied the algorithm described in the previous section to the H_2 molecule with the reduced model presented in Section 3.1.3, for various interatomic distances r. All the energies (in Hartree) reported in this section include the additional Coulomb repulsion of the two nuclei. We have used for these calculations the double zeta Dunning's correlation consistent atomic basis set (cc-pVDZ), for which $N_b = 10$. It is composed of two s-type functions and one p-type function (for each space coordinate), which are then centered at the two atoms. All the computations have been made with $K = 4$ (i.e., four natural orbitals and four c_i's coefficients).

The number of iterations in Step C (path optimization) necessary to reach a given convergence criterion strongly depends on the choice of the initial guess. In that respect, the randomly perturbed initial paths constructed in Step B of our algorithm are of better quality than the one given by formula (29).

3.3.1. Analysis of the results for r = 1 Å

Let us first analyze the results for a fixed interatomic distance $r = 1$ Å. The energy profiles of the successive paths generated by the path optimization procedure (Step C) have been reported on the same graph (Fig. 2). One can see that the energy profile of the initial trial path is a single hump and that those of

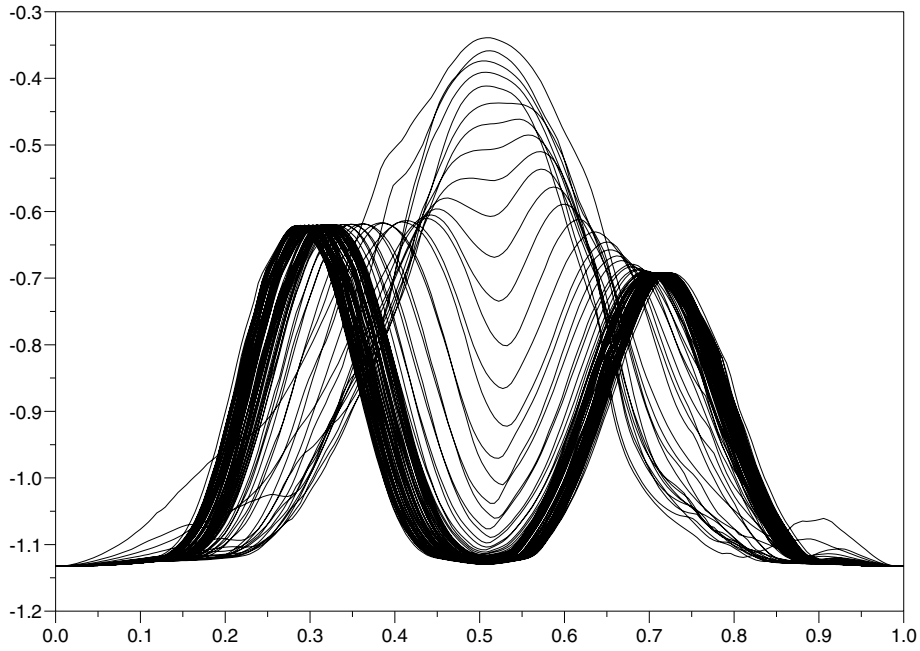


Fig. 2. Energy profiles of the successive paths generated by the path optimization procedure (H_2 molecule, interatomic distance equal to 1 Å).

the iterates progressively turn into a double hump shape. Recall that the energy profiles of the earlier iterates have a rough shape for the initial trial path results from a stochastic local deformation of a reference path (Step B of the algorithm). The optimization process rapidly smoothes the trial path. Notice that due to the reparametrization procedure, the graph of $\mathscr{E}^{\mathrm{red}}(\gamma_m)$ is not necessarily below the graph of $\mathscr{E}^{\mathrm{red}}(\gamma_{m-1})$.

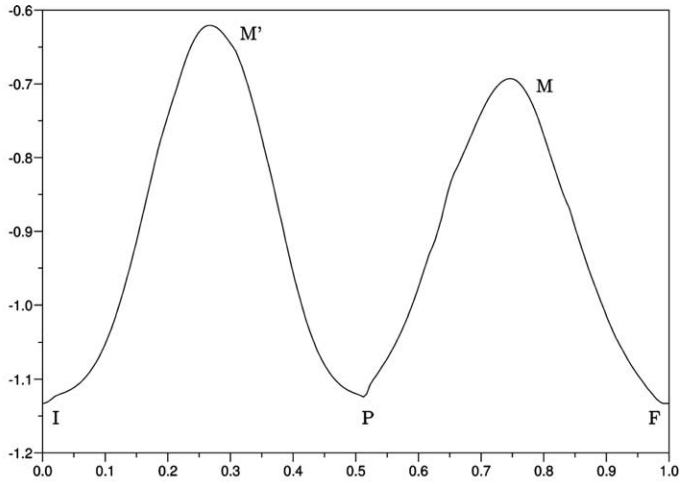
The optimal path γ obtained with our algorithm exhibits a double hump energy profile (see Fig. 3) with two local maxima. Let us point out that we have run on this case many tests with different stochastic initial trial paths; we have always obtained a double hump profile at convergence. Our method thus provides two saddle points of Morse index equal to one (denoted by M and M' in Fig. 3). Normally, the first excited state should be the one of higher energy M' but, since we have used the reduced model, we have to be careful with the interpretation of the so-obtained MCSCF states, as explained in Section 3.1.3. It might be possible to avoid the first hump by orbital rotations: the so-obtained path then is on the manifold $\mathscr{M}$ (of admissible singlet states) but not on the reduced manifold $\mathscr{M}^{\mathrm{red}}$.

This is indeed the case. We have displayed in Table 1 the results obtained after use of the Newton algorithm, using as initial states the points I (ground state), M and M' (saddle points), and P (local minimum). Recall that the coordinates in the cc-pVDZ basis of the four natural orbitals are the columns of the matrix φ. Looking first at the point I, one easily deduces from the form of φ_I and the definition of the LCAO basis that the associated wavefunction is even in $L^2(\mathbb{R}^6)$, i.e., it belongs to the Σ_g symmetry space introduced in (32). It can indeed be written as

$$\Psi_I = 0.9860929\varphi_g^1 \otimes \varphi_g^1 - 0.154182\varphi_u^1 \otimes \varphi_u^1 - 0.0548179\varphi_g^2 \otimes \varphi_g^2 - 0.0122131\varphi_u^2 \otimes \varphi_u^2, \tag{34}$$

where the orbitals φ_g^1, φ_g^2 are even and the orbitals φ_u^1, φ_u^2 are odd. This form, which is usually used as an ansatz for ground state calculations has been automatically obtained by the algorithm.

The local minimum P is also displayed in Table 1. It also belongs to Σ_g and, more importantly, the first orbital of I is found at the second place with a reversed sign in P. One might thus expect that a problem due



Point:	$I = F$	M'	P	M
Energy:	−1.132868250	−0.621666969	−1.114084700	−0.693788453

Fig. 3. Path at convergence and energies (in Hartree) of the different states after convergence of the Newton algorithm (H_2 molecule with $r = 1$ Å).

to the specific structure of the reduced manifold $\mathscr{M}^{\text{red}}$ (as explained in Section 3.1.3) is encountered, and that the first hump of the optimal path is indeed an artefact. To demonstrate that this is actually the case, we have applied a permutation of the orbitals of I by using the following permutation matrix:

$$U := \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

We have then run our main algorithm of Section 3.2, but taking as end points of the paths $I' = U \cdot (\bar{c}, \bar{\varphi})$ and $F = (-\bar{c}, \bar{\varphi})$, i.e., we have changed the order of the orbitals of the point I. The paths generated by the algorithm have been reported in Fig. 4. Only one hump is obtained at convergence, the state of higher energy being exactly the point M of the previous calculation. Since I and I' can be linked by a path on which the energy is constant in the full manifold $\mathscr{M}$, the point M is the first MCSCF excited state.[3]

The expressions of c_M and φ_M for the first excited state M are given in Table 1. It is known in Chemistry that the true first excited state of the two-body Hamiltonian belongs to the Σ_u symmetry space. We notice from the form of the coefficients of φ_M that the wavefunction Ψ_M of our approximate first excited state is only *very close to be a state of the Σ_u space*. This phenomenon is due to the non-linearity of the MCSCF model. Usually, in Quantum Chemistry one computes the excited states by restricting the whole calculation to a particular symmetry. In the quantum chemistry packages Molpro [69] or Dalton [38] for instance, the first excited state is obtained by minimizing the energy in the Σ_u symmetry space. Computing the norm of

[3] As usual in non-convex settings, one cannot be definitely sure that the calculation has converged toward the global solution of the minimax problem; nevertheless, as we performed a lot of calculations, with different initial conditions, and always got the same result, one can think that we have reached the global minmax.

Table 1
Results for the H_2 molecule with $r = 1\text{Å}$

$$c_I = \begin{bmatrix} 0.9860929 \\ -0.1564182 \\ -0.0548179 \\ -0.0122131 \end{bmatrix} \quad \varphi_I = \begin{bmatrix} -0.3870975 & -0.7828663 & 0.5785016 & 0.6474821 \\ -0.2203889 & -0.3144096 & -0.6161753 & -0.4532409 \\ -0.0161929 & -0.0022539 & 0.2808267 & -0.5622198 \\ 1.828\text{E}-10 & 5.252\text{E}-09 & -2.062\text{E}-07 & 8.258\text{E}-07 \\ -7.916\text{E}-11 & -7.147\text{E}-09 & 2.777\text{E}-07 & -0.0000012 \\ -0.3870975 & 0.7828663 & 0.5785015 & 0.6474806 \\ -0.2203889 & 0.3144096 & -0.6161752 & -0.4532386 \\ 0.0161929 & -0.0022539 & -0.2808267 & 0.5622208 \\ 2.823\text{E}-10 & -4.453\text{E}-09 & -2.043\text{E}-07 & 7.659\text{E}-07 \\ -5.130\text{E}-10 & -4.792\text{E}-10 & 3.214\text{E}-07 & 6.723\text{E}-08 \end{bmatrix}$$

$$c_{M'} = \begin{bmatrix} 0.7038748 \\ 0.7038748 \\ -0.894989 \\ -0.333252 \end{bmatrix} \quad \varphi_{M'} = \begin{bmatrix} -0.4698123 & 0.0278596 & 0.6542788 & 0.5931451 \\ -1.0258021 & -0.6809551 & -0.6653752 & -0.3626093 \\ -0.323337 & 0.0352589 & 0.2109307 & -0.5902892 \\ 5.891\text{E}-11 & -3.296\text{E}-11 & 6.744\text{E}-08 & -4.234\text{E}-09 \\ -6.193\text{E}-11 & -2.129\text{E}-11 & -1.047\text{E}-07 & 7.772\text{E}-09 \\ -0.0278596 & 0.4698123 & 0.6542789 & 0.5931451 \\ 0.6809551 & 1.0258021 & -0.6653753 & -0.3626093 \\ 0.0352589 & -0.0323337 & -0.2109307 & 0.5902892 \\ -2.652\text{E}-11 & 5.740\text{E}-11 & -6.901\text{E}-08 & 4.458\text{E}-09 \\ 9.971\text{E}-11 & -9.123\text{E}-11 & 7.500\text{E}-08 & -3.661\text{E}-09 \end{bmatrix}$$

$$c_M = \begin{bmatrix} -0.7086355 \\ 0.7051798 \\ 0.0166917 \\ 0.0166917 \end{bmatrix} \quad \varphi_M = \begin{bmatrix} -0.2226240 & -0.4225444 & 0.1688094 & 0.1688100 \\ 0.9214741 & -1.1049996 & 0.4331117 & 0.4331118 \\ -0.0594902 & -0.0468558 & 0.2761505 & 0.2761512 \\ -2.155\text{E}-09 & -2.425\text{E}-09 & 0.0700146 & -0.0700146 \\ -1.250\text{E}-09 & 4.569\text{E}-09 & 0.0250928 & -0.0250926 \\ -0.4292603 & -0.2232948 & -0.5111576 & -0.5111600 \\ -1.098875 & 0.9146022 & 0.1362389 & 0.1362409 \\ 0.0476153 & 0.561209 & 0.7893125 & 0.7893134 \\ 3.765\text{E}-09 & 4.683\text{E}-09 & -0.6813492 & 0.6812500 \\ 5.273\text{E}-08 & 4.317\text{E}-08 & -0.2441938 & 0.2441876 \end{bmatrix}$$

$$c_P = \begin{bmatrix} -0.0662199 \\ 0.9955221 \\ -0.0662199 \\ -0.0128695 \end{bmatrix} \quad \varphi_P = \begin{bmatrix} -0.0494342 & 0.3820365 & 1.1451155 & 0.0000014 \\ -1.4910565 & 0.2243957 & -1.2001197 & -9.117\text{E}-07 \\ 0.0531034 & 0.0188249 & 0.0971197 & -6.510\text{E}-07 \\ 5.600\text{E}-09 & 8.392\text{E}-10- & 1.599\text{E}-07 & -0.7648589 \\ 1.439\text{E}-07 & -6.942\text{E}-10 & -3.630\text{E}-07 & 0.3206257 \\ 0.0494342 & 0.3820365 & -1.1451155 & 0.0000013 \\ 1.4910565 & 0.2243957 & 1.2001197 & -0.0000012 \\ 0.0531034 & -0.0188249 & 0.0971197 & 9.437\text{E}-07 \\ -9.718\text{E}-09 & -2.769\text{E}-09 & 1.600\text{E}-07 & 0.7648589 \\ -7.654\text{E}-08 & -1.826\text{E}-09 & 3.383\text{E}-07 & -0.3206257 \end{bmatrix}$$

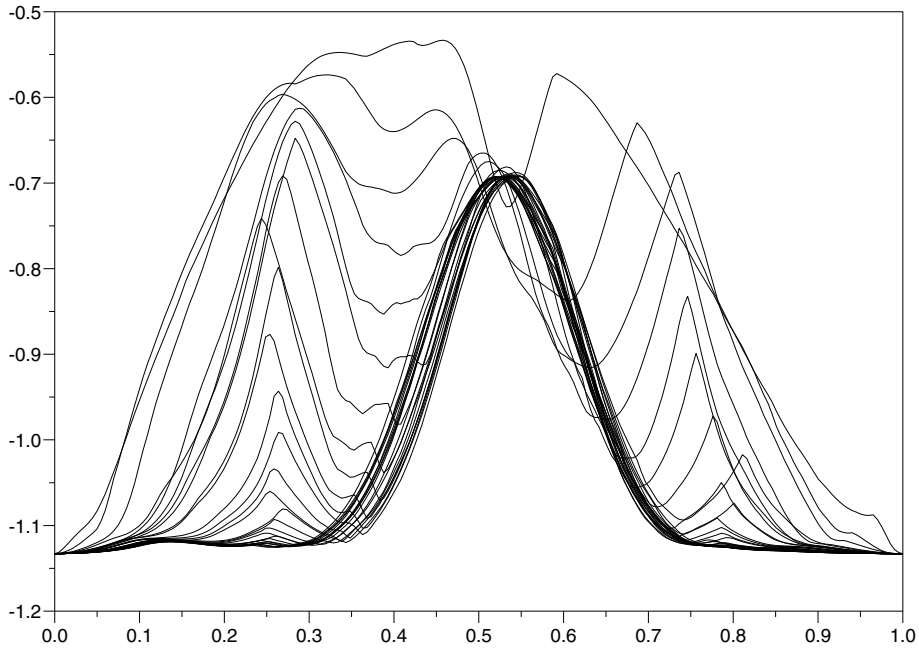


Fig. 4. Energy profiles during the path optimization procedure, when the ends points are $I' = U \cdot (\bar{c}, \bar{\phi})$ and $F = (\bar{c}, \bar{\phi})$ (H_2 molecule, $r = 1$ Å).

the gradient of the states provided by these two programs, we have checked that they are *not* critical points of the MCSCF energy. More precisely, the norm of their gradient is only of order 10^{-4}, whereas our states all have a gradient which has a norm of the order of 10^{-8}. Therefore, adding a symmetry requirement on the wavefunction may not be compatible with a non-linear model such as MCSCF, for excited states.

Furthermore, we note that the energy of M corresponds to the *first eigenvalue of the matrix* $H(\varphi)$ appearing in Eq. (27), as predicted in [37]. This shows that the definition (12) is not relevant: if one does not impose any restriction on the symmetry of the state, the first excited state of the H_2 molecule cannot be obtained as the state which minimizes the second eigenvalue of the Hamiltonian matrix with respect to the orbitals variations (definition (12)). Of course, computing its *full* Hessian matrix, we see that the point M has a Morse index equal to 1. In fact, by Theorem 1, it satisfies the three conditions (a)–(b)–(c) of Section 2. In particular, its energy is an upper bound to the true second eigenvalue of the two-body Hamiltonian.

Notice that c_M possesses two dominant coefficients. This shows the usefulness of the MCSCF method for the calculation of excited states: the Hartree–Fock method is not able to correctly describe such a two-configuration state (recall that the square of the coefficients of c are the weights of the different configurations of the multiconfiguration wavefunction).

Looking now at the state M' in Table 1, we see that it belongs to the Σ_g symmetry space. But since it has a Morse index equal to one, it is probably not a good approximation of the second true excited state (which is known to belong to the Σ_g space also). Indeed, its energy is -0.621666969 whereas the true second excited state has an energy which equals approximately -0.363201395 as computed by Molpro [69] and Dalton [38]. M' is therefore a spurious state which has no physical interpretation. We notice that M' has an energy which is the second eigenvalue of the associated Hamiltonian matrix $H(\varphi)$ and which is a local minimum with respect to orbitals variations. We therefore strongly believe that the point M' is indeed a solution of the eigenvalue minimization problem (12) which reads in this context $\inf\{\lambda_2(\varphi),\ \varphi \in \mathscr{W}_k^{N_b}\}$, $\lambda_2(\varphi)$ being by definition

the second eigenvalue of $H(\varphi)$. This would prove that using (12) with no symmetry restriction indeed leads to an unphysical result, as already predicted in [37], but we do not have a mathematical proof of this claim. Let us emphasize that this problem is not due to a degeneracy of the eigenvalue of the matrix $H(\varphi)$ as described in Section 2.1: even when the optimized eigenvalue of $H(\varphi)$ is not degenerated, (12) may lead to a wrong result.

3.3.2. The potential energy surface of the H_2 molecule

We have applied our algorithm to the reduced model of the H_2 molecule for different values of r between 0.5 and 4 Å. We have always obtained optimal paths with two humps. The first excited state always corresponds to the smallest hump, the highest one being a non-physical solution of the MCSCF equations. In Fig. 5, we have displayed the Hartree–Fock and MCSCF ground state potential energy surface (PES), and the first MCSCF singlet excited state PES (together with the PES of the spurious state M') obtained by our method.

When $r \geqslant 1.5$ Å, the optimal path exhibits the same characteristics as for the case $r = 1$ Å reported above, but the optimal path is more difficult to obtain than for smaller values of r. We have actually observed that in this range of values of r, the choice of the convergence criteria plays a crucial role in the quality of the results. Indeed, the difference $\|E_m^{\max} - E_{m-1}^{\max}\|$ can be very small during many consecutive iterations, just as if convergence was reached. But if we run many additional iterations, the algorithm finally escapes this trap and converges toward the (supposed) optimal path. We have observed that such a sequence of small changes in $E_m^{\max}$ occurs when the energy profile of the trial path turns from a single hump shape into a double hump shape (see Fig. 6).

It is very likely that during the optimization process, the state of higher energy along the trial path passes in the vicinity of an MCSCF critical point with a Morse index two. Since we do not use any second order information in the method of Section 3.2, it is difficult in this case to find the appropriate

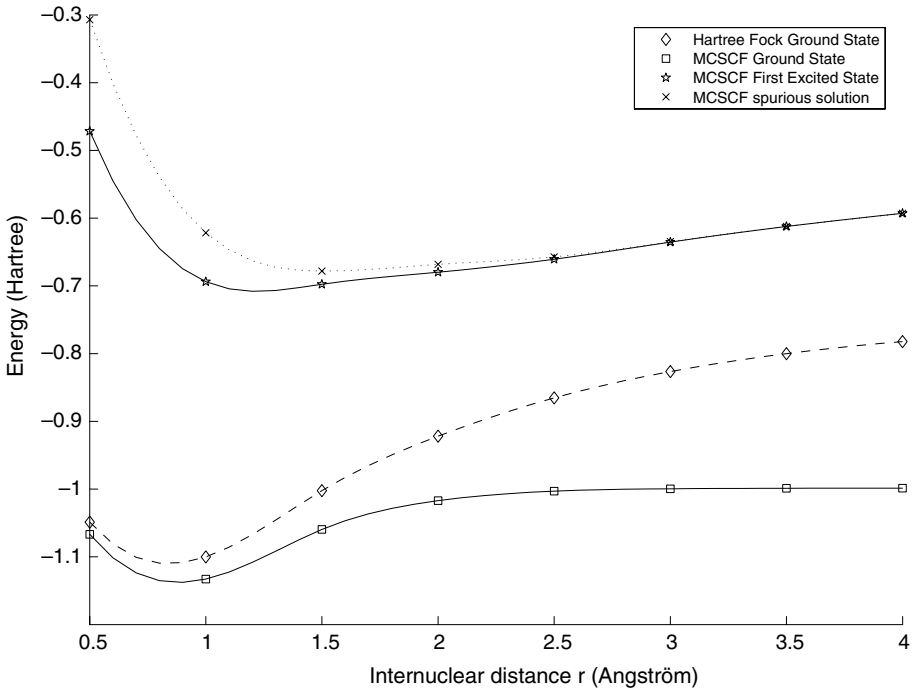


Fig. 5. Potential energy surfaces (PES) of the H_2 molecule.

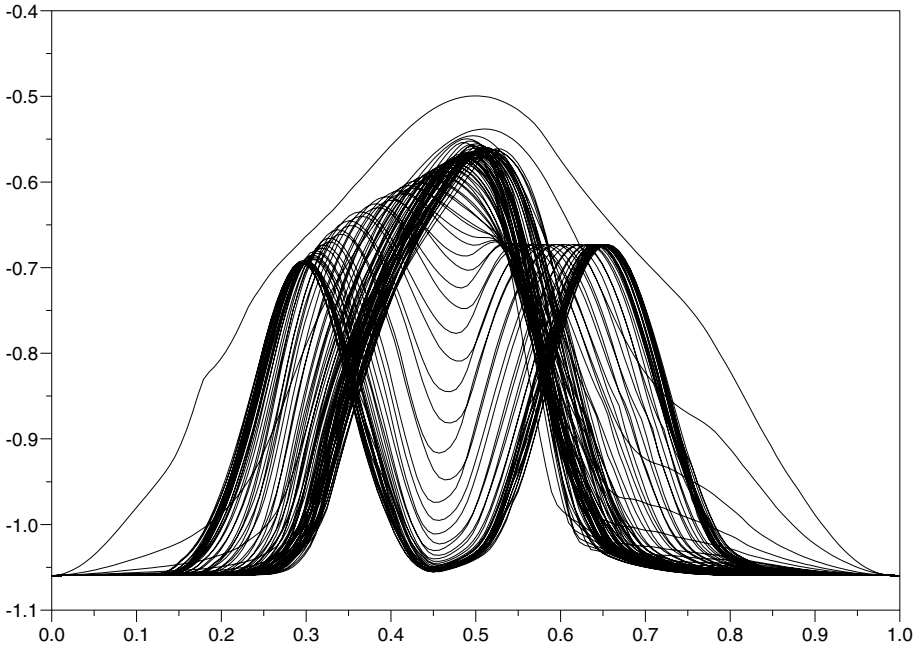


Fig. 6. Iterates of γ_k during the optimization process until convergence, for $r = 1.5$ Å.

descent direction for the deformation of the trial path, explaining the behaviour described in Fig. 6. Many solutions can be proposed to avoid this drawback. One of them consists in using a continuation method: one injects the optimal path obtained for $r = r_0$ as initial guess in the calculation for $r = r_0 + \mathrm{d}r$. We have used this method to compute the first excited state PES of the H_2 molecule for values of r in the range [1.5; 4].

We have also computed the first excited state of Helium-like atoms (one nucleus of charge $Z \geqslant 2$ and two electrons). The same behaviour is observed: we obtain double hump paths, the first excited state being the maximum of the smallest hump. The highest hump provides a spurious MCSCF state having no physical interpretation and which we believe to be a solution of (12). This is consistent with the results obtained for the H_2 molecule with small interatomic distances.

Acknowledgments

We thank Andreas Savin, Hans Jørgen Aagaard Jensen, Claude Le Bris and Éric Séré for useful discussions and advice.

References

[1] M. Head-Gordon, Quantum chemistry and molecular processes, J. Phys. Chem. 100 (1996) 13213–13225.
[2] P. Löwdin, Quantum theory of many-particle systems. III. Extension of the Hartree–Fock scheme to include degenerate systems and correlation effects, Phys. Rev. 97 (6) (1955) 1509–1520.
[3] P. Löwdin, Correlation problem in many-electron quantum mechanics. I. Review of different approaches and discussion of some current ideas, Adv. Chem. Phys. 2 (1959) 207–322.

[4] R. Shepard, The multiconfiguration self-consistent field method. Ab initio methods in quantum chemistry – II, Adv. Chem. Phys. 69 (1987) 63–200.
[5] M. Lewin, Solutions of the multiconfiguration equations in quantum chemistry, Arch. Rat. Mech. Anal. 171 (1) (2004) 83–114.
[6] G. Friesecke, The multiconfiguration equations for atoms and molecules: charge quantization and existence of solutions, Arch. Rat. Mech. Anal. 169 (2003) 35–71.
[7] P. Atkins, R. Friedman, Molecular Quantum Mechanics, third ed., Oxford University Press, Oxford, 1997.
[8] E. Cancès, M. Defranceschi, W. Kutzelnigg, C. Le Bris, Y. Maday, Computational quantum chemistry: a primer, Handbook of Numerical Analysis, vol. X, Elsevier, Amsterdam, 2003, pp. 3–270.
[9] M. Born, R. Oppenheimer, Quantum theory of molecules, Ann. Phys. 84 (1927) 457–484.
[10] G.M. Zhislin, Discussion of the spectrum of Schrodinger operators for systems of many particles, Trudy Moskovskogo matematiceskogo obscestva 9 (1960) 81–120 (in Russian).
[11] P. Löwdin, Quantum theory of many-particle systems. II. Study of the ordinary Hartree–Fock approximation, Phys. Rev. 97 (6) (1955) 1490–1508.
[12] W. Kohn, L. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev. 140 (1965) A1133–A1138.
[13] R. Dreizler, E. Gross, Density Functional Theory, Springer-Verlag, Berlin, 1990.
[14] P. Löwdin, Quantum theory of many-particle systems. I. Physical interpretations by mean of density matrices, natural spin-orbitals, and convergence problems in the method of Configurational Interaction, Phys. Rev. 97 (6) (1955) 1474–1489.
[15] E. Lieb, B. Simon, The Hartree–Fock theory for Coulomb systems, Commun. Math. Phys. 53 (1977) 185–194.
[16] P.-L. Lions, Solutions of Hartree–Fock equations for Coulomb systems, Commun. Math. Phys. 109 (1987) 33–87.
[17] C. Le Bris, A general approach for multiconfiguration methods in quantum molecular chemistry, Ann. Inst. H. Poincaré Anal. Non linéaire 11 (6) (1994) 441–484.
[18] H.-J. Werner, Matrix-formulated direct multiconfiguration self-consistent field and multiconfiguration reference Configuration–Interaction methods. Ab initio methods in quantum chemistry – II, Adv. Chem. Phys. 69 (1987) 1–62.
[19] H.-J. Werner, W. Meyer, A quadratically convergent multiconfiguration-self-consistent field method with simultaneous optimization of orbitals and CI coefficients, J. Chem. Phys. 73 (5) (1980) 2342–2356.
[20] H.-J. Werner, P. Knowles, A second order multiconfiguration SCF procedure with optimum convergence, J. Chem. Phys. 82 (11) (1985) 5053–5063.
[21] R. Eade, M. Robb, Direct minimization in MCSCF theory. The quasi-Newton method, Chem. Phys. Lett. 83 (2) (1981) 362–368.
[22] M. Frisch, I. Ragazos, M. Robb, H. Schlegel, An evaluation of three direct MCSCF procedures, Chem. Phys. Lett. 189 (6) (1992) 524–528.
[23] P. Jørgensen, J. Olsen, D. Yeager, Generalizations of Newton–Raphson and multiplicity independent Newton–Raphson approaches in multiconfigurational Hartree–Fock theory, J. Chem. Phys. 75 (12) (1981) 5802–5815.
[24] D. Yeager, D. Lynch, J. Nichols, P. Jørgensen, J. Olsen, Newton–Raphson approaches and generalizations in multiconfigurational self-consistent field calculations, J. Phys. Chem. 86 (1982) 2140–2153.
[25] P. Jørgensen, P. Swanstrøm, D. Yeager, Guaranteed convergence in ground state multiconfigurational self-consistent field calculations, J. Chem. Phys. 78 (1) (1983) 347–356.
[26] H. Jensen, Electron correlation in molecules using direct second order MCSCF. Relativistic and Electron Correlation Effects in Molecules and Solids, Plenum Press, New York, 1994, pp. 179–206.
[27] B.O. Roos, The complete active space self-consistent field method and its applications in electronic structure calculation. Ab initio methods in quantum chemistry – II, Adv. Chem. Phys. 69 (1987) 399–446.
[28] J. Golab, D. Yeager, P. Jørgensen, Multiple stationary point representation in MCSCF calculations, Chem. Phys. 93 (1985) 83–100.
[29] E. Hylleraas, B. Undheim, Numerische Berechnung der 2 *s*-Terme von Orthound Par-Helium, Z. Phys. 65 (1930) 759–772.
[30] J. MacDonald, Successive approximations by the Rayleigh–Ritz variation method, Phys. Rev. 43 (1933) 830–833.
[31] J. Olsen, P. Jørgensen, D. Yeager, Multiconfigurational Hartree–Fock studies of avoided curve crossing using the Newton–Raphson technique, J. Chem. Phys. 76 (1) (1982) 527–542.
[32] L. Cheung, S. Elbert, K. Ruedenberg, MCSCF optimization through combined use of natural orbital and the Brillouin–Levy–Berthier theorem, Int. J. Quantum Chem. 16 (1979) 1069–1101.
[33] A. Lewis, M. Overton, Eigenvalue optimization, Acta Numer. 5 (1996) 149–190.
[34] H.-J. Werner, W. Meyer, A quadratically convergent MCSCF method for the simultaneous optimization of several states, J. Chem. Phys. 74 (10) (1981) 5794–5801.
[35] M. Reed, B. Simon, Methods of modern mathematical physics. Analysis of Operators, vol. IV, Academic Press, New York, 1978.
[36] K. Docken, J. Hinze, LiH potential curves and wavefunctions, J. Chem. Phys. 57 (11) (1972) 4928–4936.
[37] M. McCourt, J. McIver Jr., On the SCF calculation of excited states: singlet states in the two-electron problem, J. Comput. Chem. 8 (4) (1987) 454–458.
[38] DALTON, a molecular electronic structure program, See http://www.kjemi.uio.no/software/dalton/dalton.html.

[39] H. Jensen, P. Jørgensen, A direct approach to second-order MCSCF calculations using a norm extended optimization scheme, J. Chem. Phys. 80 (3) (1984) 1204–1214.
[40] H. Schlegel, Optimization of equilibrium geometries and transition structures, Adv. Chem. Phys. 67 (1987) 249–286.
[41] W. Quapp, D. Heidrich, Analysis of the concept of minimum energy path on the potential energy surface of chemically reaction systems, Theoret. Chim. Acta 66 (1984) 245–260.
[42] G. Henkelman, G. Jóhannesson, H. Jónsson, in: S.D. Schwartz (Ed.), Methods for Finding Saddle Points and Minimum Energy Paths, Kluwer Academic Publishers, Dordrecht, 2000.
[43] G. Henkelman, H. Jónsson, A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives, J. Chem. Phys. 111 (15) (1999) 7010–7022.
[44] G. Henkelman, H. Jónsson, B. Uberuaga, A climbing image nudged elastic band method for finding saddle points and minimum energy paths, J. Chem. Phys. 113 (22) (2000) 9901–9904.
[45] G. Henkelman, H. Jónsson, Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points, J. Chem. Phys. 113 (22) (2000) 9978–9985.
[46] S.-L. Chiu, J. McDouall, I. Hiller, Prediction of whole reaction paths for large molecular systems, J. Chem. Soc., Faraday Trans. 90 (12) (1994) 1575–1579.
[47] R. Elber, M. Karplus, A method for determining reaction paths in large molecules: application to myoglobin, Chem. Phys. Lett. 139 (5) (1987) 375–380.
[48] P. Jasien, R. Shepard, A general polyatomic potential energy surface fitting method, Int. J. Quantum Chem. 22 (1988) 183–198.
[49] W. Quapp, H. Hirsch, O. Imig, D. Heidrich, Searching for saddle points of potential energy surfaces by following a reduced gradient, J. Comput. Chem. 19 (9) (1998) 1087–1100.
[50] R. Czerminski, R. Elber, Self-avoiding walk between two fixed points as a tool to calculate reaction paths in large molecular systems, Int. J. Quantum Chem. 24 (1990) 167–185.
[51] W. E, W. Ren, E. Vanden-Eijnden, String method for the study of rare events, Phys. Rev. B 66 (2002) 052301.
[52] Y. Choi, P. McKenna, A mountain pass method for the numerical solution of semilinear elliptic problems, Nonlinear Anal. 20 (4) (1993) 417–437.
[53] Y. Choi, P. McKenna, M. Romano, A mountain pass method for the numerical solution of semilinear wave equations, Numer. Math. 64 (4) (1993) 487–509.
[54] D. Liotard, J. Penot, Study of Critical Phenomena, Springer, Berlin, 1981, p. 213.
[55] A. Edelman, T. Arias, S. Smith, The geometry of algorithms with orthogonality constraints, SIAM J. Matrix Anal. Appl. 20 (2) (1998) 303–353.
[56] W. Quapp, Reaction pathways and projection operators: application to string methods, J. Comput. Chem. 25 (10) (2004) 1277–1285.
[57] L. Stachó, M. Bán, Theor. Chim. Acta 83 (1992) 433–440.
[58] L. Stachó, M. Bán, Theor. Chim. Acta 84 (1993) 535–543.
[59] L. Stachó, M. Bán, Procedure for determining dynamically defined reaction path, Comp. Chem. 17 (1) (1993) 21–25.
[60] M. Bán, G. Dömötör, L. Stachó, Dynamically defined reaction path (DDRP) method, J. Mol. Struct.: THEOCHEM 311 (1994) 29.
[61] L. Landau, E. Lifchitz, Quantum Mechanics, Pergamon Press, Oxford, 1977.
[62] W. Hehre, L. Radom, P. Schleyer, J. Pople, Ab initio Molecular Orbital Theory, Wiley, New York, 1986.
[63] P.-O. Löwdin, H. Shull, Natural orbitals in the quantum theory of two-electron systems, Phys. Rev. 101 (6) (1956) 1730–1739.
[64] T. Ando, Properties of fermions density matrices, Rev. Mod. Phys. 35 (3) (1963) 690–702.
[65] A. Coleman, Structure of fermions density matrices, Rev. Mod. Phys. 35 (3) (1963) 668–689.
[66] A. Coleman, V. Yukalov, Reduced Density Matrices: Coulson's Challenge, Springer-Verlag, Berlin, 2000.
[67] C. Gomez, C. Bunks, J. Chancelier, F. Delebecque, M. Goursat, R. Nikoukhah, S. Steer, Engineering and Scientific Computing with Scilab, Birkauser, 1999.
[68] M. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, V. Zakrzewski, J. Montgomery, R. Stratmann, J. Burant, S. Dapprich, J. Millam, A. Daniels, K. Kudin, M. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. Petersson, P. Ayala, Q. Cui, K. Morokuma, D. Malick, A. Rabuck, K. Raghavachari, J. Foresman, J. Cioslowski, J. Ortiz, B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, G. Gomperts, R. Martin, D. Fox, T. Keith, M. Al-Laham, C. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. Gill, B. Johnson, W. Chen, M. Wong, J. Andres, M. Head-Gordon, E. Replogle, J. Pople, Gaussian 98 (Revision A.7), Gaussian Inc., Pittsburgh, PA, 1998.
[69] H.-J. Werner, P.J. Knowles, R. Lindh, M. Schütz, P. Celani, T. Korona, F.R. Manby, G. Rauhut, R.D. Amos, A. Bernhardsson, A. Berning, D.L. Cooper, M.J.O. Deegan, A.J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, A.W. Lloyd, S.J. McNicholas, W. Meyer, M.E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, U. Schumann, H. Stoll, A.J. Stone, R. Tarroni, T. Thorsteinsson, Molpro, version 2002.6, A package of ab initio programs, 2003. Available from: <http://www.molpro.net>.