Theor Chim Acta (1992) 84: 55–83

Theoretica Chimica Acta

On the global convergence of MCSCF wave function optimization: The method of trigonometric interpolation*

Ron Shepard

Theoretical Chemistry Group, Chemistry Division, Argonne National Laboratory, Argonne, IL 60439, USA

Received August 21, 1991/Accepted January 23, 1992

Summary. A new method of MCSCF wave function optimization is presented. This method is based on a nonlinear transformation of the wave function variation coordinates along with the construction of a global interpolating function. This interpolating function is constructed for each MCSCF iteration in such a way that it reproduces certain known behavior of the exact energy function. It reproduces exactly the energy, gradient, and hessian at the expansion point, at an infinite number of isolated points, and at points on the surfaces of an infinite number of nested multidimensional balls within the wave function variational space. The optimization of the wave function correction parameters on this interpolating function does not require integral transformations or density matrix constructions, although one-index transformation and transition density matrix techniques may be used if desired. The nonlinear coordinate transformations, along with the necessary derivatives, are computed with simple matrix operations, and require only $O(N_{orb}^3)$ effort. The new method differs from previous optimization methods in several respects. (1) It reproduces certain behavior of the exact energy function that is not displayed by previous approaches. (2) The orbital-state coupling is included explicitly via the partitioned orbital hessian matrix. (3) The minimization of the approximate energy function is simpler than with previous similar approaches. (4) The treatment of redundant orbital rotations is straightforward, since the exact and approximate energy functions display the same qualitative behavior with respect to these wave function variations. (5) Finally, the present method may be implemented as a simple extension to essentially any existing second-order MCSCF code, the required changes being localized within a rather small part of the overall iterative procedure. Examples of the convergence of the new method are presented, along with numerical demonstrations of some of the relevant features of the exact and interpolated energy functions.

Key words: MCSCF – Optimization – Convergence – Nonlinear – Interpolation – Global

* Dedicated in honor of Prof. Klaus Ruedenberg

1. Introduction

Professor Ruedenberg and his research group have made many significant contributions to the methodology of the computation and to the interpretation of MCSCF wave functions. The author's own work in this field has been influenced particularly by Professor Ruedenberg's contributions in the late 1970s when, as a student, I had the opportunity to see him speak on several occasions on the FORS (Fully Optimized Reaction Space) wave function expansion, and on MCSCF wave function optimization using natural orbitals from the single-excitation configuration interaction wave function. It is a particular pleasure to have made the first presentation of this MCSCF method at the Ames Laboratory, Iowa State University Workshop in Honor of Professor Ruedenberg, "*Ab Initio* Methods in Quantum Chemistry" and to include this contribution in the proceedings of this meeting.

The problem of local convergence of MCSCF wave function optimization was essentially solved with the development of the exponential operator formalism [1–5]. This provided the basis for a coupled treatment of the orbital rotation parameters and the configuration state function (CSF) rotation parameters, and led to the first implementations of fully second-order convergent optimization procedures applicable to general MCSCF wave function expansions [4–6]. This formalism, along with its accompanying commutator expansions, also provided the basis for a more complete understanding of redundant orbital rotations [7], of the relation between optimization methods based on Taylor expansions and those based on hamiltonian operator representations [8, 9], of the relation between response theory and the multiconfigurational generalization of coupled-perturbed Hartree–Fock theory [1, 10], and of the nature of analytical energy derivatives [11–15].

The use of matrix partitioning techniques, leading to the analysis based on the *partitioned orbital hessian* matrix [8, 16], provided further insight into the nature of the MCSCF wave function, particularly for excited electronic states. For excited states, it was demonstrated that the minimization of the appropriate hamiltonian matrix *eigenvalue* with respect to orbital rotations was not equivalent to minimization of the *expectation value* with respect to orbital rotations. This equivalence had historically been a tacit assumption of MCSCF orbital optimization. The correct conceptual connection is between the minimization of the hamiltonian eigenvalue and the minimization of an effective energy expression, the first- and second-order terms of which are the orbital gradient and the partitioned orbital hessian. This subtle distinction between an *expectation value* and an *eigenvalue* results in the somewhat surprising condition that an eigenvalue minimum may, in fact, correspond to a maximum of the expectation value with respect to a particular orbital rotation variation [7, 8].

However, the straightforward use of second-order expansions does not eliminate the global convergence problems associated with MCSCF wave function optimization. Some aspects of these problems have been addressed using various approaches. Mode-specific level-shifting methods [17], approximate Super-CI methods [8, 16] (also called augmented hessian methods [5], rational function methods [7], and norm-extended methods [18]), the iterative and perturbative inclusion of specific higher-order terms [18], and nonlinear changes of expansion variables [6, 19] are examples of some of these approaches. All of these approaches are closely related, particularly from the point of view that an *approximate energy function* is constructed in some manner, the optimization of

which is easier than the optimization of the *exact energy function*, and the wave function corrections derived from this approximate energy function are applied iteratively to the exact wave function until convergence is achieved.

The present method employs the idea of a nonlinear change of coordinate in order to reproduce within an approximate energy function certain known properties of the exact energy function. It differs from previous optimization methods in several respects. (1) It reproduces certain behavior of the exact energy function that is not displayed by previous approaches. (2) The orbital-state coupling is included explicitly via the partitioned orbital hessian matrix. (3) The minimization of the approximate energy function is simpler than with previous similar approaches. (4) The treatment of redundant orbital rotations is straightforward, since the exact and approximate energy functions display the same qualitative behavior with respect to these wave function variations. (5) Finally, the present method may be implemented as a simple extension to essentially any existing second-order MCSCF code, the required changes being localized within a rather small part of the overall iterative procedure.

The following section contains a brief summary of the previous MCSCF methods that are most closely related to the new method, herein called *multidimensional trigonometric interpolation*. Section 3 describes this new method and the associated nonlinear coordinate transformation. Section 4 extends the discussion of exact and approximate energy functions to include the effects of redundant orbital rotation variables. Section 5 contains numerical examples of some of the relevant issues related to the method of trigonometric interpolation. Examples of MCSCF optimization using the new method are given in Sect. 6. Concluding remarks and discussions of future directions of research are then given in Sect. 7.

2. Review of previous work

The exponential operator parameterization of the orbital rotation space is based on the intimate connections between the hilbert space operator $\exp(K)$, the matrix representation of this operator in a many-electron basis, $\langle m|\exp(K)|n\rangle$, and the orbital rotation matrix $\exp(\boldsymbol{K})$ of dimension N_{orb}. The orbital rotations are defined by the matrix:

$$\boldsymbol{U} = \exp(\boldsymbol{K}) = \mathbf{1} + \boldsymbol{K} + \tfrac{1}{2}\boldsymbol{K}^2 + \cdots + \frac{1}{n!}\boldsymbol{K}^n + \cdots . \tag{2.1}$$

As usual for molecular calculations, it will be assumed that $\boldsymbol{K}$, and therefore $\boldsymbol{U}$, are real matrices. If $\boldsymbol{K}$ is antisymmetric (i.e. $\boldsymbol{K} = -\boldsymbol{K}^T$; also, such matrices are called skew-symmetric), then $\boldsymbol{U}$ is an orthogonal matrix with $\text{Det}(\boldsymbol{U}) = +1$. The operator $\exp(K)$ is defined as:

$$\exp(K) = 1 + K + \tfrac{1}{2}K^2 + \cdots \frac{1}{n!}K^n + \cdots \tag{2.2}$$

in which the one-electron operator K is defined in terms of the matrix elements of $\boldsymbol{K}$ as:

$$K = \sum_{r=1}^{N_{orb}} \sum_{s=1}^{N_{orb}} K_{rs} E_{rs} . \tag{2.3}$$

The operators $E_{rs} = a^{\dagger}_{r\alpha}a_{s\alpha} + a^{\dagger}_{r\beta}a_{s\beta}$ are usually called "generators" of the unitary group due to their commutation properties (see Ref. [7] and references therein and additionally Ref. [20] for further details). As indicated, these generators may be represented in terms of electron creation and annihilation operators, thereby allowing the hamiltonian operator to be written as:

$$H = \sum_{rs} h_{rs}E_{rs} + \tfrac{1}{2}\sum_{rstu} g_{rstu}e_{rstu} \tag{2.4}$$

with $e_{rstu} = E_{rs}E_{tu} - \delta_{st}E_{ru}$ being the normal-order product, and with the summations ranging over the entire orthonormal orbital basis. The arrays $\boldsymbol{h}$ and $\boldsymbol{g}$ are the one-electron hamiltonian and two-electron repulsion integrals respectively. The rotations within the CSF space are parameterized with the unitary operator $\exp(P)$, defined with:

$$P = \sum_{n(\neq 0)} p_n(|n\rangle\langle 0| - |0\rangle\langle n|) \tag{2.5}$$

in which $|0\rangle$ is the reference wave function and the summation ranges over the orthogonal complement to $|0\rangle$ within the MCSCF expansion space. Computationally, it is most convenient to represent the operator P in the overcomplete CSF basis, or the closely related, linearly dependent, projected basis. The merits of various bases for the representation of the operator P are discussed in detail in Ref. [7].

The unitary operators $\exp(K)$ and $\exp(P)$ allow the exact MCSCF energy expression in terms of the rotation parameters $\{K_{rs}, p_n\}$ to be written as:

$$E^{Exact}(\boldsymbol{K}, \boldsymbol{p}) = \langle 0|\exp(-P)\exp(-K)H\exp(K)\exp(P)|0\rangle. \tag{2.6}$$

Expansion through second order in these parameters leads to the matrix expression

$$E^{[2]}(\boldsymbol{\kappa}, \boldsymbol{p}) = E(0) + (\boldsymbol{\kappa}^T\boldsymbol{p}^T)\begin{pmatrix}\boldsymbol{w}\\ \boldsymbol{v}\end{pmatrix} + \tfrac{1}{2}(\boldsymbol{\kappa}^T\boldsymbol{p}^T)\begin{pmatrix}\boldsymbol{B} & \boldsymbol{C}\\ \boldsymbol{C}^T & \boldsymbol{M}\end{pmatrix}\begin{pmatrix}\boldsymbol{\kappa}\\ \boldsymbol{p}\end{pmatrix}. \tag{2.7}$$

As customary, the vector $\boldsymbol{\kappa}$ contains only the unique, essential (i.e. nonredundant), elements of the orbital rotation array $\boldsymbol{K}$. The vectors $\boldsymbol{w}$ and $\boldsymbol{v}$ are the gradient vectors for the orbital and CSF rotation parameters respectively. The matrix $\boldsymbol{B}$ is the orbital-orbital hessian matrix, $\boldsymbol{C}$ is the orbital-CSF block of the hessian matrix, and $\boldsymbol{M}$ is the CSF-CSF hessian matrix.

Given a set of CSF coefficients and orbitals to define a reference wave function, the construction of the gradient vector and hessian matrix provides an approximate energy function, $E^{[2]}(\boldsymbol{\kappa}, \boldsymbol{p})$, which may be used to determine the correction vectors $\boldsymbol{\kappa}$ and $\boldsymbol{p}$. These corrections may be applied to the reference wave function (i.e. by applying the orbital transformation matrix $\boldsymbol{U}$ to the orbitals, and the operator $\exp(P)$ to the CSF coefficients) in order to define a new reference wave function. This straightforward procedure is the *wave function hessian Newton–Raphson* (WNR) optimization method.

If the optimal CSF vector is determined for the reference orbitals, then the CSF gradient vector $\boldsymbol{v}$ vanishes. The CSF correction vector $\boldsymbol{p}$ from optimization of Eq. (2.7) then satisfies the relation:

$$\boldsymbol{p} = -\boldsymbol{M}^{-1}\boldsymbol{C}^T\boldsymbol{\kappa}. \tag{2.8}$$

When substituted into $E^{[2]}(\boldsymbol{\kappa}, \boldsymbol{p})$, this leads to an energy expression that depends explicitly only on the orbital rotation parameters.

$$E^{[2]}(\boldsymbol{\kappa}) = E(0) + \boldsymbol{\kappa}^T \boldsymbol{w} + \tfrac{1}{2}\boldsymbol{\kappa}^T(\boldsymbol{B} - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)\boldsymbol{\kappa}. \tag{2.9}$$

The matrix $(\boldsymbol{B} - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)$ is the *partitioned orbital hessian matrix*. These matrix elements are the second derivatives of the hamiltonian *eigenvalue* with respect to pairs of orbital rotations. In comparison, the matrix $\boldsymbol{B}$ consists of the second derivatives of the *expectation value* for fixed CSF coefficients with respect to pairs of orbital rotations. Although the distinction between the meanings of these two arrays is rather subtle, this difference leads to important consequences. The most obvious of these differences is that in the neighborhood of a correct solution, the partitioned orbital hessian matrix is always positive semidefinite (i.e. there are no negative eigenvalues). This property is essentially a restatement of the eigenvalue minimization definition of the MCSCF wave function. That this is true for ground electronic states is not surprising; the fact that it is also true for excited electronic states leads to great simplifications in the optimization of excited states. (See Refs. [7] and [8] for further discussion of this issue.)

The function, $E^{[2]}(\boldsymbol{\kappa})$, may also be considered as the second-order perturbation approximation of the eigenvalue of the hamiltonian matrix with elements $H_{mn} = \langle m|\exp(-K)H\exp(K)|n\rangle$. This relation is useful when considering the form of higher-order coupling between the $\boldsymbol{p}$ and $\boldsymbol{\kappa}$ vectors [7].

The straightforward use of $E^{[2]}(\boldsymbol{\kappa})$ to determine the orbital corrections is called the *partitioned orbital hessian Newton–Raphson* (PNR) optimization method. It should perhaps be mentioned that it is not necessary to actually invert the potentially large matrix $\boldsymbol{M}$ in order to use Eq. (2.9) as an approximating function. It is primarily for notational convenience that Eq. (2.9) is written in the folded form. Reference [7] may be consulted for details of the solutions to various orbital optimization equations using only subspace representations of the subblocks of the hessian matrix, the construction of which requires only the computation of matrix-vector products.

When not in the neighborhood of a correct solution, the partitioned orbital hessian matrix may possess negative eigenvalues. However, the approximate energy expression $E^{[2]}(\boldsymbol{\kappa})$ cannot be used directly to move to regions of wave function variational space which have the desired positive-semidefinite hessian property. The nature of a second-order expansion is such that along any line (e.g. parameterized as $t\boldsymbol{\kappa}$ for a fixed direction $\boldsymbol{\kappa}$) there is either an energy maximum or a minimum, but not both. The existence of a maximum along some line (implying at least one negative partitioned orbital hessian eigenvalue) indicates that the trial function is not in the neighborhood of a correct solution, but moving downhill in search of a minimum does not help since $E^{[2]}(t\boldsymbol{\kappa}) \to -\infty$ as $t \to \pm\infty$. Specifically, the approximating function $E^{[2]}(t\boldsymbol{\kappa})$ is unbound with constant second derivatives, and therefore it does not mimic, in these respects, the exact energy expression, which is not only bound but oscillatory along certain direction vectors $\boldsymbol{\kappa}$.

To overcome this limitation of second-order expansions, several approaches have been taken. Among the first were mode-specific level-shifting methods [17]. The central idea of these methods is the identification of a subspace within which the accuracy of the local expansion is sacrificed in order to achieve qualitatively correct energy behavior in some nonlocal region. Such techniques had met with some success within the context of first-order-convergent single-configuration (SCF) schemes. However, these techniques require the input of several empirical

parameters for the subspace selection and subsequent local modifications. These parameters must be chosen with great care, else the overall convergence rate deteriorates inordinately. Generally, such schemes were found unsatisfactory, and their use has all but disappeared.

Other approaches that address the global-convergence problems were based on the formal similarities between the hessian matrix and the matrix representation of the hamiltonian operator in the first-order variational space. This space is defined as $\{(E_{pq} - E_{qp})|0\rangle, |0\rangle, |n\rangle\}$ with $(p > q)$ essential. In modern terminology, this would be called an "internally contracted" single-excitation CI calculation. Experience had shown that these "Super-CI" (SCI) or "single-excitation-CI" optimization techniques were very reliable, yielding 2–3 correct decimals of energy in only a few iterations, starting with almost any set of initial orbitals [21–25]. However, SCI procedures ultimately display only first-order convergence, so convergence to 6–8 decimals in the wave function (equivalent to 12–16 decimals in the energy), which is routine for second-order procedures, is impractical. Another difficulty with the traditional SCI approach was that the contraction of the hamiltonian operator was achieved only with some difficulty; although the methodology was never completely mature, the effort for the overall procedure scaled roughly as the product, $(N_{csf}^2 N_{pq}^2)$, where N_{csf} is the MCSCF expansion length and N_{pq} is the number of orbital rotations. This is in contrast to the commutator-based second-order methods whose effort typically scales only as the sum $(N_{csf}^2 + N_{pq}^2)$ for the corresponding hessian and gradient construction steps.

These observations led several researchers to the approximate SCI methods [9, 5, 16]. Like the true SCI methods, these methods result in an eigenvalue equation, with the wave function corrections being determined from the elements of an eigenvector. In Ref. [7], general forms for these energy expressions are introduced by replacing the truncated second-order expansion with a rational function approximation. The form of the rational function is such that the first and second derivatives at the expansion point are unchanged (and, therefore, exact). Consequently, these methods display rigorous second-order convergence, in contrast to the true internally contracted SCI methods which are only first-order convergent. Furthermore, the $t \to \pm\infty$ asymptotic behavior is damped in such a manner as to keep the approximate rational energy function bound (although the bound is somewhat weaker than that of the internally-contracted SCI methods). In the case of the approximate SCI method based on the partitioned orbital hessian matrix (the PSCI method), this also ensures the existence of a single unique minimum of the approximating function, greatly simplifying excited state wave function optimization. The convergence neighborhood using the rational function is much larger than that of the truncated $E^{[2]}$, allowing convergence to be obtained with no empirical parameters in many cases for which the straightforward Newton–Raphson method is divergent.

However, even the rational function approach does not eliminate all convergence problems, and additional level-shifting and/or vector-scaling is typically employed particularly in the initial MCSCF iterations [7]. The associated idea of a "trust radius" [26] has also been applied in order to limit the step size to within the well-approximated local region of wave function variation space [18]. In this approach, the level-shifting parameter is dynamically adjusted until the magnitude of the computed correction vector is less than an empirically adjusted value. The vast majority of MCSCF wave function optimization procedures use such

rational-function-based approaches, both when far from the desired solution and in the local neighborhood of the desired solution.

Finally, the last approach to the global convergence problem relevant to the present work is the nonlinear exponential coordinate transformation [6, 19]. In this method, the energy is expanded to second order in the parameters T_{ts} defined by:

$$\boldsymbol{T} = \boldsymbol{U} - \mathbf{1} = (\exp(\boldsymbol{K}) - \mathbf{1}) = \boldsymbol{K} + \tfrac{1}{2}\boldsymbol{K}^2 + \cdots \frac{1}{n!}\boldsymbol{K}^n + \cdots . \tag{2.10}$$

This approach has the feature that the approximate energy function $E^{[T]}$ is exact at the infinite number of points which satisfy the relation $\exp(\boldsymbol{K}) = \mathbf{1}$. By construction, the approximate energy function reproduces the exact first and second derivatives at this infinite number of isolated points also. Due to these features, $E^{[T]}$ may be regarded as a "global" approximating energy function, whereas the truncated energy functions and rational functions are only valid in a "local" region about the reference wave function.

However, as implemented in [19], the approximating function $E^{[T]}$ does not have the same redundant orbital rotation behavior as the exact energy expression. This results in unsatisfactory behavior when the wave function contains "almost" redundant variables, such as active orbitals with occupations only slightly less than two. In fact, as originally implemented, $E^{[T]}$ was used only for the initial iterations, the final iterations being performed with one of the previously described Newton–Raphson procedures which display the correct redundant orbital rotation behavior [6]. This characteristic of the $E^{[T]}$ method has been addressed [19], after noting that E^{Exact} is a fourth-order function of $\boldsymbol{T}$, by including a sufficient subset of the third- and fourth-order expansion terms into the $E^{[T]}$ approximating function to overcome these convergence deficiencies.

Another deficiency of this approach is that the true energy function is oscillatory along certain lines $\boldsymbol{K}$, but the period should be π rather than 2π. For example, in the simple case of one orbital rotation parameter, the two-orbital subblock of the transformation matrix may be written as

$$\boldsymbol{U} = \begin{pmatrix} \cos(\kappa) & -\sin(\kappa) \\ \sin(\kappa) & \cos(\kappa) \end{pmatrix}. \tag{2.11}$$

At $k = \pi$ the exact energy, gradient, and hessian are identical to those at $\kappa = 0$, since these quantities depend only on the relative, not the absolute, phase of the two orbitals. However, since

$$\boldsymbol{T}(0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

but

$$\boldsymbol{T}(\pi) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix},$$

this periodicity is not automatically reproduced by an approximate energy expression based on a truncated $\boldsymbol{T}$ expansion. This incorrect behavior of $E^{[T]}$ with respect to periodicity is displayed most clearly in Fig. 3 of Ref. [19].

3. Multidimensional trigonometric interpolation

From the discussion of the previous section, it is clear that a global approximating function is desirable, but to the extent possible, such a function should display the correct oscillatory behavior, the correct dependence on redundant orbital rotations, and it should be based on commutator expansions in order to provide smooth connections to other related theories (response theory, analytic energy derivative theory, etc.).

General coordinate transformations of the form $F(\boldsymbol{K})$, where F is analytic, are now considered. A more complete discussion of the choice of $F(\boldsymbol{K})$ is given in Ref. [27]. The relevant conclusions are summarized here.

$\boldsymbol{K}$ is a normal matrix (i.e. it commutes with its adjoint), and may be diagonalized by a unitary matrix $\boldsymbol{V}$. The eigenvalues of $\boldsymbol{K}$ are purely imaginary and occur in complex-conjugate pairs, $(\ldots i\lambda_j, -i\lambda_j \ldots)$; the associated eigenvectors may also be paired, $(\ldots \boldsymbol{v}_j, \boldsymbol{v}_j^* \ldots)$. For odd matrix dimensions, at least one eigenvalue of $\boldsymbol{K}$ is zero, and its associated eigenvector may be chosen to be real. These observations allow the matrix $\boldsymbol{K}$ to be written in the forms:

$$\boldsymbol{K} = \boldsymbol{V}(i\boldsymbol{\lambda})\boldsymbol{V}^\dagger \tag{3.1}$$

$$= \sum_{j=1}^{INT[N_{orb}/2]} i\lambda_j(\boldsymbol{v}_j\boldsymbol{v}_j^\dagger - \boldsymbol{v}_j^*\boldsymbol{v}_j^{*\dagger}) + (0)\boldsymbol{v}_0\boldsymbol{v}_0^\dagger \tag{3.2}$$

where the last contribution in Eq. (3.2), included here for definiteness, only appears for odd dimensions. For notational simplicity, it is hereafter assumed that N_{orb}, the dimension of $\boldsymbol{K}$, is even in the remainder of this section, thereby allowing the (0)-eigenvalue term to be dropped in the following equations, although the final results will be correct in any case. It may be verified that each of the rank-2 contributions to $\boldsymbol{K}$ in the summation of Eq. (3.2) of the form, $i\lambda_j(\boldsymbol{v}_j\boldsymbol{v}_j^\dagger - \boldsymbol{v}_j^*\boldsymbol{v}_j^{*\dagger})$, is a real antisymmetric matrix.

The expansion of the analytic function F may be written:

$$F(\boldsymbol{K}) = F_0\boldsymbol{1} + F_1\boldsymbol{K} + \frac{F_2}{2}\boldsymbol{K}^2 + \cdots \frac{F_n}{n!}\boldsymbol{K}^n + \cdots . \tag{3.3}$$

Following the heuristic derivation of the rational function approaches, one of the truncated second-order energy expansions will be modified. The expansion based on the partitioned orbital hessian matrix of Eq. (2.9) is considered explicitly; the derivation of the wave function hessian expressions follows by analogy. Consider substitutions of the form $F_{rs}(\boldsymbol{K}) \leftarrow K_{rs}$. In a one-dimensional example of $E^{[2]}(\boldsymbol{\kappa})$, i.e. $f(\kappa) \leftarrow \kappa$, this results in

$$E^{[F]}(\kappa) = E(0) + wf + \tfrac{1}{2}(B - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)f^2 \tag{3.4}$$

$$E^{[F]}(\kappa)' = w(f') + (B - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)f(f') \tag{3.5}$$

$$E^{[F]}(\kappa)'' = w(f'') + (B - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)(f(f'') + (f')^2). \tag{3.6}$$

B and w are scalars, and $\boldsymbol{C}$ is a row vector in these equations. The derivatives, denoted with primes, are with respect to the scalar κ in these equations. In order for $E^{[F]}(0) = E(0)$ to be satisfied for arbitrary w and $(B - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)$ in Eq. (3.4), it follows that $f(0) = F_0 = 0$. Similarly, for $E^{[F]}(0)' = w$ in Eq. (3.5), it follows that $f(0)' = F_1 = 1$. Finally, for $E^{[F]}(0)'' = (B - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)$ in Eq. (3.6), it follows that $f(0)'' = F_2 = 0$. These conditions carry over to the multidimen-

sional approximating function $E^{[F]}(\boldsymbol{\kappa})$, resulting in the conditions $F_0 = 0$, $F_1 = 1$, and $F_2 = 0$ in the matrix expansion of Eq. (3.3). In fact, the conditions on F_0 and F_2 are satisfied automatically if $F(\boldsymbol{K})$ is simply chosen to be antisymmetric, simplifying the $F_{rs} \leftarrow K_{rs}$ association. (In general an odd scalar function $F(x)$ implies an antisymmetric matrix function $F(\boldsymbol{K})$; this association will be used interchangeably in the following discussions.)

It may be noted that the $\boldsymbol{T}$ expansion described in the previous section already violates the antisymmetry requirement, along with the $F_2 = 0$ condition, and therefore is not a candidate expansion form for the present work. This is one reason for the relatively complicated iterative procedure required to optimize $E^{[T]}$; much effort is devoted to "correcting" the second-order terms induced by the nonzero F_2 terms in $T(\boldsymbol{K})$ at the expansion point $\boldsymbol{K} = \boldsymbol{0}$.

Finally, it is desired for the approximate function to be oscillatory in $\boldsymbol{K} = (\boldsymbol{K}_0 + t\Delta)$ with period $t = \pi$. There are, of course, an unlimited number of ways to define such a function. In order to mimic $E^{Exact}[\exp(\boldsymbol{K})]$ as closely as possible, the functions $\exp(2\boldsymbol{K})$ and $\exp(-2\boldsymbol{K})$ might be considered. Each of these would have the correct periodicity, but they possess nonzero F_0 and F_2 terms. However, these undesired contributions may be eliminated by taking the sum $(\exp(2\boldsymbol{K}) - \exp(-2\boldsymbol{K}))$. This function satisfies all of the requirements except for the $F_1 = 1$ condition, which is easily restored by scaling to give the final coordinate transformation:

$$F(\boldsymbol{K}) = \tfrac{1}{4}(\exp(2\boldsymbol{K}) - \exp(-2\boldsymbol{K})) = \tfrac{1}{2}\sinh(2\boldsymbol{K}). \tag{3.7}$$

The relation $\sinh(ix) = i\sin(x)$ allows the antisymmetric matrix $F(\boldsymbol{K})$ to be written as:

$$F(\boldsymbol{K}) = \tfrac{1}{2}\boldsymbol{V}i\sin(2\boldsymbol{\lambda})\boldsymbol{V}^{\dagger} \tag{3.8}$$

displaying clearly the oscillatory behavior of the function $F(\boldsymbol{K})$.

Defining the vector $\boldsymbol{f}$ as the essential, unique elements of the antisymmetric matrix $F(\boldsymbol{K})$ allows the approximate global interpolating function to be written as:

$$E^{TIF}(\boldsymbol{\kappa}) = E(0) + \boldsymbol{f}^T\boldsymbol{w} + \tfrac{1}{2}\boldsymbol{f}^T(\boldsymbol{B} - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)\boldsymbol{f}. \tag{3.9}$$

Although only second-order in $\boldsymbol{f}$, this approximate energy function is infinite-order in the orbital rotation parameters $\boldsymbol{K}$. Note also that, due to the use of the partitioned orbital hessian, this functional form implies a relation $\boldsymbol{f}_{csf} = -\boldsymbol{M}^{-1}\boldsymbol{C}^T\boldsymbol{f}$ (compare to Eq. (2.8)), showing that the (implicit) CSF rotation parameters $\boldsymbol{p} = F_{csf}^{-1}(\boldsymbol{f}_{csf})$ (with an herein unspecified mapping $F_{csf}^{-1}(\ )$) are also infinite-order and oscillatory in the rotation parameters $\boldsymbol{K}$. It should be stressed that the form of Eq. (3.7) is by no means unique. Indeed general combinations of terms of the form $1/2n\sinh(2n\boldsymbol{K})$ may be chosen to satisfy the above, rather liberal, constraints imposed on the coordinate transformation.

Due to the form of Eq. (3.8), the method currently described is herein called *multidimensional trigonometric interpolation*, or more briefly, *trigonometric interpolation*, and the interpolating function of Eq. (3.9) is called the *trigonometric interpolating function* (TIF).

The one-dimensional function, Eq. (3.4), may now be written explicitly as

$$E^{TIF}(\kappa) = E(0) + \tfrac{1}{2}w\sin(2\kappa) + \tfrac{1}{8}(B - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)\sin^2(2\kappa) \tag{3.10}$$

$$= E(0) + \tfrac{1}{2}w\sin(2\kappa) + \tfrac{1}{16}(B - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)(1 - \cos(4\kappa)). \tag{3.11}$$

The last form in particular reveals that the κ dependence of the interpolating function consists of a combination of two terms: a $\sin(2\kappa)$ term with period π whose contribution is determined by the magnitude of the gradient w, and a $\cos(4\kappa)$ term of period $\pi/2$ whose contribution is determined by the magnitude of the (scalar) second-derivative $(B - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)$. Normally it would be expected that when far from convergence, the gradient component would dominate the functional behavior, and when close to convergence, the shorter-period (higher-frequency) hessian term would dominate. However, as will be seen later, the hessian terms are sometimes dominant even when far from convergence, such as near maxima or (in higher dimensions) saddle points. It can also be seen from Eq. (3.11) that at convergence, $w = 0$, the interpolating function will have purely period $\pi/2$ behavior. In most cases, such κ behavior would be spurious, as E^{Exact} is expected usually to display period π behavior. This is not detrimental to the optimization problem for which E^{TIF} is herein developed, but it should be taken into consideration if E^{TIF} were to be extended beyond this limited application (e.g. to MCSCF response theory application). Another feature of E^{TIF}, which is apparent from inspection of Eq. (3.10), is that a minimum is closer than $\pi/4$ to the expansion point only when $2|w| < |(B - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)|$; otherwise, the gradient terms dominate E^{TIF}, and the "coordinate-induced" minimum at $\pm\pi/4$ (with the sign determined by the signs of $(B - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)$ and w) is the minimization solution.

The above arguments are now extended to the general multidimensional case. Suppose a solution has been found to the orbital transformation equation $\exp(t\boldsymbol{K}) = \boldsymbol{S}$ in which $\boldsymbol{S}$ is a diagonal sign matrix with elements $S_{mn} = \pm\delta_{mn}$. The variables to be considered are the line *direction*, determined by the matrix $\boldsymbol{K}$, and the distance along this line from the origin parameterized by the scalar t. At points $t\boldsymbol{K}$ which satisfy the general equation, it follows that $E^{Exact}(0) = E^{Exact}(t\boldsymbol{K})$ since, as argued previously, the energy is independent of orbital phases. The most general solutions to this equation are unknown to the author at the present time. In the absence of a general solution, several special cases will be discussed.

Examine first E^{Exact} for a two-parameter orbital transformation of the form:

$$\boldsymbol{K} = k_{\alpha\beta}\boldsymbol{e}^{\alpha\beta} + k_{\mu\nu}\boldsymbol{e}^{\mu\nu}. \tag{3.12}$$

The matrices of the form $\boldsymbol{e}^{\alpha\beta}$ are *infinitesimal generator coordinates* of the rotation group, and in order to avoid confusion with the unitary group generators E_{pq} which are hilbert space operators, these matrices will herein be called simply *rotation coordinates*. These are antisymmetric matrices constructed from unit vectors as $\boldsymbol{e}^{\alpha\beta} = (\boldsymbol{e}^{\alpha}(\boldsymbol{e}^{\beta})^T - \boldsymbol{e}^{\beta}(\boldsymbol{e}^{\alpha})^T)$ with, for example, $\boldsymbol{e}^{\alpha}$ being the α-th column of the unit matrix $(\boldsymbol{e}^{\alpha})_j = \delta_{j\alpha}$. Consider first the cases in which none of the rotation coordinate labels are coincident, i.e. $\alpha \neq \beta \neq \mu \neq \nu$. In this case, $[\boldsymbol{e}^{\alpha\beta}, \boldsymbol{e}^{\mu\nu}] = \boldsymbol{0}$, and the associated orbital transformation factors according to:

$$\boldsymbol{U} = \exp(\boldsymbol{K}) = \exp(k_{\alpha\beta}\boldsymbol{e}^{\alpha\beta})\exp(k_{\mu\nu}\boldsymbol{e}^{\mu\nu}) = \exp(k_{\mu\nu}\boldsymbol{e}^{\mu\nu})\exp(k_{\alpha\beta}\boldsymbol{e}^{\alpha\beta}) \tag{3.13}$$

allowing the two individual factors to be considered independently. Consequently, by generalizing the previous single-parameter arguments, $E^{Exact}(k_{\alpha\beta} = 0, k_{\mu\nu} = 0) = E^{Exact}(m\pi, n\pi)$ for arbitrary integers m, n. This defines a regular 2-dimensional grid of knots (using spline terminology) on the plane defined by cartesian product of the rotation coordinates $\boldsymbol{e}^{\alpha\beta} \otimes \boldsymbol{e}^{\mu\nu}$, and the orbital transformation matrix is sinusoidal on this grid with period 2π. This results in an oscillatory E^{Exact} which is periodic in this plane with period π.

In order to determine the behavior of the transformed coordinates $F(\boldsymbol{K}) = \frac{1}{2}\sinh(2\boldsymbol{K})$, it is recognized that since $\boldsymbol{e}^{\alpha\beta}$ and $\boldsymbol{e}^{\mu\nu}$ commute they may be diagonalized simultaneously by some unitary matrix $\boldsymbol{V}$. This allows $\sinh(2\boldsymbol{K})$ to be written in the factored form:

$$\sinh(2\boldsymbol{K}) = \sinh(2k_{\alpha\beta}\boldsymbol{e}^{\alpha\beta} + 2k_{\mu\nu}\boldsymbol{e}^{\mu\nu}) \tag{3.14}$$

$$= \cosh(2k_{\alpha\beta}\boldsymbol{e}^{\alpha\beta})\sinh(2k_{\mu\nu}\boldsymbol{e}^{\mu\nu}) + \sinh(2k_{\alpha\beta}\boldsymbol{e}^{\alpha\beta})\cosh(2k_{\mu\nu}\boldsymbol{e}^{\mu\nu}). \tag{3.15}$$

Recognizing the identity, $\cosh(ix) = \cos(x)$, this factored form demonstrates that the transformed coordinates are periodic along both coordinate directions within the plane with period π. This behavior results in an oscillatory E^{TIF} on the plane with period π. Therefore in the plane, both E^{Exact} and E^{TIF} display oscillatory behavior with period π.

This argument is now extended to the general multidimensional case by adding all possible patterns of mutually commuting rotation coordinate matrices in the definition of $\boldsymbol{K}$. These *grid patterns* may be labeled by taking all possible sets of pairs of integers from the range $[1 \ldots N_{orb}]$. This results in

$$N^{[Grid]}(N_{orb}) = \begin{cases} (N_{orb}-1)!! & ;N_{orb}\,even \\ N_{orb}!! & ;N_{orb}\,odd \end{cases} \tag{3.16}$$

possible grid patterns for a given matrix dimension N_{orb}. Each such grid pattern defines a cartesian product subspace of dimension $INT[N_{orb}/2]$, and within each of these subspaces both E^{Exact} and E^{TIF} are oscillatory with period π along each rotation coordinate direction.

To clarify this subspace concept, consider some specific examples. In a 2-orbital case, there is $N^{[Grid]}(2) = 1$ subspace of dimensionality $N_{orb}/2 = 1$ (a line). This is the trivial case considered previously. In a 3-orbital case, there are $N^{[Grid]}(3) = 3$ cartesian subspaces, each of dimensionality $INT[N_{orb}/2] = 1$; in other words there are three lines within the full space along which both E^{Exact} and E^{TIF} are oscillatory with period π. These lines are along the coordinate axes $\boldsymbol{e}^{21}$, $\boldsymbol{e}^{31}$, and $\boldsymbol{e}^{32}$. In a 4-orbital case, there are $N^{[Grid]}(4) = 3$ cartesian subspaces, each of dimensionality $N_{orb}/2 = 2$; in other words, there are three planes within the full space on which both E^{Exact} and E^{TIF} display oscillatory behavior with period π. These planes are specified by the rotation coordinate pairs $(\boldsymbol{e}^{21} \otimes \boldsymbol{e}^{43})$, $(\boldsymbol{e}^{31} \otimes \boldsymbol{e}^{42})$, and $(\boldsymbol{e}^{41} \otimes \boldsymbol{e}^{32})$.

Next, the energy behavior along particular lines consisting of linear combinations of *noncommuting* rotation coordinates is examined, and it is shown that the approximating function displays the qualitatively correct oscillatory behavior. Specifically, consider the behavior of $E^{Exact}(\boldsymbol{K})$, $E^{TIF}(\boldsymbol{K})$, $\exp(\boldsymbol{K})$, and $F(\boldsymbol{K})$ along the lines parameterized with the scalar t according to $\boldsymbol{K} = (\boldsymbol{K}_0 + t\boldsymbol{\Delta})$. The real, antisymmetric matrices $t\boldsymbol{\Delta}$ are constructed as:

$$t\boldsymbol{\Delta} = t\boldsymbol{V}(i\boldsymbol{m})\boldsymbol{V}^{\dagger} = t\sum_j im_j(\boldsymbol{v}_j\boldsymbol{v}_j^{\dagger} - \boldsymbol{v}_j^{*}\boldsymbol{v}_j^{*\dagger}) \tag{3.17}$$

with $\{m_j \mid j = 1, 2 \ldots N_{orb}/2\}$ being a set of integers, $\boldsymbol{m} = \mathrm{diag}(\ldots m_j, -m_j, \ldots)$, and $\boldsymbol{v}_j$ the eigenvectors of $\boldsymbol{K}_0$. For the special case, $\boldsymbol{K}_0 = \boldsymbol{0}$, any set of orthonormal vectors satisfying the pairing condition may be used. (A rotation coordinate direction is a special case given by a pattern in which all m_j are zero except for a single element, which has the value of unity, and by a particular choice of vectors $\boldsymbol{V}$. As noted previously, in these cases E^{Exact} and E^{TIF} both display $t = \pi$ periodicity.)

For simplicity, only sets $\{m\}$ for which the elements are either all even or all odd are considered. As discussed further below, this is more restrictive than necessary, but these sets of direction patterns are sufficiently flexible for the present demonstration purposes. According to the construction given above, $[\boldsymbol{K}_0, \boldsymbol{\Delta}] = \boldsymbol{0}$, thereby admitting the factorizations:

$$\exp(\boldsymbol{K}_0 + t\boldsymbol{\Delta}) = \exp(\boldsymbol{K}_0)\exp(t\boldsymbol{\Delta}) \tag{3.18}$$

and

$$\sinh(2\boldsymbol{K}_0 + 2t\boldsymbol{\Delta}) = \sinh(2\boldsymbol{K}_0)\cosh(2t\boldsymbol{\Delta}) + \cosh(2\boldsymbol{K}_0)\sinh(2t\boldsymbol{\Delta}). \tag{3.19}$$

Diagonalization of $\exp(t\boldsymbol{\Delta})$, $\cosh(2t\boldsymbol{\Delta})$, and $\sinh(2t\boldsymbol{\Delta})$ reveals the oscillatory behavior of these quantities as a function of t. For example, the matrix:

$$\exp(t\boldsymbol{\Delta}) = \boldsymbol{V}\exp(it\boldsymbol{m})\boldsymbol{V}^\dagger = \sum_j \exp(itm_j)\boldsymbol{v}_j\boldsymbol{v}_j^\dagger + \exp(-itm_j)\boldsymbol{v}_j^*\boldsymbol{v}_j^{*\dagger} \tag{3.20}$$

is oscillatory along the line with period $t = 2\pi$. At $t = 2n\pi$, $\exp(t\boldsymbol{\Delta})$ is a unit matrix, $+\boldsymbol{1}$. At $t = (2n+1)\pi$, the matrix $\exp(t\boldsymbol{\Delta})$ is diagonal of the form $\pm\boldsymbol{1}$, with the sign determined by whether the m_j values of the direction pattern are even or odd. Using the argument that E^{Exact} is independent of these orbital phases, it follows that E^{Exact} is oscillatory along the line $t\boldsymbol{\Delta}$ with period $t = \pi$. Similarly, diagonalization of $\sinh(2t\boldsymbol{\Delta})$ and $\cosh(2t\boldsymbol{\Delta})$ in Eq. (3.19) reveals that $F(\boldsymbol{K})$ is oscillatory with period $t = \pi$. It then follows from this oscillatory nature of the coordinates, that E^{TIF} is oscillatory with period π. In this sense, E^{Exact} and E^{TIF} display the same periodic behavior along lines parameterized by $t\boldsymbol{\Delta}$ of the form of Eq. (3.19). It is mentioned in passing that since odd N_{orb} bases must have a zero $\boldsymbol{\Delta}$ eigenvalue, $t = \pi$ periodicity is displayed generally only by even-$\{m\}$ direction patterns in this case. Several examples of this behavior are demonstrated in Sect. 5.

If a direction pattern $\{m\}$ contains both mixed even and odd integers, then the resulting E^{Exact} generally is oscillatory only with period $t = 2\pi$. (There may also exist mixed direction patterns $\{m\}$ for which $\exp(t\boldsymbol{\Delta})$ is diagonal with mixed ± 1 diagonal elements occurring at regular intervals $t = n\pi$; this is the case with the coordinate grid patterns discussed previously. Other exceptions involving redundant orbital rotations are discussed in the following section.) For the period $t = 2\pi$ cases however, since $\sinh(2t\boldsymbol{\Delta})$ and $\cosh(2t\boldsymbol{\Delta})$ both have period $t = \pi$, the resulting E^{TIF} displays incorrect $t = \pi$ periodicity. The correction of this behavior within the interpolating function E^{TIF} does not appear simple. However, it should be noted that even in the worst situation, this causes a step to be somewhat shorter than optimal only when large steps are being taken. Since these large steps are to the least-accurate regions of the approximate E^{TIF} energy surface, it remains to be seen the extent to which this behavior adversely affects the convergence.

4. Redundant orbital rotations

It is the usual practice to remove explicitly the redundant orbital rotation variables from the wave function variational space. In particular, with an appropriate ordering of the orbitals, the orbital rotation matrix $\boldsymbol{K}$ takes a "block-off-diagonal form" [28], and the vector $\boldsymbol{\kappa}$ contains only the unique nonzero entries of $\boldsymbol{K}$. The "diagonal-blocks" correspond to invariant orbital

subspaces [13, 15]. To examine the effects of redundant orbital rotations on the approximating function E^{TIF}, suppose that, for some reason, this explicit elimination has not been done for at least one orbital rotation variable.

In Ref. [7] it is shown that, when the CSF gradient vector is zero, the gradient components corresponding to redundant orbital rotations vanish, and that all matrix elements of the partitioned orbital hessian matrix corresponding to pairs of redundant rotations vanish.

$$w_{(rs)} = 0 \quad ;(rs)\ redundant \tag{4.1}$$

$$(\boldsymbol{B} - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)_{(pq),(rs)} = 0 \quad ;(rs)\ redundant,\ (pq)\ redundant. \tag{4.2}$$

This implies that the contributions to the approximate E^{TIF} of the form $(w_{(rs)}f_{(rs)})$ and $(f_{(pq)}(\boldsymbol{B} - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)_{(pq),(rs)}f_{(rs)})$ vanish just as they do for E^{Exact}. It is also shown in Ref. [7] that for off-diagonal elements of the partitioned orbital hessian matrix the relation:

$$(\boldsymbol{B} - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)_{(xy),(rs)} = w_{ry}\delta_{sx} + w_{sx}\delta_{ry} + w_{xr}\delta_{ys} + w_{ys}\delta_{xr} \tag{4.3}$$

holds for (rs) redundant and (xy) arbitrary. For the coordinate grid patterns discussed previously (i.e. $x \neq y \neq r \neq s$), these off-diagonal elements are also zero. The contributions to E^{TIF} of the form $(f_{(xy)}(\boldsymbol{B} - \boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{C}^T)_{(xy),(rs)}f_{(rs)})$ vanish just as they do for E^{Exact}. That is, not only does E^{TIF} display the correct oscillatory behavior, but it also displays the correct invariance to the redundant rotation coordinate value within the mutually-commuting, cartesian-product subspace.

These relations demonstrate the conceptual advantage of formulating E^{TIF} in terms of the partitioned orbital hessian matrix instead of the wave function hessian matrix. There are no similar relations to Eqs. (4.2) and (4.3) on the matrix elements of $\boldsymbol{B}$, $\boldsymbol{C}$, or $\boldsymbol{M}$ taken individually; it is due to a subtle cancellation of contributions within individual matrix elements that these relations are satisfied.

As a final example of the effects of redundant orbital rotation operators, assume now that the wave function possesses some invariant orbital subspace. Let an arbitrary element of the redundant orbital rotation subgroup be denoted by $\boldsymbol{V}^{[R]}$. Further assume that E^{Exact} and E^{TIF} are invariant to rotations among the orbitals within this subspace, $E^{Exact}(0) = E^{Exact}[\boldsymbol{V}^{[R]}] = E^{TIF}(0) = E^{TIF}[\boldsymbol{V}^{[R]}]$. Suppose there exists some matrix $\boldsymbol{K}$ involving essential (i.e. nonredundant) rotations such that $E^{Exact}(0) = E^{Exact}[\exp(\boldsymbol{K})] = E^{TIF}(0) = E^{TIF}[\exp(\boldsymbol{K})]$ are also satisfied. Such a matrix $\boldsymbol{K}$ could be a knot point on a cartesian-product subspace grid, or at a $t = n\pi$ point corresponding to some direction pattern $\{m\}$, of the previous sections. The important characteristic of $\boldsymbol{K}$ for the present argument is that the energy invariance results from the general properties of wave function invariance, such as discussed in the previous sections, and not on some accidental numerical relationship. It is now asserted that the relations $E^{Exact}(0) = E^{Exact}[\boldsymbol{V}^{[R]}\exp(\boldsymbol{K})\boldsymbol{V}^{[R]\dagger}] = E^{TIF}(0) = E^{TIF}[\boldsymbol{V}^{[R]}\exp(\boldsymbol{K})\boldsymbol{V}^{[R]\dagger}]$ also hold.

This assertion rests on the following interpretation of the above product. Let the first $\boldsymbol{V}^{[R]}$ define a passive rotation of the original coordinate system. In this context, a passive rotation is associated with an integral transformation and adjustment of the CSF coefficient vector, such that all quantities are expressed relative to this new coordinate system. Since $\boldsymbol{V}^{[R]}$ is an element of the redundant subgroup, the relations, $E^{Exact}(0) = E^{Exact}[\boldsymbol{V}^{[R]}] = E^{TIF}(0) = E^{TIF}[\boldsymbol{V}^{[R]}]$, still hold. The matrix $\exp(\boldsymbol{K})$ then defines a point within the *rotated* coordinate system. If

the invariance property is independent of the orbital basis, then the relations, $E^{Exact}(0) = E^{Exact}[\boldsymbol{V}^{[R]} \exp(\boldsymbol{K})]$ and $E^{TIF}(0) = E^{TIF}[\boldsymbol{V}^{[R]} \exp(\boldsymbol{K})]$, will still hold in this new coordinate system, just as they did in the original system. The last $\boldsymbol{V}^{[R]\dagger}$ then rotates passively the coordinate system back to the original orientation, again leaving the energy unchanged. In this picture, the essential rotations are associated with coordinate axes, whereas the redundant rotations $\boldsymbol{V}^{[R]}$ are associated with rotations of these axes about the origin. The final result is that E^{Exact} and E^{TIF} are invariant to the matrix product, $(\boldsymbol{V}^{[R]} \exp(\boldsymbol{K}) \boldsymbol{V}^{[R]\dagger})$, and this relation must also hold if the matrix product is considered in the active sense. Therefore, any rotation matrix which is *similar* to another rotation matrix which is known to display energy invariance by one of the previous arguments, and for which the transformation matrix $\boldsymbol{V}^{[R]}$ belongs to the redundant orbital rotation group, will also display energy invariance.

Since $\boldsymbol{V}^{[R]}$ is chosen independently of $\boldsymbol{K}$, the behavior of the energy does not depend on the matrix elements of $\boldsymbol{V}^{[R]}$. Different choices for $\boldsymbol{V}^{[R]}$, would result in different rotations of the point $\exp(\boldsymbol{K})$ to the new point $(\boldsymbol{V}^{[R]} \exp(\boldsymbol{K}) \boldsymbol{V}^{[R]\dagger})$, all of which satisfy $E(0) = E^{exact}[\exp(\boldsymbol{K})] = E^{exact}[\boldsymbol{V}^{[R]} \exp(\boldsymbol{K}) \boldsymbol{V}^{[R]\dagger}] = E^{TIF}[\exp(\boldsymbol{K})] = E^{TIF}[\boldsymbol{V}^{[R]} \exp(\boldsymbol{K}) \boldsymbol{V}^{[R]\dagger}]$. This structure results in a denumerably infinite sequence of multidimensional nested balls, on the surface of which the reference energy is reproduced. If the matrices $\exp(\boldsymbol{K})$ and $\boldsymbol{V}^{[R]}$ commute, then the energy-equivalence relation is trivially satisfied, in that the representation of $\exp(\boldsymbol{K})$ in both the original and rotated systems is the same. Since the representation of the point $\boldsymbol{K}$ is the same in both coordinate systems, the rotation $\boldsymbol{V}^{[R]}$ does not induce a nested ball structure. Therefore, the dimensionality of the surface of the balls is determined by the number of redundant orbital rotation parameters in the wave function which do not commute with the contributing rotation coordinate matrices in the matrix $\boldsymbol{K}$.

In a sense, this oscillatory behavior is independent from that which results from the direction patterns $\{m\}$ of the previous section, being more intimately connected with the wave function expansion form rather than simply the orbital transformation matrix. Indeed, direction patterns $\{m\}$ with mixed even-odd integers do not normally result in $t = \pi$ periodicity (displaying instead only period $t = 2\pi$ behavior). However, in the presence of redundant orbital rotations, it is possible for the "offending" terms in the even-odd set to be flattened, leaving only the coherent, period $t = \pi$ terms remaining to contribute to the energy. An example of this behavior is given in the next section.

This discussion is concluded by pointing out that this nested ball structure is shared by both E^{Exact} and E^{TIF}. These common features lend further support to the argument that E^{TIF} is a desirable approximating function to E^{Exact} with which to interpolate between the surfaces of the nested balls or between the isolated points for which $E^{TIF} = E^{Exact}$. Future work will be directed towards a deeper understanding of the nature of this interesting structure of MCSCF wave functions.

5. Numerical examples

Some of the features of coordinate transformations that have been discussed in the previous sections will now be demonstrated. For this purpose, it is convenient to define a model hamiltonian with which several model MCSCF optimization problems will be constructed. For simplicity, a three-orbital basis is chosen.

Table 1. The model hamiltonian integrals[a]

$(r, s) =$	(1, 1)	(2, 1)	(2, 2)	(3, 1)	(3, 2)	(3, 3)
$h_{rs} =$	−2.00	0.10	−1.75	0.10	0.10	−1.50
$g_{11rs} =$	0.80	—	—	—	—	—
$g_{21rs} =$	−0.18	0.27	—	—	—	—
$g_{22rs} =$	0.52	−0.17	0.70	—	—	—
$g_{31rs} =$	−0.16	0.03	−0.12	0.26	—	—
$g_{32rs} =$	−0.11	0.02	−0.14	0.01	0.25	—
$g_{33rs} =$	0.50	−0.10	0.51	−0.15	−0.13	0.60

[a] Only the canonical integrals are specified. Units are E_h

This has the advantage of allowing sufficient complexity to realize some of the more subtle points discussed above, while avoiding some of the conceptualization difficulties associated with higher dimensionality. Table 1 displays the 1-*e* hamiltonian and 2-*e* repulsion integrals for this model hamiltonian. With these integrals, all of the results of this section may be reproduced and examined in more detail by the reader. It should perhaps be mentioned in passing that these integrals were originally taken from a CH^+ cation calculation with four electrons frozen in the core and with all but the lowest three remaining sigma orbitals deleted. The integrals from this calculation were "simplified" to the final form given in Table 1. Consequently, these results should model typical electronic structure calculations.

It is first demonstrated that E^{Exact} is oscillatory along the rotation coordinate directions for a general MCSCF wave function. For this purpose, a two-CSF singlet expansion space, defined as $\{|300\rangle, |030\rangle\}$ using step-vector notation [7, 20], is employed. There are no redundant orbital rotations with this expansion space, and consequently the energy depends on all three essential orbital rotation coordinates, e^{21}, e^{31}, and e^{32}. In Fig. 1 the ground-state MCSCF energy is shown as a function of t along each of these three coordinates. It may be seen in particular that the energy is oscillatory with period $t = \pi$ along all three

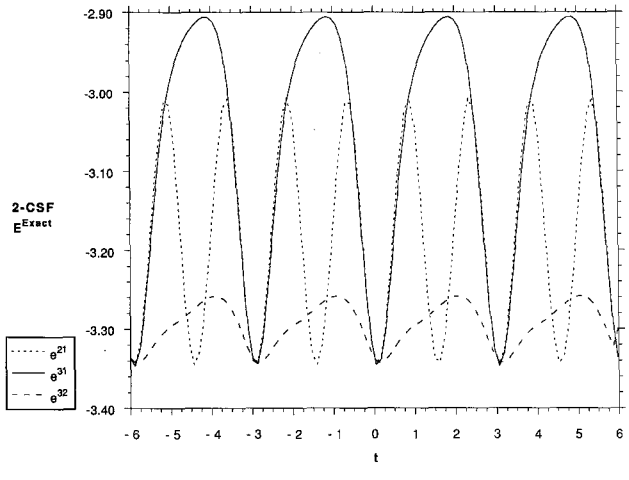


Fig. 1. 2-CSF E^{Exact} along the coordinate axes e^{21}, e^{31}, and e^{32}

coordinates. In fact, in the te^{21} direction, the energy is actually oscillatory with period $t = \pi/2$. This is a consequence of the symmetric appearance of orbitals 1 and 2 in the CSF expansion space; at every $t = \pi/2$ increment, the two expansion CSFs essentially swap roles. According to the previous analysis for this three-orbital problem, there are no more cartesian grids which are periodic, since $N^{[Grid]}(3) = 3$, and there are no higher-dimensionality subspaces, such as planes, which are periodic, since $INT[3/2] = 1$.

Next, consider E^{Exact} along an arbitrary line within the plane defined by the e^{31} and e^{32} coordinates. The line $t\Delta$ may be parameterized by an angle θ as $\Delta = \cos(\theta)e^{31} + \sin(\theta)e^{32}$. With this parameterization, it may be verified that the eigenvalues of Δ are $(\pm i, 0)$ and are independent of θ. The plot in Fig. 2 is for $\theta = -\pi/5$. This plot shows clearly that E^{Exact} is only oscillatory with period $t = 2\pi$ for this case. As discussed in Sect. 3, the energy along a line with a mixed even-odd direction pattern, $\{1, 0\}$ in this case, is not expected to display $t = \pi$ periodicity except in special cases. For this model problem, it may be verified that $\theta = \pm\pi/4$ are such special cases and do, in fact, result in $t = \pi$ periodicity (again a consequence of the role-swapping of the two expansion CSFs). However even in the general case, it would appear quite reasonable to *approximate* the curve in Fig. 2 with a period $t = \pi$ oscillatory function, since the exact function is "almost" $t = \pi$ periodic; this approximation is exactly what is done by the E^{TIF} interpolating function.

In order to test the periodicity arguments given in Section 3, the behavior of E^{Exact} along lines $t\Delta$ parameterized by $\Delta = 2(\cos(\theta)e^{31} + \sin(\theta)e^{32})$ is now examined. With this scaling, Δ has eigenvalues $(\pm 2i, 0)$, corresponding to the all-even direction pattern of $\{2, 0\}$. Figure 3 shows plots for $\theta = -\pi/5$. With this scaling of the coordinates (compared to Fig. 2), the curve is now oscillatory with period $t = \pi$.

The effects of redundant variables are considered next. With the chosen model hamiltonian, this may be achieved by expanding to the 3-CSF expansion space $\{|300\rangle, |120\rangle, |030\rangle\}$. This is, of course, a 2-orbital/2-electron full-CI expansion space, so the rotation e^{21} is now redundant. Figure 4 is analogous to Fig. 1 for the 2-CSF expansion space, showing the ground state energy, E^{Exact},

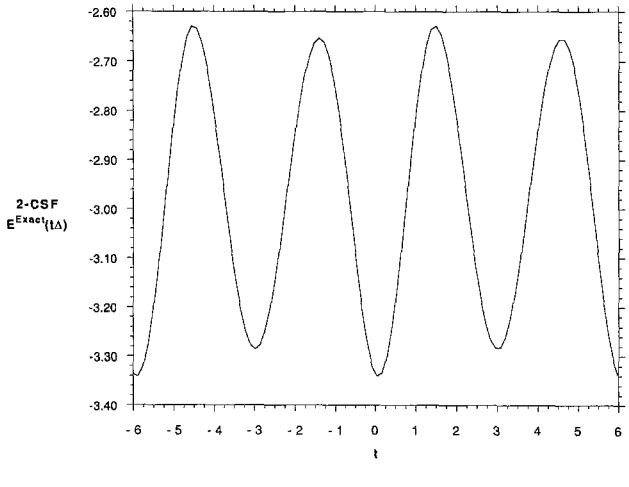


Fig. 2. 2-CSF $E^{Exact}(t\Delta)$ for $\Delta = (\cos(\theta)e^{31} + \sin(\theta)e^{32})$ with $\theta = -\pi/5$. This corresponds to the mixed even-odd direction pattern $\{1, 0\}$, and therefore has period $t = 2\pi$ instead of period $t = \pi$

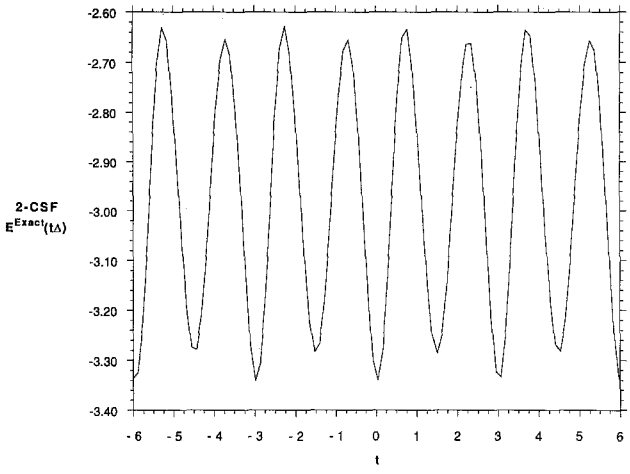


Fig. 3. 2-CSF $E^{Exact}(t\Delta)$ for $\Delta = 2(\cos(\theta)e^{31} + \sin(\theta)e^{32})$ with $\theta = -\pi/5$. This corresponds to the all-even direction pattern $\{2, 0\}$, and therefore has period $t = \pi$

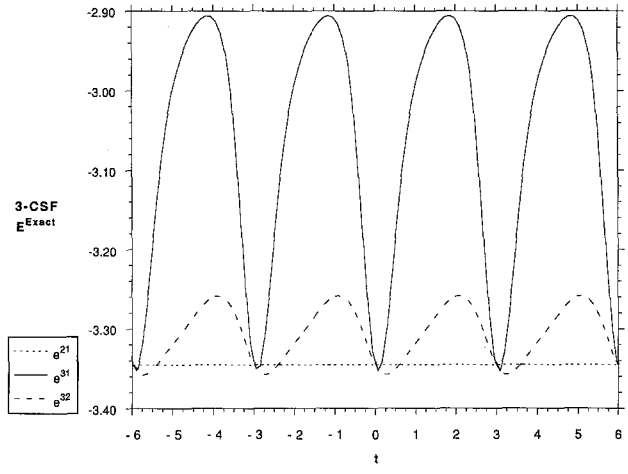


Fig. 4. 3-CSF E^{Exact} along the coordinate axes e^{21}, e^{31}, and e^{32}

along each of the three orbital rotation coordinates. As expected, $E^{Exact}(te^{21}) = E^{Exact}(0)$ for all values of t. The energy dependence along the other two coordinate directions is unchanged qualitatively compared to the 2-CSF case, although quantitatively all of the energy values are now lower due to the bracketing theorem [7].

Figure 5 displays the 3-CSF $E^{Exact}(t\Delta)$ along the lines $\Delta = \cos(\theta)e^{31} + \sin(\theta)e^{32}$ for $\theta = \pm\pi/5, \pm 2\pi/5$. The $\theta = -\pi/5$ curve in particular may be compared directly to the 2-CSF curve in Fig. 2. It is seen that the introduction of the redundant orbital rotation coordinate e^{21} causes the energy to display $t = \pi$ periodicity in all four curves, even for the mixed even-odd direction pattern $\{1, 0\}$. There are two ways to interpret this feature. The first is to note that any matrix of the form

$$\Delta = \begin{pmatrix} 0 & 0 & -k_{31} \\ 0 & 0 & -k_{32} \\ k_{31} & k_{32} & 0 \end{pmatrix},$$

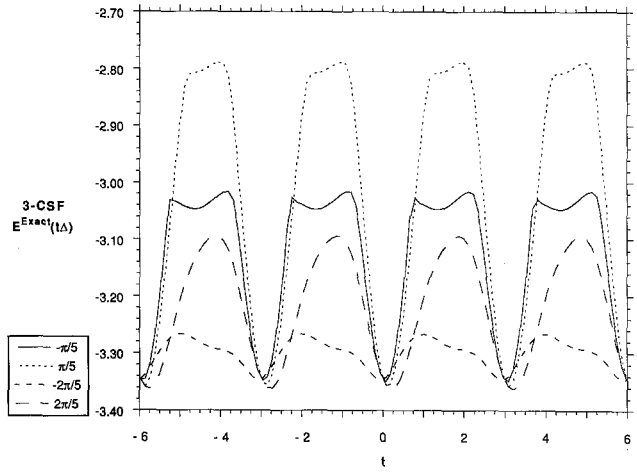


Fig. 5. 3-CSF $E^{Exact}(t\Delta)$ at $\theta = \pm\pi/5, \pm 2\pi/5$

is similar to a matrix of the form

$$\begin{pmatrix} 0 & 0 & -|k| \\ 0 & 0 & 0 \\ |k| & 0 & 0 \end{pmatrix},$$

using a plane Givens rotation in the $e^1 \otimes e^2$ plane. In this transformed basis, the representation of Δ obviously leads to periodic behavior, and since the Givens rotation is within the redundant orbital rotation group, the original representation of Δ must also be oscillatory with period $t = \pi$. This argument generalizes quite readily to arbitrary matrices of the form $\Delta = (e^n x^T - x(e^n)^T)$ for arbitrary dimensions N_{orb}. With $|x| = 1$, $V = (1 - (2/y^T y)yy^T)(1 - 2e^1(e^1)^T)$ and $y = x + |x|e^1$, Δ is seen to be similar to the matrix e^{n1}, leading trivially to an oscillatory energy with period $t = \pi$. As long as x contains nonzero elements only in rows corresponding to the invariant orbital subspace, and n corresponds to an orbital that is not in this set, then the given product of Householder matrices is an element of the redundant orbital rotation subgroup.

The other interpretation is to *start* with the matrix

$$\Delta^{31} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

which trivially leads to oscillatory energy with period $t = \pi$, and then to rotate a point $n\pi\Delta^{31}$ about the origin with an arbitrary element of the (one-parameter) redundant orbital rotation subgroup parameterized by the angle $-\theta$. This subgroup matrix may be written as $V^{[R]} = \exp(-\theta e^{21})$. The desired result, $V^{[R]}\Delta^{31}V^{[R]\dagger} = \Delta$, follows immediately. In this simple example, the close relation between these two points of view is almost trivial; even the same algebraic manipulations are used in both cases, although in slightly different orders. In this second interpretation, it is clear how, as a function of θ, a series of points $n\pi\Delta^{31}$ are mapped into a series of nested circles. The dimensionality of the surfaces of the nested balls is determined by the number of redundant orbital rotation operators (i.e. 1 in this example).

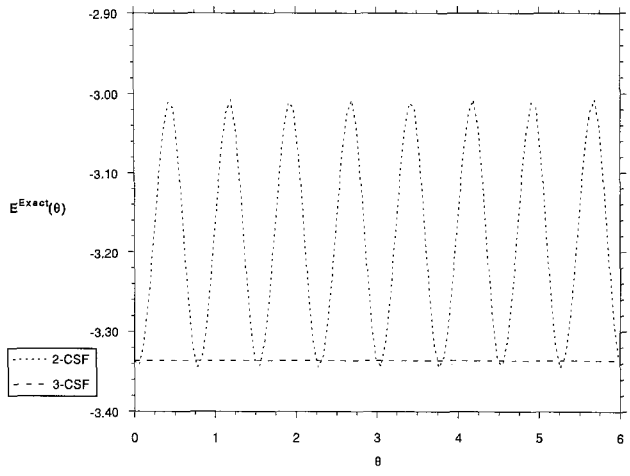
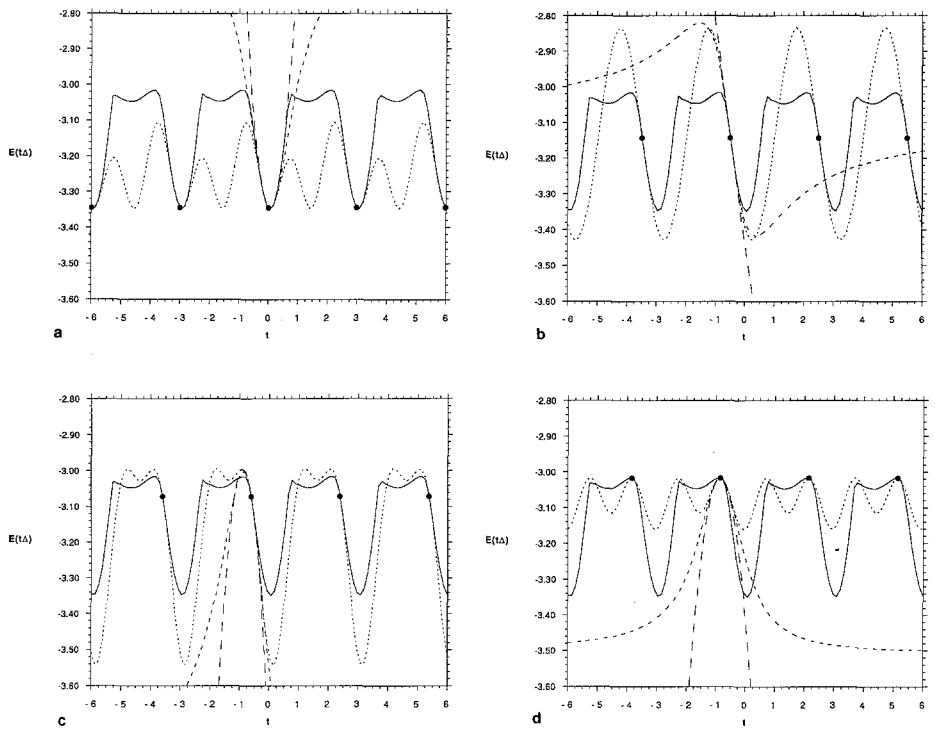


Fig. 6. Comparison of 2-CSF E^{Exact} and 3-CSF E^{Exact} on a circle of radius π

A more direct verification of this effect may be seen by plotting the energy on a circle of radius π within the $e^{31} \otimes e^{32}$-plane. In Fig. 6, the 2-CSF and 3-CSF energies are displayed on the curve $t\Delta$ with $\Delta = \cos(\theta)e^{31} + \sin(\theta)e^{32}$ at fixed $t = \pi$ as a function of θ. The 3-CSF energy is invariant on this curve, whereas the 2-CSF energy is oscillatory. This is a clear demonstration of the nested ball structure of the 3-CSF MCSCF energy in the presence of a redundant orbital rotation coordinate; this same constant-energy plot would be obtained at $t = 2\pi, 3\pi, \ldots n\pi \ldots$

The 3-CSF functions E^{Exact} and E^{TIF} along the line $t\Delta$ with $\Delta = \cos(\theta)e^{31} + \sin(\theta)e^{32}$ for $\theta = -\pi/5$ will now be examined in detail. This particular angle is chosen because, from among the curves in Fig. 5, it has the most complicated structure to be approximated. The solid line in Figs. 7a–7d is the same as the $\theta = -\pi/5$ curve in Fig. 5, and the broken lines are some of the local energy approximations discussed in Section 2. $E^{[2]}(t\Delta)$ is a simple parabola which always reproduces the energy, gradient, and hessian at the expansion point. In these cases, the rational function approximation is given by $E^{PSCI}(t\Delta) = E(0) + (E^{[2]}(t\Delta) - E(0))/(1 + t^2)$. This function also reproduces the energy, gradient, and hessian at the expansion point, but displays quite different $t \to \pm\infty$ behavior from $E^{[2]}$. It is immediately clear from these graphs that the "local" approximating functions, $E^{[2]}$ and E^{PSCI}, are indeed only accurate within a rather small neighborhood of the expansion point, whereas E^{TIF} displays periodic behavior, and thereby reproduces correctly the energy, gradient, and hessian not only at the expansion point, but also at an infinite number of points displaced by $n\pi$ from the expansion point.

In Fig. 7a, the expansion point, $t_0 = 0$, is close to the minimum of the surface. It may be observed that when the reference point is near any of the (equivalent) minima on the curve, then E^{TIF} predicts quite accurately the behavior of the energy in the regions near the minima. At this expansion point, all three approximating functions, E^{TIF}, $E^{[2]}$, and E^{PSCI}, describe well the position of the nearest minimum. As discussed previously, as the gradient becomes small, the higher-frequency, period $\pi/2$ contributions from the hessian are overestimated in the E^{TIF} curve. This results in the shallow, high-energy minimum, displaced



Figs. 7a–d. Comparison of 3-CSF —— E^{Exact}, ----- E^{TIF}, — — $E^{[2]}$, and - - - E^{PSCI} along $t\Delta$ for $\Delta = (\cos(\theta)e^{31} + \sin(\theta)e^{32})$ with $\theta = -\pi/5$ for various expansion points. The expansion points (● t_0) are 0.0, −0.5, −0.625, and −0.84, respectively

about $\pi/2$ from the minimum in E^{Exact}, being predicted to have approximately the same energy as the minimum.

In Fig. 7b, the expansion point, $t_0 = -0.5$, is somewhat further away and is about midway between the energy minimum and the maximum. The hessian is small and positive at this expansion point. The E^{TIF} behavior is dominated by the period π gradient contribution, this time underestimating the period $\pi/2$ behavior from the hessian and missing entirely the high-energy structure of the E^{Exact} curve. The position of the minimum is predicted fairly well by E^{TIF}. $E^{[2]}$ clearly does not predict the behavior near the minimum; this expansion point is outside of the radius of convergence for the straightforward Newton–Raphson iterative procedure. At this expansion point, E^{PSCI} still approximates the position of the nearest minimum reasonably well.

In Fig. 7c, the expansion point, $t_0 = -0.625$, is slightly further still from the minimum. The hessian is small and negative at this expansion point. Both the high-energy structure and the position of the minimum are approximated reasonably well by E^{TIF}. $E^{[2]}$ cannot predict the position of the minimum, but it does predict accurately the position of the nearest maximum. E^{PSCI} also predicts accurately the position of the maximum, but the predicted minimum, at approximately 2.97 is nowhere near the position of the nearest minimum. Empirical trust-radius constraints would be required in order to converge reliably to the nearest minimum using either local-approximation method.

In Fig. 7d, the expansion point, $t_0 = -0.84$, is close to the energy maximum. Again, $E^{[2]}$ has no minimum and the E^{PSCI} minimum (at approximately 20.36) does not predict even qualitatively the real position of the nearest minimum. The TIF approximates not only the neighborhood of the maximum, but also the positions of both nearby minima. Since the gradient contribution is relatively small at this expansion point, the period $\pi/2$ hessian contribution is again overestimated. If the expansion point had been chosen exactly at the maximum, then pure period $\pi/2$ behavior would have been predicted, and the predicted energies at the two minima would have been equal. It may be inferred from this property of E^{TIF} that only the positions of the nearest minima are expected to be predicted accurately, and that the relative energies of the nonequivalent minima are not quantitatively accurate. It remains to be determined what information about the other predicted minima can be extracted from E^{TIF}.

It is encouraging that as the reference point is moved away from the minimum, the position of the approximate E^{TIF} minimum is usually very close to the actual minimum, although the computed approximate energy value at this minimum is not always quantitatively accurate. For the purpose of optimization, it is of course the location of the minimum that is most important, and not the quantitative prediction of the energy value at that point.

A more complete view of the 3-CSF E^{Exact} surface is given with the contour plot in Fig. 8. For reasons to be discussed below, the plot coordinates have been chosen relative to the point $(k_{31}^0, k_{32}^0) = (-1.5, 0.0)$. Specifically, the value at a coordinate point (k_{31}, k_{32}) corresponds to E^{Exact} defined with the transformation $\boldsymbol{U}^{Total}$:

$$\begin{aligned}\boldsymbol{U}^{Total}(k_{31}, k_{32}) &= \exp(-1.5\boldsymbol{e}^{31}) \exp(k_{31}\boldsymbol{e}^{31} + k_{32}\boldsymbol{e}^{32}) \\ &= \exp\begin{pmatrix} 0 & 0 & 1.5 \\ 0 & 0 & 0 \\ -1.5 & 0 & 0 \end{pmatrix} \exp\begin{pmatrix} 0 & 0 & -k_{31} \\ 0 & 0 & -k_{32} \\ k_{31} & k_{32} & 0 \end{pmatrix}\end{aligned} \tag{5.1}$$

relative to the basis integrals of Table 1. (In reproducing the results of this section, the reader should note in particular that since $[\boldsymbol{e}^{31}, \boldsymbol{e}^{32}] \neq 0$, then $\boldsymbol{U}^{Total}(k_{31}, k_{32}) \neq \exp((k_{31} - 1.5)\boldsymbol{e}^{31} + k_{32}\boldsymbol{e}^{32})$.) As a consequence of this shift, the origin of Fig. 8 is slightly on the downhill side of the maximum on the surface, and is approximately on the steepest-descent path toward the saddle point.

The oscillatory behavior along lines $t\boldsymbol{\Delta}$ may be observed directly from the contour plot of Fig. 8 as follows. From any point (k_{31}, k_{32}) draw a line through the origin. All points (k'_{31}, k'_{32}) on this line that are a distance of $n\pi$ from (k_{31}, k_{32}) will have the same energy. Using this geometrical model, a set of closely spaced points near the origin is seen to map onto a set of widely spaced points near the circle of radius π centered at the origin – a lever-arm effect which results in the overall general landscape of the contour plot. The origin point itself, maps exactly onto the circles at radius $n\pi$, providing the geometrical representation of the nested ball structure.

From this geometrical representation, it may be verified that the two minima of the surface at approximately $(-1.5, -0.3)$ and $(1.5, 0.3)$ are actually two equivalent representations of the same wave function. Similarly, the three maxima at approximately $(0.4, -0.5)$, $(-1.5, 1.9)$, and $(2.4, -3.0)$ are actually

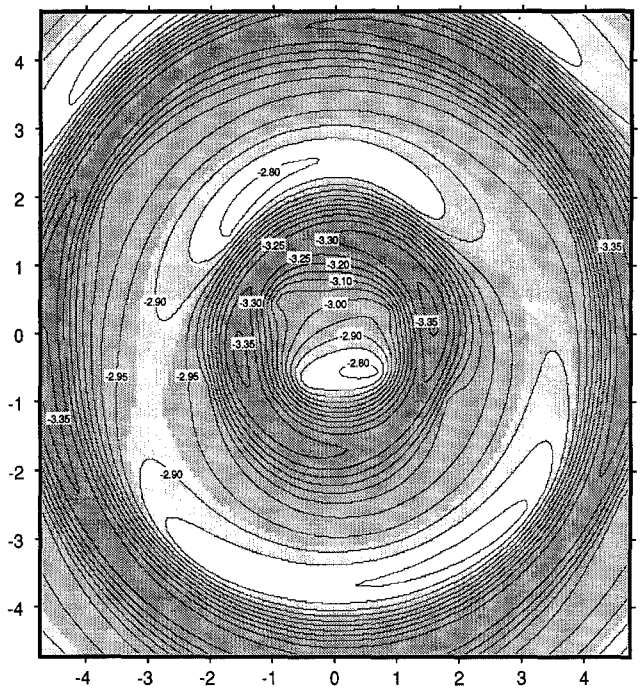


Fig. 8. Contour plot of 3-CSF E^{Exact} relative to the intermediate basis, $\exp(-1.5e^{31})$. The horizontal and vertical axes are the values of k_{31} and k_{32}, respectively

three equivalent representations of the same wave function. It is this global oscillatory behavior that the TIF-based MCSCF optimization presented in this work attempts to exploit.

An alternative view of the 3-CSF E^{Exact} is given in Fig. 9. Such a graph is called a "Random Dot Stereogram" [29, 30]. To view this image, the reader should place the graph on a flat surface at a comfortable reading distance. Focus past the actual surface, until the two guide dots at the bottom of the graph are seen as four dots (i.e. two dots for each eye). Adjust the focus point until the two "middle" dots merge together and three dots are perceived. When the three-dot pattern is stable, the reader may slowly look up to the graph and perceive the full three-dimensional character of the surface. With a few minutes of practice, the typical viewer can dispense with using the guide dots and can focus almost instantly on the three-dimensional image itself. It should also be mentioned that the larger the graph, the more dramatic is the effect. Enlarged copies of the graph may be produced quite easily with an office copy machine, and the integrity of the image is reduced insignificantly with repeated enlargements.

As plotted in Figs. 8 and 9, the constant-energy circle at radius π is seen to cut across the north face of the southeast peak, passing through a high saddle point around to the north face of the northwest peak, then passing through another image of the high saddle-point back to the north face of the bottom peak. Although the origin is not a saddle point, the two high-energy saddle points are actually images of the origin point. It may be verified that if $\boldsymbol{w}(\mathbf{0}) \neq \mathbf{0}$, then images of the origin at two points on each $n\pi$-circle will be saddle points. This is because the hessian at these two points in polar coordinates has the form $\begin{pmatrix} 0 & h_{t\theta} \\ h_{t\theta} & h_{tt} \end{pmatrix}$ resulting in a positive and a negative eigenvalue when transformed back to the k_{31}–k_{32} coordinate system. The two equivalent minima to the southwest

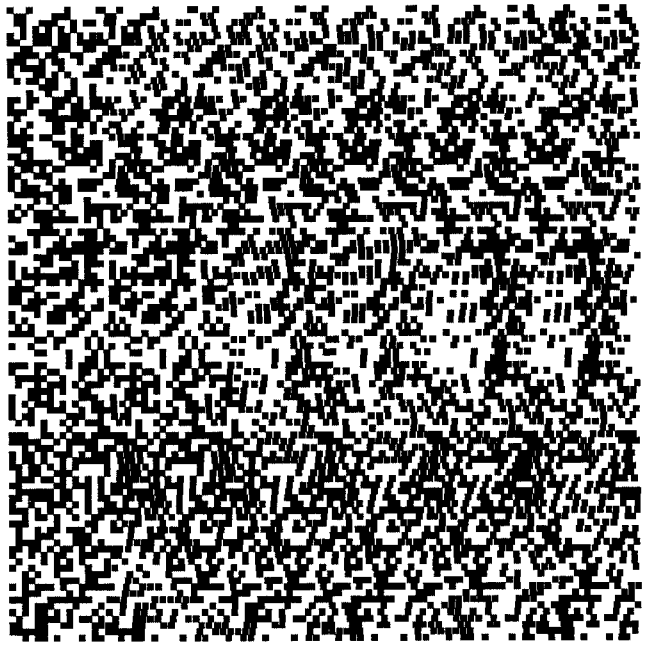


Fig. 9. Random Dot Stereogram of 3-CSF E^{Exact}. Compare to the contour plot of this same surface shown in Fig. 8

and northeast are connected with two equivalent low-energy saddle points to the northwest and southeast, approximately at the coordinates $(-0.9, 1.1)$ and $(1.0, -1.3)$. The energies and coordinates of these surface features are given in Table 2. These surface features will be used for reference as the problem of MCSCF optimization on this surface is considered.

6. MCSCF optimization

Optimization of the MCSCF wave function using the method of trigonometric interpolation proceeds according to the following steps.

DO

(1) Compute $E(0)$, $\boldsymbol{w}$, and optionally $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{M}$ for the current orbitals, defined by $\boldsymbol{U}$. This involves the partial integral transformation, solution of the hamiltonian eigenvector, density-matrix construction, and optionally, the construction of the matrices $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{M}$.

(2) Optimize $\boldsymbol{K}$ on E^{TIF} as defined in Eq. (3.9). This is a nonlinear optimization step, and may itself involve (micro-)iterative procedures of various kinds. $\boldsymbol{K}$ is computed relative to the current orbitals.

(3) Check for convergence, and EXIT if appropriate.

(4) Update $\boldsymbol{U}^{new} = \boldsymbol{U}^{old} \exp(\boldsymbol{K})$.

ENDDO

Table 2. E^{Exact} surface features[a]

Description	Energy/E_h	k_{31}	k_{32}
Minima	−3.362113258981	4.6036283989	0.9895685549
		1.5321927067	0.3293510230
		−1.5392429855	−0.3308665089
		−4.6106786777	−0.9910840409
Low-energy saddle points	−3.283000398520	−2.8617781006	3.6234151099
		−0.9146105618	1.1580260987
		1.0325569770	−1.3073629124
		2.9797245159	−3.7727519236
High-energy saddle points (axes origin)	−2.917251878752	−3.0119508693	−0.8931720751
		0.0	0.0
		3.0119508693	0.8931720751
Maxima	−2.788220893587	−3.5179019269	4.3073900388
		−1.5306681700	1.8741809650
		0.4565655868	−0.5590281087
		2.4437993437	−2.9922371825

[a] All coordinates are relative to the intermediate orbital basis $\exp(-1.5e^{31})$

There are, of course, various ways to optimize the nonlinear function E^{TIF} in step (2) of the above sequence. The most efficient are those which utilize not only the approximate energy, $E^{TIF}(\boldsymbol{K}_0)$, at various expansion points $\boldsymbol{K}_0$, but also the derivatives,

$$\left.\frac{\partial E^{TIF}}{\partial K_{\alpha\beta}}\right|_{\boldsymbol{K}=\boldsymbol{K}_0} \quad \text{and possibly} \quad \left.\frac{\partial^2 E^{TIF}}{\partial K_{\alpha\beta}\,\partial K_{\mu\nu}}\right|_{\boldsymbol{K}=\boldsymbol{K}_0}.$$

Of these, the most practical and useful methods are subspace methods which require only matrix-vector products of the matrices $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{M}$ with trial vectors. For small problems, these matrices may be explicitly constructed, and the required products computed directly with the stored matrices. For problems involving larger basis sets and/or larger CSF expansion spaces, these matrix-vector products may be computed using 1-index transformation and transition density matrix techniques. See Ref. [7] for further discussion of these details. As shown in Ref. [27], under these conditions, the construction of the matrix $F(\boldsymbol{K}_0)$, and the required first and second derivative contributions,

$$\left.\frac{\partial F(\boldsymbol{K})}{\partial K_{\alpha\beta}}\right|_{\boldsymbol{K}=\boldsymbol{K}_0} \quad \text{and} \quad \left.\frac{\partial^2 F(\boldsymbol{K})}{\partial K_{\alpha\beta}\,\partial K_{\mu\nu}}\right|_{\boldsymbol{K}=\boldsymbol{K}_0},$$

require only $O(N_{orb}^3)$ computational effort.

In the present work, it is mentioned only that a rational-function-based optimization method with dynamic level-shifting (a slight variation of the usual PSCI MCSF optimization procedure), is used herein to optimize E^{TIF}. To date, there have been no unexpected difficulties in any of these optimizations, but it is

Table 3. MCSCF optimization using trigonometric interpolation

(k_{31}, k_{32})[a]	(−1.50, −0.48)	(−1.45, −0.62)	(1.0, −1.0)	(0.0, 0.0)	(0.725, −0.75)
Hessian[b]	2.2888521138	2.284340525	1.8292524009	0.3189427581	−0.6113559802
Eigenvalues	0.0870895590	0.023280740	−0.4363317324	−0.1772591097	−2.8577918026
1	−3.3603960748	−3.3561962157	−3.2112067263	−2.9172518787	−2.8602918655
2	−3.3620151864	−3.3544485010	−3.3083150188	−3.0333225484	−3.2708583480
3	−3.3621131802	−3.3620602613	−3.3606462404	−3.2918719500	−3.3325673816
4	−3.3621132589_	−3.3621132557	−3.3621002529	−3.3291123418	−3.3615851643
5		−3.3621132589_	−3.3621132572	−3.3617290940	−3.3621119396
6			−3.3621132589_	−3.3621124779	−3.3621132589_
7				−3.3621132589_	

[a] The initial starting points are relative to the intermediate orbital basis $\exp(-1.5e^{31})$

[b] Hessian eigenvalues at the minimum are 2.2972896439 and 0.1491485099

anticipated that more efficient methods, particularly those which exploit the known nonlinear nature of E^{TIF} and which reduce the number of required matrix-vector products, will be developed in the near future.

In Table 3, examples of the overall convergence using the trigonometric interpolation method with various initial guesses are presented. The first row of Table 3 gives the coordinates of the initial starting point (again relative to the $\exp(-1.5e^{31})$ intermediate basis). The second two rows of Table 3 are the hessian eigenvalues at the initial starting point, relative to the intermediate orbital basis. In this basis, the eigenvalues of the hessian at the minimum are 2.2972896439 and 0.1491485099. (For completeness, the hessian eigenvalues are 2.3125549489 and 0.3672649727 in the converged orbital basis. This difference is due to a "fish-eye" type of distortion of the coordinate system as the origin is translated with the $e^{31} \otimes e^{32}$ plane.) The subsequent rows of Table 3 are the energies of each MCSCF iteration.

The first calculation in Table 3 is chosen to typify a situation for which reasonable starting orbitals are available, for example, from a nearby geometry on a potential energy surface. This starting point corresponds to a rotation of the minimum in the $e^{31} \otimes e^{32}$ plane about the origin by $\pi/32$, or about 5.6°. The hessian contains all positive eigenvalues, and the eigenvalues are not significantly different from those at the minimum. The straightforward Newton–Raphson procedure also converges for this starting point, and in the same number of iterations as the new trigonometric interpolation method. The energy residual in the fourth iteration is about $7.0 \cdot 10^{-14} E_h$, thereby verifying numerically the second-order convergence of the TIF-based MCSCF method.

The next column in Table 3 corresponds to a rotation of $\pi/16$, or about 11°. Although the initial hessian is positive definite, the straightforward Newton–Raphson procedure does not converge from this starting point. In fact, the first iteration of the TIF procedure oversteps with a maximal rotation of $\pi/4$, indicating that the period $\pi/2$ contributions from the hessian matrix are underestimated in E^{TIF}. An iteration could be saved by damping this step, but since one of the goals of the TIF optimization procedure is the elimination of such empirical adjustments, such damping was not employed in any of the calculations of Table 3. Future work will examine the

possibility of eliminating these spurious large steps in a well-founded, nonempirical, manner.

The third column of Table 3 corresponds to an initial point near the low-energy saddle point of the surface. The initial hessian matrix has a positive and a negative eigenvalue, characteristic of the saddle point region. The initial energy is about 95 kcal/mol higher than the minimum, a value which is outside of what might be considered "normal" starting orbitals for MCSCF optimization. Nonetheless, the TIF procedure succeeds in locating the quadratic region of the minimum.

The fourth column of Table 3 corresponds to the rather complicated region of the energy surface near the origin of Figs. 8 and 9. The hessian matrix in this region possesses mixed positive and negative eigenvalues, but it is not in the vicinity of a saddle point. (However, as mentioned previously, isolated *images* of the origin generally do correspond to saddle points.) Inspection of the second root of the hamiltonian in this region reveals that this complicated structure is due to a crossing of the lowest two roots within the $e^{31} \otimes e^{32}$ wave function variational space. The local minimum and the other high-energy features shown in the line plots in Figs. 5 and 7 are also due to this crossing. This crossing can be seen distinctly in Figs. 8 and 9, in which an otherwise smooth surface appears "sliced-off" in the region near the origin (and its images). Consequently, the origin represents a rather difficult starting point for optimization, since the gradient and hessian describe qualitatively this anomalous local surface patch rather than that associated with the low-energy surface regions. The first iteration succeeds in moving away from the maximum of the surface, but the E^{TIF} minimum erroneously lies almost on top of the actual saddle point. It is because of this feature that the origin of the plots of Figs. 8 and 9 are chosen as they are; this provides a stringent test of the method, and moves the interesting region of the surface to where these details are most readily observed. Maximal $\pi/4$ steps are taken during the first three iterations, and thereafter the reference point moves into the region of the minimum.

The last column of Table 3 corresponds to an initial point that is close to the maximum, but on the other side of the peak from the complicated root-crossing region discussed above. The initial hessian has two negative eigenvalues. Although the initial energy is higher than that of the previous calculation (over 300 kcal/mol above the minimum), this actually represents a somewhat easier optimization problem. The first step is again a maximal $\pi/4$ step, picking up 250 kcal/mol at once, and then the subsequent steps wind around the curved path into the quadratic region of the minimum.

Considering some of the complications of this model surface, the TIF optimization method does quite well at optimizing the energy from various starting points, particularly within low-energy regions that are connected simply to the neighborhood of the minimum. When not in the low-energy region, the best behavior appears to be observed when the initial point possesses comparable contributions from both the gradient and hessian terms. If the hessian contributions are relatively small, then maximal $\pi/4$ steps are taken, and although these are often fruitful, they sometimes lead to overstepping or oscillatory behavior. If the gradient contributions are small relative to the hessian terms, then the high-frequency period $\pi/2$ contributions from the hessian appear to be overestimated within E^{TIF}. This is observed in particular if the initial guess is exactly at one of the saddle points or at a maxima. In all cases examined so far, the effects

on the MCSCF convergence of this undesired behavior can be eliminated with the imposition of stepsize restrictions, indicating that this does not appear to be an insurmountable problem inherent with the interpolation procedure. It is hoped, however, that these remaining problems can be addressed using nonempirical approaches.

7. Discussion and conclusions

As mentioned in the previous sections, there are several features of the new MCSCF optimization method which should be examined more fully. These include the development of more efficient optimization schemes for the E^{TIF} optimization within an MCSCF iteration. Another issue to be addressed is the "spurious" nature of the $\pi/2$ periodicity of E^{TIF} when the gradient is small at the expansion point. One possible approach involves the incorporation of information from previous MCSCF iterations into the current E^{TIF}. For example, by generalizing the functional forms of E^{TIF} and/or $F(\boldsymbol{K})$, it is possible to reproduce exactly energies and gradient vectors determined from previous iterations at the appropriate nonexpansion points. Another interesting possibility would be to solve for the CSF correction vector $\boldsymbol{p}$ to infinite-order for fixed-order $F(\boldsymbol{K})$, and then to update the densities directly from this relaxed wave function correction. Such schemes have been proposed for the other MCSCF methods (for a more complete discussion, see Ref. [7] and references therein) as a way to reduce the number of expensive repulsion-integral transformation steps, particularly for large orbital-basis-function sets.

MCSCF methods that have been examined in the past and that proved ineffective in the context of local expansions, might now be reexamined in the context of the global interpolating function approach. The iterative incorporation of higher-order derivatives at the expansion point is one such likely candidate for reexamination.

The trigonometric interpolation method has been presented primarily in the context of energy minimization. As mentioned in the previous sections, such interpolation methods may also be useful in following other stationary solutions, e.g. in a region of geometry space at which multiple MCSCF solutions exist. Finally, although the present work has been presented from the point of view of general MCSCF wave function optimization with second-order convergence, it is of course straightforward to apply these methods to the special case of single-configuration SCF wave function optimization and perhaps also to cases in which the second derivatives are approximated in some reliable manner.

The present understanding of the periodic nature of E^{Exact} is based on rather pedestrian, linear algebra, concepts. Although several special cases have been analyzed, it is somewhat disconcerting that a single general formalism, which unites all of these cases in a simple manner, has not yet been found. It might be anticipated that some of the powerful methodology of Lie algebras and group theory can be used to understand this problem, and such an investigation is currently underway. Hopefully, new insights into the periodic nature of E^{Exact} gained in this manner can be incorporated into the interpolating function E^{TIF}. Therefore, rather than solving all global convergence problems of MCSCF methods, the current work is expected to be only an initial step in this new direction.

In summary, a new method of MCSCF wave function optimization has been described. This new method is based on multidimensional trigonometric interpolation of the energy within a coordinate space which is a nonlinear transformation of the individual orbital rotation parameters. The interpolating function reproduces exactly the energy, gradient, and hessian at the expansion point, and, due to its periodic nature, also at an infinite number of isolated points within the wave function variational space. Similarities between the coordinate transformation and the orbital coefficient transformation result in the interpolating function also reproducing the exact energy behavior on regular grids within multidimensional, cartesian-product subspaces of the full wave function variational space. The number of such subspaces and their dimensionalities are discussed. In the presence of redundant orbital rotations, the interpolating function and the exact function also share several properties involving continuous coordinate variations. These properties include a multidimensional nested ball structure and the invariance to wave function changes along particular lines within the full space. Numerical examples of MCSCF optimization show that it is possible to converge to energy minimum solutions starting from essentially arbitrary orbitals without resorting to the use of empirical stepsize restrictions.

Acknowledgements. The author acknowledges many discussions concerning various aspects of this work with several colleagues including Drs. A. Banerjee and R. J. Harrison and Professors R. M. Pitzer, I. Shavitt, and H.-J. Werner. This work was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, under contract W-31-109-ENG-38.

References

1. Dalgaard E, Jørgensen P (1978) J Chem Phys 69:3833
2. Yeager DL, Jørgensen P (1979) J Chem Phys 71:755
3. Roothaan CCJ, Detrich JH, Hopper DG (1979) Int J Quantum Chem S13:93; Roothaan CCJ, Detrich JH (1983) Phys Rev A 27:29
4. Shepard R, Simons J (1980) Int J Quantum Chem S14:211
5. Lengsfield III BH (1980) J Chem Phys 73:382
6. Werner HJ, Meyer W (1980) J Chem Phys 73:2342
7. Shepard R (1987) in: Lawley KP (ed) *Ab initio* methods in quantum chemistry II, Advances in Chemical Physics 69. Wiley, NY, p 63
8. Shepard R, Shavitt I, Simons J (1982) J Chem Phys 76:543
9. Roos BO (1980) Int J Quantum Chem S14:175; Roos BO (1983) in: Diercksen GHF, Wilson S (eds) Methods in computational physics, NATO ASI Series C, Vol. 113, Reidel, Dordrecht, p 161
10. Olsen J, Jørgensen P (1985) J Chem Phys 82:3235
11. Helgaker T, Almlöf J (1984) Int J Quantum Chem 26:275; Almlöf J, Taylor PR (1985) Int J Quantum Chem 27:743
12. Shepard R (1987) Int J Quantum Chem 31:33
13. Shepard R, Shavitt I, Pitzer RM, Comeau DC, Pepper M, Lischka H, Szalay PG, Ahlrichs R, Brown FB, Zhao JG (1988) Int J Quantum Chem S22:149
14. Helgaker T, Jørgensen P (1988) Adv Quantum Chemistry 19:183
15. Shepard R, Lischka H, Szalay PG, Kovar T, Ernzerhof M (1992) J Chem Phys 96:2085
16. Shepard R (1981) in: Recent developments and applications of multiconfigurational Hartree–Fock methods, NRCC Proc. No. 10. p. 117, Report LBL-12157, Lawrence Berkeley Laboratory, Univ of California, Berkeley, CA

17. Yeager DL, Albertsen P, Jørgensen P (1980) J Chem Phys 73:2811
18. Olsen J, Yeager DL, Jørgensen P (1983) Adv Chem Phys 54:1
19. Werner HJ (1987) in: Lawley KP (ed) *Ab initio* methods in quantum chemistry II, Advances in Chemical Physics 69. Wiley, NY, p 1
20. Shavitt I (1988) in: Truhlar DG (ed) Mathematical frontiers in computational and chemical physics. Springer-Verlag, Berlin, p 300
21. Grein F, Banerjee A (1974) Chem Phys Lett 25:255; Grein F, Banerjee A (1975) Int J Quantum Chem S9:147; Banerjee A, Grein F (1977) J Chem Phys 66:1054
22. Ruedenberg K, Cheung LM, Elbert ST (1979) Int J Quantum Chem 16:1069; Cheung LM, Sundberg KR, Ruedenberg K (1979) Int J Quantum Chem 16:1103
23. Chang TC, Schwarz WHE (1977) Theor Chim Acta 44:45
24. Ruttink PJA, van Lenthe JH (1977) Theor Chim Acta 44:97
25. Shepard R, Banerjee A, Simons J (1979) J Am Chem Soc 101:6174
26. Fletcher R (1980) Practical methods of optimization, Vol. 1, Wiley, NY
27. Shepard R (manuscript in preparation)
28. King HF, Camp RN, McIver JW (1984) J Chem Phys 80:1171
29. Bar-Natan D (1991) Mathematica Journal 1:69
30. Stork DG, Rocca C (1989) Behavior research methods, instruments, & computers 21:525