

An exact reformulation of the diagonalization step in electronic structure calculations as a set of second order nonlinear equations

WanZhen Liang, and Martin Head-Gordon

Citation: *The Journal of Chemical Physics* **120**, 10379 (2004);

View online: <https://doi.org/10.1063/1.1729870>

View Table of Contents: <http://aip.scitation.org/toc/jcp/120/22>

Published by the [American Institute of Physics](#)

Articles you may be interested in

[Analysis of charge transfer effects in molecular complexes based on absolutely localized molecular orbitals](#)

The Journal of Chemical Physics **128**, 184112 (2008); 10.1063/1.2912041

[Analytical energy gradients for local second-order Møller–Plesset perturbation theory using density fitting approximations](#)

The Journal of Chemical Physics **121**, 737 (2004); 10.1063/1.1760747

[Distance-dependent Schwarz-based integral estimates for two-electron integrals: Reliable tightness vs. rigorous upper bounds](#)

The Journal of Chemical Physics **136**, 144107 (2012); 10.1063/1.3693908

[A general intermolecular force field based on tight-binding quantum chemical calculations](#)

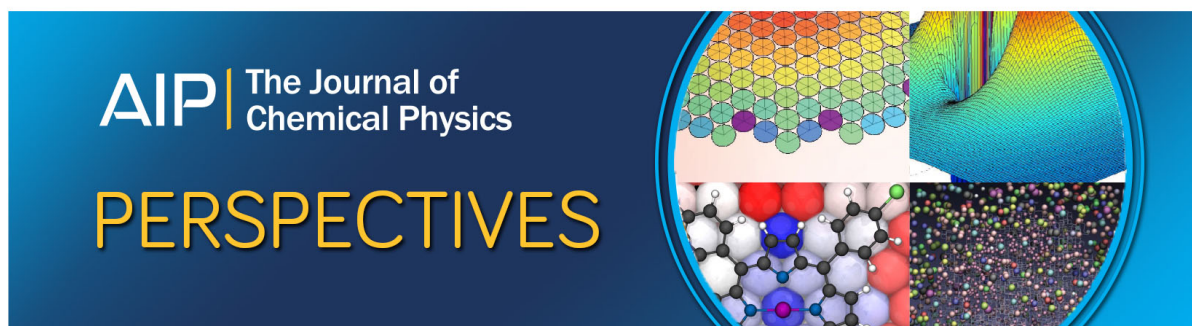
The Journal of Chemical Physics **147**, 161708 (2017); 10.1063/1.4991798

[Incremental full configuration interaction](#)

The Journal of Chemical Physics **146**, 104102 (2017); 10.1063/1.4977727

[Announcement: Top reviewers for *The Journal of Chemical Physics* 2016](#)

The Journal of Chemical Physics **146**, 100201 (2017); 10.1063/1.4978399



An exact reformulation of the diagonalization step in electronic structure calculations as a set of second order nonlinear equations

WanZhen Liang^{a)} and Martin Head-Gordon^{b)}

Department of Chemistry, University of California, Berkeley, California 94720

(Received 29 December 2003; accepted 9 March 2004)

A new formulation of the diagonalization step in self-consistent-field (SCF) electronic structure calculations is presented. It exactly replaces the diagonalization of the effective Hamiltonian with the solution of a set of second order nonlinear equations. The density matrix and/or the new set of occupied orbitals can be directly obtained from the resulting solution. This formulation may offer interesting possibilities for new approaches to efficient SCF calculations. The working equations can be derived either from energy minimization with respect to a Cayley-type parametrization of a unitary matrix, or from a similarity transformation approach. © 2004 American Institute of Physics. [DOI: 10.1063/1.1729870]

I. INTRODUCTION

The self-consistent-field (SCF) model is the fundamental building block of modern electronic structure theory.¹ Within wave-function-based methods, it corresponds to the mean-field Hartree–Fock description of electronic structure theory. As well as being a useful theoretical model chemistry in its own right, the molecular orbitals that result are commonly used as a starting point for the description of electron correlation. Within Kohn–Sham density functional theory,^{2,3} solving for the Kohn–Sham orbitals and the associated Kohn–Sham effective Hamiltonian is also an SCF problem.⁴

The SCF procedure consists of two main steps,¹ effective Hamiltonian construction, and update of the occupied orbital subspace, which are repeated in an iterative fashion until the occupied orbitals and the associated effective Hamiltonian are self-consistent with each other. The first step is to construct the effective Hamiltonian given the current set of occupied molecular orbitals. Algorithms whose computational cost scales asymptotically only linearly with the size of the molecule are available for all three expensive parts of this operation—Coulomb interactions,^{5–9} exact exchange interactions,^{10,11} and exchange–correlation functional evaluation^{12,13} in DFT. The success of these methods means that Fock matrix construction, which was formerly the computationally dominant step, is now no longer necessarily so, at least for large molecules.

The second step, which is the update of the occupied subspace, has been traditionally solved by so-called Roothaan steps,¹⁴ in which the current Fock matrix is diagonalized, and the lowest n orbitals are occupied. Convergence is usually accelerated with Pulay's DIIS methods.¹⁵ The Roothaan step involves matrix manipulations which scale as the third power of molecular size. Therefore for large enough

SCF calculations, it becomes the rate-determining step in SCF calculations. Typically this may occur for calculations involving several thousand basis functions.⁸

There is a large and growing literature on ways to avoid cubic scaling Roothaan step by replacing it with diagonalization-free alternatives. The first phase of development occurred before interest in linear scaling. Methods were developed for direct solution for the density matrix,¹⁶ as well as for direct variation of the SCF energy with respect to molecular orbital degrees of freedom.^{17–20} More recently, the quest for linear scaling has caused a resurgence of interest, and these new developments have been reviewed several times recently.^{21–23} Two of the main alternatives may be briefly summarized as follows. One may attempt to update the one-particle density matrix itself,^{24–28} rather than the molecular orbitals. Recent examples that we have pursued include the Chebyhev approach²⁹ and so-called curvy steps.³⁰ Second, one may attempt to obtain localized molecular orbitals^{31–35} rather than the delocalized orbitals that diagonalize the Fock matrix. Some systematic comparisons of different approaches have been reported in the context of semi-empirical electronic structure methods.³⁶

The purpose of this paper is to introduce an exact reformulation of the Roothaan step that we think offers promise for further development of diagonalization-free methods. We show that the Roothaan diagonalization step can be replaced without approximation by the solution of a set of second order nonlinear equations. The theory is developed as follows. Given the Fock matrix built from the density matrix associated with the current occupied subspace, the density matrix after the Roothaan step is a unitary transformation of the present one, $\mathbf{P}' = \mathbf{U}\mathbf{P}\mathbf{U}^\dagger$, just as the new orbitals are a unitary transformation of the existing orbitals, $\mathbf{C}' = \mathbf{C}\mathbf{U}$. Our theory is based on a representation of a unitary matrix, \mathbf{U} , that while surely known, is seldom used in electronic structure theory. We discuss this representation in some detail first. Expressions for the transformed density matrix are then obtained, and finally we examine the equations that fully determine the unitary transformation. We discuss some of the

^{a)}Present address: Laboratory of Bond Selective Chemistry, University of Science and Technology of China, Hefei, Anhui 230026, People's Republic of China.

^{b)}Author to whom correspondence should be addressed. Electronic mail: mhg@cchem.berkeley.edu

possibilities for developing efficient algorithms based on this formulation. We present a few preliminary numerical examples to illustrate that the equations can be solved, and indeed yield energies identical with a conventional Roothaan step. A companion paper, submitted elsewhere, solves these equations to obtain a dual basis correction (for basis set incompleteness effects) to a self-consistent density functional theory calculation in a small basis set.³⁷

II. THEORY

A. Unitary transformations

There are a number of well-known parametrizations of a unitary transformation. The most widely used one within self-consistent-field theory is the exponential of an antisymmetric matrix,^{28,38} $\mathbf{U} = \exp(\Delta)$. Another alternative is as the product of elementary rotation matrices (Givens transforma-

tions), which has also occasionally been employed.³⁹ A third possibility, which we use here, is the square root of the Cayley form,⁴⁰

$$\mathbf{U} = (\mathbf{1} + \Delta^\dagger \Delta)^{-1/2} (\mathbf{1} + \Delta) = (\mathbf{1} + \Delta) (\mathbf{1} + \Delta^\dagger \Delta)^{-1/2}, \quad (1)$$

where Δ is an antisymmetric matrix, $\Delta^\dagger = -\Delta$. It may be verified that $\mathbf{U}^\dagger \mathbf{U} = \mathbf{U} \mathbf{U}^\dagger = \mathbf{1}$.

In the SCF problem, the variables in Δ associated with mixings within the occupied space do not affect the energy, as can be easily shown. Similarly, mixings of virtual orbitals do not affect the energy. We thus retain only the occupied–virtual and virtual–occupied mixings. Denoting the (rectangular) matrix of nonredundant unknowns as \mathbf{X}_{VO} , we have the explicit parametrization

$$\mathbf{U} = (\mathbf{1} + \mathbf{X}_{VO}^\dagger \mathbf{X}_{VO} + \mathbf{X}_{VO} \mathbf{X}_{VO}^\dagger)^{-1/2} (\mathbf{1} + \mathbf{X}_{VO} - \mathbf{X}_{VO}^\dagger). \quad (2)$$

This can be equivalently expressed as

$$\mathbf{U} = \begin{bmatrix} (\mathbf{1}_{OO} + \mathbf{X}_{VO}^\dagger \mathbf{X}_{VO})^{-1/2} & -(\mathbf{1}_{OO} + \mathbf{X}_{VO}^\dagger \mathbf{X}_{VO})^{-1/2} \mathbf{X}_{VO}^\dagger \\ \mathbf{X}_{VO} (\mathbf{1}_{OO} + \mathbf{X}_{VO}^\dagger \mathbf{X}_{VO})^{-1/2} & (\mathbf{1}_{VV} + \mathbf{X}_{VO} \mathbf{X}_{VO}^\dagger)^{-1/2} \end{bmatrix}. \quad (3)$$

The transformation as written above can only act upon matrices that are defined in the current molecular orbital basis.

We can lift this restriction by introducing appropriate projection operators onto the occupied and virtual subspaces of the full single particle space. Let us denote the projector onto the occupied subspace as \mathbf{P} , and the complementary projector onto the virtual space as $\mathbf{Q} = \mathbf{1} - \mathbf{P}$. Any matrix \mathbf{A} can then be decomposed into OO , VO , OV , and VV subblocks according to

$$\mathbf{A} = (\mathbf{P} + \mathbf{Q}) \mathbf{A} (\mathbf{P} + \mathbf{Q}) = \mathbf{PAP} + \mathbf{QAP} + \mathbf{PAQ} + \mathbf{QAQ}. \quad (4)$$

The matrix of unknowns, denoted as \mathbf{X} henceforth, is only defined in the virtual–occupied block, and thus is required to satisfy the projection conditions

$$\mathbf{X} = \mathbf{QX} = \mathbf{XP} = \mathbf{QXP}. \quad (5)$$

As an immediate consequence we are guaranteed that

$$\mathbf{X}^2 = (\mathbf{QXP})(\mathbf{QXP}) = \mathbf{0}. \quad (6)$$

Equation (2) still holds, but now could either involve separate occupied and virtual expansion spaces, or simply the satisfaction of the projection conditions, Eq. (5), in any convenient orthogonal single particle basis. To summarize, at this stage, we have

$$\begin{aligned} \mathbf{U} &= (\mathbf{1} + \mathbf{X}^\dagger \mathbf{X} + \mathbf{X} \mathbf{X}^\dagger)^{-1/2} (\mathbf{1} + \mathbf{X} - \mathbf{X}^\dagger) \\ &= (\mathbf{1} + \mathbf{X} - \mathbf{X}^\dagger) (\mathbf{1} + \mathbf{X}^\dagger \mathbf{X} + \mathbf{X} \mathbf{X}^\dagger)^{-1/2}. \end{aligned} \quad (7)$$

In any basis, the components of \mathbf{U} are explicitly given [using Eqs. (3) and (4)] as

$$\mathbf{U}_{OO} = \mathbf{P} (\mathbf{1} + \mathbf{X}^\dagger \mathbf{X})^{-1/2} \mathbf{P}, \quad (8)$$

$$\mathbf{U}_{VO} = -(\mathbf{U}_{OV})^\dagger = \mathbf{X} (\mathbf{1} + \mathbf{X}^\dagger \mathbf{X})^{-1/2} \mathbf{P} = \mathbf{Q} (\mathbf{1} + \mathbf{X} \mathbf{X}^\dagger)^{-1/2} \mathbf{X}, \quad (9)$$

$$\mathbf{U}_{VV} = \mathbf{Q} (\mathbf{1} + \mathbf{X} \mathbf{X}^\dagger)^{-1/2} \mathbf{Q} \quad (10)$$

which sum together via Eq. (4) to yield the full matrix, \mathbf{U} .

B. One particle density matrix and Fock matrix

From the unitary transformation (7), and the initial density matrix \mathbf{P} , the transformed density matrix becomes

$$\mathbf{P}' = \mathbf{U} \mathbf{P} \mathbf{U}^\dagger = (\mathbf{P} + \mathbf{X}) (\mathbf{1} + \mathbf{X}^\dagger \mathbf{X})^{-1} (\mathbf{P} + \mathbf{X}^\dagger). \quad (11)$$

Only the occupied–occupied block of the inverse term is required, because the indices of the first and third terms is contracted with span only the occupied space. We can equivalently express the transformed density matrix as

$$\mathbf{P}' = \mathbf{V} \mathbf{V}^\dagger, \quad (12)$$

where the matrix \mathbf{V} comes directly from the OO and VO blocks of \mathbf{U} [cf. Eq. (3)],

$$\mathbf{V} = (\mathbf{P} + \mathbf{X}) (\mathbf{1} + \mathbf{X}^\dagger \mathbf{X})^{-1/2}. \quad (13)$$

C. Stationary conditions for the energy

The energy associated with a fixed Fock matrix,

$$E = \text{Tr}(\mathbf{P}' \mathbf{F}) = \text{Tr}(\mathbf{P} \mathbf{F}') \quad (14)$$

is to be made stationary with respect to variations in the independent variables, \mathbf{X} . This can be accomplished, according to the variational principle, by diagonalizing \mathbf{F} and then occupying the n lowest energy spin–orbitals, if n is the number of electrons. However, the energy is invariant to mixings of the occupied orbitals amongst themselves and the virtual orbitals amongst themselves, so it is sufficient to zero just the occupied–virtual couplings,

$$\mathbf{F}'_{VO} = \mathbf{Q} (\mathbf{U}^\dagger \mathbf{F} \mathbf{U}) \mathbf{P} = [\mathbf{U}_{OV} + \mathbf{U}_{VV}]^\dagger \mathbf{F} [\mathbf{U}_{OO} + \mathbf{U}_{VO}] = \mathbf{0}. \quad (15)$$

Explicitly this is given by

$$(\mathbf{1} + \mathbf{X}\mathbf{X}^\dagger)^{-1/2}(-\mathbf{X} + \mathbf{Q})\mathbf{F}(\mathbf{P} + \mathbf{X})(\mathbf{1} + \mathbf{X}^\dagger\mathbf{X})^{-1/2} = \mathbf{0}. \quad (16)$$

Observing that the leading and trailing prefactors are invertible matrices, we premultiply and post-multiply appropriately to eliminate them, leaving us with the final condition,

$$(\mathbf{Q} - \mathbf{X})\mathbf{F}(\mathbf{P} + \mathbf{X}) = \mathbf{0}. \quad (17)$$

Equations (17) are a second order set of nonlinear equations whose solution, subject to the projection conditions (5), determines the Roothaan step. It is the central result of this paper, which offers an exact re-expression of the conventional diagonalization problem. The density matrix may then be constructed via either Eqs. (11) or (12). Equation (17) may also be obtained from the derivative of the energy with respect to the independent degrees of freedom, \mathbf{X} . After some algebra, one may show that

$$\frac{\partial E}{\partial \mathbf{X}} = (\mathbf{1} + \mathbf{X}\mathbf{X}^\dagger)^{-1}(\mathbf{Q} - \mathbf{X})\mathbf{F}(\mathbf{P} + \mathbf{X})(\mathbf{1} + \mathbf{X}^\dagger\mathbf{X})^{-1}. \quad (18)$$

Thus satisfying Eq. (17) is sufficient to zero out the gradient of the energy with respect to \mathbf{X} .

Finally, we note that if one wants to solve the second order nonlinear equations in an arbitrary orthogonal basis, and automatically obey the projection conditions, then the Fock operator appearing in Eqs. (17) must be properly projected in each term of the equations. Thus we have the alternative form of the nonlinear equations equivalent to the Roothaan diagonalization step,

$$\mathbf{QFP} + (\mathbf{QFQ})\mathbf{X} - \mathbf{X}(\mathbf{PFP}) - \mathbf{X}(\mathbf{PFQ})\mathbf{X} = \mathbf{0}. \quad (19)$$

D. Connection to the exponential of an antisymmetric matrix

Let us relate the variables \mathbf{X} that we have used to parameterize the unitary transformation, Eq. (7), to the exponential of an antisymmetric matrix, $\mathbf{U} = \exp(\mathbf{Y} - \mathbf{Y}^\dagger)$, where the matrix \mathbf{Y} must satisfy the projection properties given in Eq. (5). This is most conveniently done by expressing \mathbf{U} in terms of the full antisymmetric matrix, $\Delta = \mathbf{X} - \mathbf{X}^\dagger$, as given in Eq. (1), which we shall relate to $\exp(\Theta)$, where $\Theta = \mathbf{Y} - \mathbf{Y}^\dagger$. Thus we seek to satisfy

$$\exp(\Theta) = (\mathbf{1} - \Delta\Delta)^{-1/2}(\mathbf{1} + \Delta) = (\mathbf{1} + \Delta)^{1/2}(\mathbf{1} - \Delta)^{1/2}. \quad (20)$$

From basic relationship for hyperbolic functions we immediately obtain

$$\Theta = \frac{1}{2} \ln \frac{(\mathbf{1} + \Delta)}{(\mathbf{1} - \Delta)} = \text{artanh}(\Delta). \quad (21)$$

For small displacements, we may use the power series definition of artanh to obtain the explicit matrix polynomial expression

$$\Theta = \Delta + \frac{\Delta^3}{3} + \frac{\Delta^5}{5} + \dots. \quad (22)$$

Inserting the definitions $\Delta = \mathbf{X} - \mathbf{X}^\dagger$ and $\Theta = \mathbf{Y} - \mathbf{Y}^\dagger$, and using the projection properties (5), we obtain a power series expression for \mathbf{Y} in terms of \mathbf{X} , which is reminiscent of the power series for artan,

$$\mathbf{Y} = \mathbf{X} - \frac{\mathbf{X}\mathbf{X}^\dagger\mathbf{X}}{3} + \frac{\mathbf{X}\mathbf{X}^\dagger\mathbf{X}\mathbf{X}^\dagger\mathbf{X}}{5} - \dots. \quad (23)$$

This equation suggests that the magnitude of the matrix \mathbf{Y} will be less than that of the matrix \mathbf{X} , once the displacements become noninfinitesimal. Indeed this must be the case, because it has been proved⁴¹ that $\mathbf{U} = \exp(\mathbf{Y} - \mathbf{Y}^\dagger)$ is a geodesic (or shortest length) displacement on a Grassman manifold, and therefore any other (inequivalent) representation cannot also yield geodesic steps. The new variables \mathbf{X} thus bear a resemblance to the tangent of the usual rotation angles \mathbf{Y} —a relationship that is precisely true for the toy problem considered later. In the limit of very small values, the variables are the same. As one approaches this limit, solving the new equations is related to pseudodiagonalization,⁴² which is the use of one Jacobi sweep⁴⁰ (or alternatively several such sweeps) to replace exact diagonalization.

E. Discussion

The central result, Eqs. (17) or (19), can also be derived in other ways. One is by using similarity transformations of the effective Hamiltonian, $\bar{\mathbf{F}} = \mathbf{\Omega}^{-1}\mathbf{F}\mathbf{\Omega}$, which preserve its eigenvalue spectrum, and therefore can be used as an alternative way to zero the virtual-occupied couplings. Such methods are used for constructing effective Hamiltonians in model spaces, for example, such as in nuclear physics,⁴³ and in multireference perturbation theory,⁴⁴ in molecular electronic structure theory. The most closely related choice for a nonunitary similarity transformation is one that reduces to Eq. (7) in the limit of \mathbf{X} small,

$$\mathbf{\Omega} = \mathbf{1} + \mathbf{X} - \mathbf{X}^\dagger. \quad (24)$$

\mathbf{X} is required to satisfy the projection conditions (5), as before. The condition for uncoupling occupied and virtual blocks of the transformed effective Hamiltonian can be shown as

$$\begin{aligned} \bar{\mathbf{F}}_{VO} &= \mathbf{Q}(\mathbf{\Omega}^{-1}\mathbf{F}\mathbf{\Omega})\mathbf{P} \\ &= (\mathbf{1} + \mathbf{X}\mathbf{X}^\dagger)^{-1}(\mathbf{Q} - \mathbf{X})\mathbf{F}(\mathbf{P} + \mathbf{X}) = \mathbf{0}. \end{aligned} \quad (25)$$

This leads us to the conditions derived before, Eq. (17), after premultiplying by $(\mathbf{1} + \mathbf{X}\mathbf{X}^\dagger)$.

To obtain a Hermitian density matrix corresponding to the similarity-transformed Hamiltonian, we must connect the similarity transformation back to an equivalent unitary transformation. Since the conditions on \mathbf{X} are identical to those derived earlier from the unitary transformation (7), it is clear that this desired matrix is Eq. (7). Indeed \mathbf{U} could be defined from the polar decomposition,⁴⁵ $\mathbf{\Omega} = (\mathbf{\Omega}\mathbf{\Omega}^\dagger)^{1/2}\mathbf{U}$, of the similarity transformation, so that $\mathbf{U} = (\mathbf{\Omega}\mathbf{\Omega}^\dagger)^{-1/2}\mathbf{\Omega}$, which leads directly to Eq. (7) again. Thus all of the results obtained within a unitary framework can be equivalently derived from the similarity transformation, Eq. (24).

The central equations can also be represented in other ways. We have used an orthogonal basis to represent all

quantities thus far, because it permits a simple matrix notation. However, while convenient, this is not essential. This can be seen by regarding X as a single particle operator that induces coupling between the starting set of occupied and virtual states. In operator terms, Eq. (17) and the projection conditions (5), are written as

$$\hat{Q}(1 + \hat{X}\hat{P})\hat{F}(1 - \hat{Q}\hat{X})\hat{P} = 0. \quad (26)$$

This operator equation is invariant to any change of basis, as long as the transformed functions still span the same space. This is most conveniently expressed in practice using tensor concepts and notation.^{46–48} One can also eliminate the requirement for projection conditions, by employing separate expansion spaces for the occupied and virtual indices. This may be convenient in practice—for example, one could employ localized occupied orbitals to span the occupied space, and projected atomic orbitals to span the virtual space.

III. APPLICATIONS

A. Toy example

We have previously used minimal basis HeH^+ to explore the nature of displacements when minimizing a tight-binding or self-consistent-field (SCF) energy.^{47,49} It is also a suitable vehicle to illustrate the nature of a Roothaan step via solution of Eq. (17). In the symmetrically orthogonalized basis, we take the SCF Hamiltonian (corresponding to the converged STO-3G basis density matrix at bondlength 0.93 Å) as

$$\mathbf{F} = \begin{bmatrix} -1.333 & -0.451 \\ -0.451 & -0.459 \end{bmatrix}. \quad (27)$$

The matrix \mathbf{X} in this case is characterized by a single degree of freedom, x , which couples the one (initial guess) occupied orbital \mathbf{o} and the one (corresponding) virtual orbital \mathbf{v} ,

$$\mathbf{X} = x\mathbf{v}\mathbf{o}^\dagger. \quad (28)$$

The alpha (or beta) density matrix is simply $\mathbf{P} = \mathbf{o}\mathbf{o}^\dagger$, and the optimal density matrix, \mathbf{P}_{opt} , which minimizes the energy by diagonalizing \mathbf{F} is

$$\mathbf{P}_{\text{opt}} = \begin{bmatrix} 0.848 & 0.359 \\ 0.359 & 0.152 \end{bmatrix}. \quad (29)$$

We can now explore how the energy depends on x when we start from initial guess density matrices that are not optimal. Figure 1 shows how the energy depends on x as defined above, when the initial guess density matrix corresponds to two electrons localized on the He atom,

$$\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \quad (30)$$

For comparison the fainter dashed curve in the upper panel also indicates how the energy depends on the single degree of freedom which generates unitary transformations as the exponential of the corresponding antisymmetric matrix. In the lower panel of the figure is the deviation of the equations (17), from the optimal value of zero, as a function of x . This initial guess is close to the minimum of the energy, and it is evident from Fig. 1 that only a small displacement from the

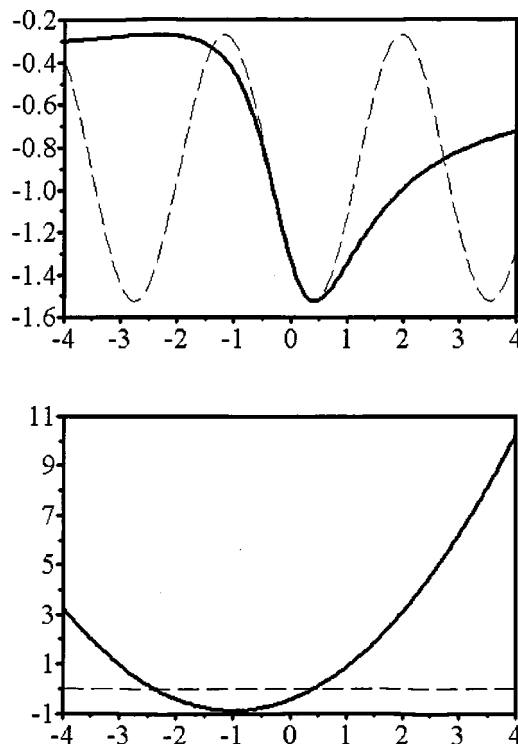


FIG. 1. The energy in minimum basis HeH^+ starting from a good initial guess, where the two electrons are both localized on the He atom. The upper panel shows the energy (in a.u.) as a function of the single degree of freedom, x , on the bold curve. The dashed curve in the background, for comparison, is the energy of the corresponding degree of freedom that generates a unitary transformation as the exponential of an antisymmetric matrix. In the lower panel is the deviation of the quadratic equations from the desired value of 0, as a function of x .

starting point $x=0$ is necessary to reach the zero of the equations which corresponds to the minimum energy. There is a second zero, much further from the starting point, which maximizes the energy. Figure 2 shows the same data, but for a poor initial guess, where the initial guess density matrix corresponds to having both electrons localized on the H atom. In this case a quite large displacement is required to move from the initial guess ($x=0$) to the optimal value ($x=2.4$) that minimizes the energy.

The differences in Figs. 1 and 2 between displacements via our new equations and those corresponding to exponentials of antisymmetric matrices are very interesting. The two variables (let us call them x and y , respectively) give similar steps for values that deviate only slightly from zero, because they have the same first order Taylor expansions. For larger displacements, the change in the new variables x is always larger to achieve a given energy change from the comparison between bold and dashed curves. This is a direct reflection of the fact that displacements (y) in terms of the exponential of the antisymmetric matrix are geodesic steps, and accordingly steps in x that get to the same physical density matrix from a given starting point (0) must be larger. Indeed the limiting values for $E(x)$ as x becomes large in magnitude are just those corresponding to y approaching $\pm\pi/2$. This follows directly from the exact relationship between them, Eq. (23), which for this one variable case reduces to $y = \arctan x$.

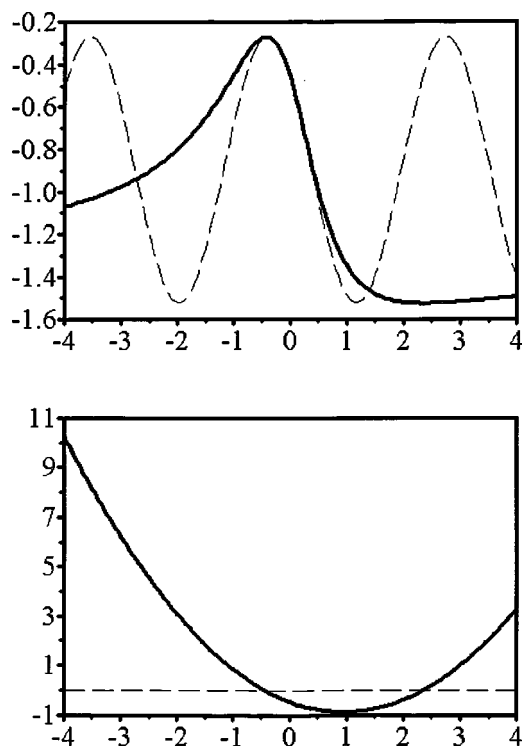


FIG. 2. The energy in minimum basis HeH^+ starting from a poor initial guess, where the two electrons are both localized on the H atom. The upper panel shows the energy (in a.u.) as a function of the single degree of freedom, x , on the bold curve. The dashed curve in the background, for comparison, is the energy of the corresponding degree of freedom that generates a unitary transformation as the exponential of an antisymmetric matrix. In the lower panel is the deviation of the quadratic equations from the desired value of 0, as a function of x .

B. Implementation and preliminary tests

We have developed a pilot iterative solver for the non-linear equations (17), to drive self-consistent-field calculations. This algorithm solves the equations in an orthogonalized basis, which derives from the Cholesky decomposition of the overlap matrix, $\mathbf{S} = \mathbf{L}\mathbf{L}^\dagger$, and the corresponding inverse factors, $\mathbf{Z} = \mathbf{L}^{-1}$, which decompose the inverse of the overlap matrix as $\mathbf{S}^{-1} = \mathbf{Z}^\dagger \mathbf{Z}$. If a row vector of the atomic orbitals is denoted as ω_{AO} , then the orthogonalized functions are $\omega = \omega_{\text{AO}} \mathbf{Z}^\dagger$, and the effective Hamiltonian in the orthogonalized basis is $\mathbf{F} = \mathbf{Z} \mathbf{F}_{\text{AO}} \mathbf{Z}^\dagger$. This basis has been commonly used in linear scaling density functional theory calculations.^{27,50} The algorithm requires a valid set of initial

guess molecular orbital coefficients (or an idempotent density matrix). We take these as the converged MO's from a minimum basis calculation (M), transformed into the (larger) working basis via $\mathbf{C} = \mathbf{S}^{-1} \mathbf{S}_2 \mathbf{C}_M$, where \mathbf{S}_2 is the matrix of overlaps between the working basis and the minimal basis. On subsequent SCF cycles the initial guess orbitals (or density) are those from the previous cycle. Additionally we apply Pulay's DIIS extrapolation¹⁵ to generate the effective Hamiltonian on the second and subsequent SCF steps, just as in a conventional SCF calculation.

On a given iteration, we solve Eq. (17) in the orthogonalized basis, and with the specified initial guesses as follows. To ensure that the projection conditions (5), for X are satisfied, the equations are solved in the form (19). First, simple conjugate gradient steps are taken to solve the linearized version of Eq. (19) (i.e., the second order terms are completely neglected). In these preliminary tests, no preconditioning is used at all. We denote the *target* error and the *actual* error in the SCF (i.e., the RMS of the gradient, QFP, with the Fock operator itself computed from P) as 10^{-T} and 10^{-D} , respectively. The *working* accuracy, 10^{-W} for converging the linear equations is then set to be $W = \min(D+3, T+3)$. The second step, after this convergence is achieved, is to perform simple conjugate gradient iterations (again with no preconditioning) to solve the nonlinear equations (using the solution to the linear equations as the starting guess), with a target working accuracy of $W = \min(D+1, T+1)$.

This procedure has been implemented in a development version of the Q-Chem electronic structure program,⁵¹ and we have employed it for several simple tests to indicate that the equations can be solved. These results, expressed in terms of the number of matrix multiples required to achieve convergence (for $T=5$), are summarized in Table I. The test cases are two relatively large molecules in the 6-31G** basis, although both represent quite easy SCF calculations to converge. The results show that the number of matrix multiplies required by this pilot scheme to solve the second order nonlinear equations is already roughly competitive with the number of matrix multiples required by a state-of-the-art linear scaling approach (the Fermi operator expansion method).²⁹ Given the very simple nature of the iterative algorithm used in this pilot implementation, this is quite encouraging.

TABLE I. The number of iterations and matrix multiples required for SCF calculations in the 6-31G** basis, using the pilot conjugate gradient method to solve the second order nonlinear equations. These calculations are examples we have used recently to assess the number of matrix multiples required to obtain the density matrix via Fermi operator expansions (Ref. 29).

Molecule (method)	Matrix multiples per SCF iteration							ΔE^a (a.u.)	Average (present) ^b	Average (FOE) ^c
	1	2	3	4	5	6	7			
$\text{C}_{60}\text{H}_{122}$ (BLYP)	48	60	66	66	74			7.4×10^{-7}	63	47
$(\text{H}_2\text{O})_{72}$ (HF)	44	50	50	58	58	66	66	2.7×10^{-6}	56	43

^aThe deviation between the converged energy, and that obtained by tightly converging the calculation using conventional SCF methods.

^bThe average number of matrix multiples required per iteration in the present method.

^cThe average number of matrix multiples required per iteration in the Fermi operator expansion (FOE) method reported in Ref. 29.

IV. SUMMARY

This paper has developed a new reformulation of the Roothaan diagonalization step, the fundamental step in a tight binding calculation, or a self-consistent-field iteration. We have shown that the diagonalization can be exactly re-expressed as the solution of a second order nonlinear set of equations. We have discussed the connection between the solutions of these equations and the corresponding representation of a unitary transformation in terms of the exponential of an antisymmetric matrix. Additionally, a number of the basic properties of this reformulation of the Roothaan diagonalization step have been illustrated with calculations on a toy model problem, and a pilot implementation that solves the equations for realistic problems has been reported, using conjugate gradients. Future work needs to address fully the question of developing more efficient algorithms that exploit locality and sparsity to solve these equations.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation through Grant No. CHE-9981997. This work was also supported indirectly by funding from Q-Chem Inc. via an SBIR subcontract from the National Institutes of Health (R43GM069255-01). M.H.G. is a part owner of Q-Chem Inc.

- ¹T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular Electronic Structure Theory* (Wiley, Chichester, 2000).
- ²R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford, New York, 1989).
- ³W. Kohn, A. D. Becke, and R. G. Parr, *J. Phys. Chem.* **100**, 12974 (1996).
- ⁴J. A. Pople, P. M. W. Gill, and B. G. Johnson, *Chem. Phys. Lett.* **199**, 557 (1992).
- ⁵C. A. White, B. G. Johnson, P. M. W. Gill, and M. Head-Gordon, *Chem. Phys. Lett.* **230**, 8 (1994).
- ⁶M. C. Strain, G. E. Scuseria, and M. J. Frisch, *Science* **271**, 51 (1996).
- ⁷C. A. White, B. G. Johnson, P. M. W. Gill, and M. Head-Gordon, *Chem. Phys. Lett.* **253**, 268 (1996).
- ⁸Y. H. Shao, C. A. White, and M. Head-Gordon, *J. Chem. Phys.* **114**, 6572 (2001).
- ⁹L. Fusti-Molnar and P. Pulay, *J. Chem. Phys.* **117**, 7827 (2002).
- ¹⁰E. Schwegler, M. Challacombe, and M. Head-Gordon, *J. Chem. Phys.* **106**, 9708 (1997).
- ¹¹C. Ochsenfeld, C. A. White, and M. Head-Gordon, *J. Chem. Phys.* **109**, 1663 (1998).
- ¹²J. M. Perezjorda and W. T. Yang, *Chem. Phys. Lett.* **241**, 469 (1995).
- ¹³R. E. Stratmann, G. E. Scuseria, and M. J. Frisch, *Chem. Phys. Lett.* **257**, 213 (1996).
- ¹⁴C. C. J. Roothaan, *Rev. Mod. Phys.* **23**, 69 (1951).
- ¹⁵P. Pulay, *J. Comput. Chem.* **3**, 556 (1982).
- ¹⁶R. McWeeny, *Rev. Mod. Phys.* **32**, 335 (1960).
- ¹⁷J. Douady, Y. Ellinger, R. Subra, and B. Levy, *J. Chem. Phys.* **72**, 1452 (1980).
- ¹⁸G. B. Bacskay, *Chem. Phys.* **61**, 385 (1981).
- ¹⁹G. Chaban, M. W. Schmidt, and M. S. Gordon, *Theor. Chem. Acc.* **97**, 88 (1997).
- ²⁰T. Van Voorhis and M. Head-Gordon, *Mol. Phys.* **100**, 1713 (2002).
- ²¹S. Goedecker, *Rev. Mod. Phys.* **71**, 1085 (1999).
- ²²G. E. Scuseria, *J. Phys. Chem. A* **103**, 4782 (1999).
- ²³D. R. Bowler, T. Miyazaki, and M. J. Gillan, *J. Phys.: Condens. Matter* **14**, 2781 (2002).
- ²⁴X. P. Li, W. Nunes, and D. Vanderbilt, *Phys. Rev. B* **47**, 10891 (1993).
- ²⁵S. Goedecker and L. Colombo, *Phys. Rev. Lett.* **73**, 122 (1994).
- ²⁶A. H. R. Palser and D. E. Manolopoulos, *Phys. Rev. B* **58**, 12704 (1998).
- ²⁷M. Challacombe, *J. Chem. Phys.* **110**, 2332 (1999).
- ²⁸T. Helgaker, H. Larsen, J. Olsen, and P. Jorgensen, *Chem. Phys. Lett.* **327**, 397 (2000).
- ²⁹W. Z. Liang, C. Saravanan, Y. H. Shao, R. Baer, A. T. Bell, and M. Head-Gordon, *J. Chem. Phys.* **119**, 4117 (2003).
- ³⁰Y. Shao, C. Saravanan, M. Head-Gordon, and C. A. White, *J. Chem. Phys.* **118**, 6144 (2003).
- ³¹F. Mauri, G. Galli, and R. Car, *Phys. Rev. B* **47**, 9973 (1993).
- ³²P. Ordejon, D. A. Drabold, M. P. Grumbach, and R. M. Martin, *Phys. Rev. B* **48**, 14646 (1993).
- ³³J. J. P. Stewart, *Int. J. Quantum Chem.* **58**, 133 (1996).
- ³⁴W. T. Yang, *Phys. Rev. B* **56**, 9294 (1997).
- ³⁵J. Van de Vondele and J. Hutter, *J. Chem. Phys.* **118**, 4365 (2003).
- ³⁶A. D. Daniels and G. E. Scuseria, *J. Chem. Phys.* **110**, 1321 (1999).
- ³⁷W. Z. Liang and M. Head-Gordon, *J. Phys. Chem. A* **108**, 3206 (2004).
- ³⁸J. Hutter, M. Parrinello, and S. Vogel, *J. Chem. Phys.* **101**, 3862 (1994).
- ³⁹M. Head-Gordon and J. A. Pople, *J. Phys. Chem.* **92**, 3063 (1988).
- ⁴⁰G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. (Johns Hopkins, Baltimore, 1989).
- ⁴¹A. Edelman, T. A. Arias, and S. T. Smith, *SIAM J. Matrix Anal. Appl.* **20**, 303 (1998).
- ⁴²J. J. P. Stewart, P. Csaszar, and P. Pulay, *J. Comput. Chem.* **3**, 227 (1982).
- ⁴³K. Suzuki and R. Okamoto, *Prog. Theor. Phys.* **92**, 1045 (1994).
- ⁴⁴I. Shavitt and L. T. Redmon, *J. Chem. Phys.* **73**, 5711 (1980).
- ⁴⁵P. O. Lowdin, *Linear Algebra for Quantum Theory* (Wiley, New York, 1998).
- ⁴⁶E. Artacho and L. M. Delbosch, *Phys. Rev. A* **43**, 5770 (1991).
- ⁴⁷C. A. White, P. Maslen, M. S. Lee, and M. Head-Gordon, *Chem. Phys. Lett.* **276**, 133 (1997).
- ⁴⁸M. Head-Gordon, P. E. Maslen, and C. A. White, *J. Chem. Phys.* **108**, 616 (1998).
- ⁴⁹M. Head-Gordon, Y. H. Shao, C. Saravanan, and C. A. White, *Mol. Phys.* **101**, 37 (2003).
- ⁵⁰J. M. Millam and G. E. Scuseria, *J. Chem. Phys.* **106**, 5569 (1997).
- ⁵¹J. Kong, C. A. White, A. I. Krylov *et al.*, *J. Comput. Chem.* **21**, 1532 (2000).