Pramāṇa – J. Phys., Vol. 30, No. 5, May 1988, pp. 387–405. © Printed in India.

# The minimax technique in relativistic Hartree-Fock calculations

S N DATTA and G DEVAIAH

Department of Chemistry, Indian Institute of Technology, Powai, Bombay 400076, India

MS received 24 August 1987; revised 14 January 1988

**Abstract.** Using the set of trial spinors $\left\{N_i\begin{pmatrix} u_i \\ \hat{\Omega}_i u_i \end{pmatrix}, i = 1, \ldots, N\right\}$ and the Dirac-Coulomb Hamiltonian ($H_{DC}$) we discuss the role of the minimax theorem in relativistic Hartree-Fock calculations. In principle, the minimax theorem guarantees the occurrence of an upper bound. We also consider a scaling of the functions $u_i$ and discuss the condition to derive the relativistic hypervirial theorem; the variational procedure represented by the condition serves as an example of the minimax technique. Single zeta calculations on $H_2^+$, $H_2$ and He are analysed. The effect of enlarging the basis is investigated for the He atom. The "upper bound" obtained by using *coherent* basis spinors differs from the result of the (random) linear variation using the kinetically balanced basis set by an amount which is at most of order $c^{-4}$. Use of the *coherent* basis set is advocated.

**Keywords.** Minimax technique; relativistic Hartree-Fock; virial theorem; coherent spinors.

**PACS Nos** 03·65; 31·15; 31·30

## 1. Introduction

Many methods of quantum chemistry are based on the use of the variational principle for solving a wave equation. These methods work whenever the eigenvalue that is sought is the absolute minimum of the expectation value of some operator as a functional of trial functions of a certain class. But, as the Dirac Hamiltonian ($H_D$) has negative-energy eigenfunctions, the minimum of the expectation value $\langle H_D \rangle$ may not be an upper bound to the lowest positive-energy eigenvalue ($\varepsilon_0$). It is quite possible that an upper bound is obtained from variation in a restricted domain; however, the bound will be in error unless the negative-energy functions are optimally removed from the trial spinor. Early investigators of relativistic atoms were aware of this fact (Kim 1967; Desclaux 1973). However, since the angular form of an atomic spinor is exactly known and the radial components can be easily estimated, this difficulty was not serious. A molecular spinor is devoid of these simplifying features (Oreg and Malli 1976; Datta and Ewig 1982; Lee and McLean 1982). Therefore, the need of a method for separating the negative energy states from the trial molecular spinors was felt (Datta 1980). Such a method must have two important features: (i) it must apply to systems beyond the non-relativistic limit, and (ii) calculations using even a small basis set should give a good estimate of the relativistic corrections. In the non-relativistic limit, relativistic corrections can be calculated by using the Pauli equation.

Many authors have attempted to find a safe method of relativistic calculations (Rosicky and Mark 1975; Datta 1980, 1984; Morrison and Moss 1980; Drake and

Goldman 1981; Datta and Ewig 1982; Lee and McLean 1982; Mark and Schwarz 1982; Schwarz and Wallmeier 1982; Schwarz and Wechsel-Trakowski 1982; Gazdy 1983; Ishikawa *et al* 1983; Ketley and Moss 1983; Ladik *et al* 1983; Datta and Jagannathan 1984; Dyall *et al* 1984; Gazdy and Ladanyi 1984; Laaksonen and Grant 1984a, b; Stanton and Havriliak 1984; Sepp and Fricke 1985; Wood *et al* 1985; Datta and Devaiah 1986; Grant 1986; Rutkowsky 1986). Here we focus our attention only on the approaches which utilize the standard variational principle (that is, extremization of $\langle H_D \rangle$) *in order to exclude* the negative-energy states.

The Dirac equation, $H_D\psi = \varepsilon\psi$ has four-component solutions (spinors). Here, $H_D = c^2\beta + c\alpha\cdot p + V, V$ being an external potential. A trial spinor is written as $N\begin{pmatrix} u \\ l \end{pmatrix}$ where the upper ($u$) and the lower ($l$) components themselves are two-component spinors. We assume that $l = \hat{\Omega}u$ and study the variation of $\langle H_D \rangle$ with respect to $\hat{\Omega}$. If $\delta\langle H_D \rangle = 0$ for all variations $\delta\hat{\Omega}$ and arbitrary $u$, then $\hat{\Omega} = \hat{\Omega}^0 = c(\varepsilon^0 + c^2 - V)^{-1}\sigma\cdot p$ where $\varepsilon^0$ is the stationary value and $\sigma$'s are the Pauli spin matrices. Rosicky and Mark (1975) pointed out that $\hat{\Omega}^0$ gives an adequate representation of the lower component. The two components, $Nu$ and $N\hat{\Omega}^0 u$, are said to be *kinetically balanced.* Lee and McLean (1982) as well as Ishikawa *et al* (1983) used $\hat{\Omega} = \hat{\Omega}_a = (2c)^{-1}\sigma\cdot p$ to form approximately balanced basis spinors and contended that the final kinetic balance can be achieved by using a large basis set. A "kinetically balanced" basis set consists of the set of upper component functions $\{u\}$ and the set of lower component functions $\{\chi_l\}$ where $\chi = \sigma\cdot p u\mathcal{N}$, $\mathcal{N}$ being an orthonormalizing factor (Mark and Schwarz 1982; Schwarz and Wallmeier 1982; Schwarz and Wechsel-Trakowski 1982). Stanton and Havriliak (1984) have shown that if the basis set is "kinetically balanced", the variational calculation is stable with the converged matrix eigenvalue correct through order $c^{-2}$. Wood *et al* (1985) investigated the variational properties associated with the partitioning of the Dirac equation in matrix representation and prescribed a set of separation theorems. A rigorous theorem on the eigenvalue distribution has been presented by Grant (1986).

As part of the ongoing study in our laboratory, we had proposed a projection operator technique to compensate for variational collapse (Datta 1980) and found by using the relativistic orbital notation suggested by Oreg and Malli (1976) that contamination from the $\approx -2c^2$ atomic positronic spinors is small for the $Be_2$ molecule (Datta and Ewig 1982). Actually, the process of finding the optimum $\hat{\Omega}$ makes the electronic spinor consistent with a trial $\approx -2c^2$ positronic spinor (Datta and Jagannathan 1984). In order to illustrate the constrained-component variation we chose $\hat{\Omega} = \hat{\Omega}_t = \Lambda\sigma\cdot p$ where $\Lambda$ is a variational parameter (Datta and Jagannathan 1984) and used it to discuss the one-electron virial theorem (Datta 1984) and to point out calculational ambiguities associated with the relativistic density functional treatment (Datta 1987). The use of $\hat{\Omega}_t$ can be described as the method of kinetically balanced *trial* spinors. In fact, the method of "kinetically balanced" basis set involves a random linear variation (subject to the constraint of orthogonality) using the set of positive-energy basis spinors

$$\left\{\phi_+ = \begin{pmatrix} u \\ 0 \end{pmatrix}\right\}$$

and the set of negative-energy basis spinors

$$\left\{\phi_- = N_-\begin{pmatrix}0\\ \sigma\cdot pu\end{pmatrix}\right\};$$

still, if the system is not too relativistic the positive-energy eigenvectors will not differ much from the *coherent* form

$$\begin{pmatrix} v \\ \Lambda^0\sigma\cdot pv\end{pmatrix},$$

where

$$v = \sum_i c_i u_i$$

and $\Lambda^0$ is an optimized parameter. Besides, the use of the coherent trial spinors gives rise to the viewpoint of the minimax technique in relativistic Hartree-Fock (HF) calculations.

There is a nice compact discussion on the variational collapse problem of the Dirac Hamiltonian in a recent contribution by Schwarz (1987). This is particularly relevant since Schwarz discusses alternative approaches to stationarity. Of course, the exact $u-l$ coupling can be achieved by the numerical integration techniques. Other approaches are: (i) application of the appropriate boundary conditions, (ii) the use of a basis set which guarantees the coupling to a high degree of accuracy, and (iii) the choice of a specific formulation of the Hamiltonian (Mark and Schwarz 1982; Kutzelnigg 1984) where the coupling is insensitive to basis set truncation errors. In all these analytical methods, matrix eigenvalues can be converged towards the exact values by extending finite basis sets in an appropriate manner but the bound is not well specified, and the convergence may even occur from below (Hegarty 1986). That the energies calculated with the appropriate boundary conditions are rigorous upper bounds "has not been substantiated" (Schwarz 1987) except for a special case (Drake and Goldman 1981; Goldman 1985). In the main part of this paper we will show, using very simple mathematics, that for any upper-component basis set the correct $u-l$ coupling ensures that the variational energy is an *upper bound*. In the light of the minimax theorem which we discuss in the next section, the variational energy is seen to be an upper bound if the step of maximization is *exactly* performed.

The detailed objectives of this paper are:

(i) to show that if both $\hat{\Omega}$ and $u$ of a trial spinor

$$N\begin{pmatrix}u\\ \hat{\Omega}u\end{pmatrix}$$

are treated as variational parameters, the minimax theorem holds, and to prescribe the minimax theorem for the average value of the Dirac-Coulomb Hamiltonian ($H_{DC}$) when the many-electron function is written as a Slater determinant,

(ii) to find the sufficient condition for the relativistic virial theorem for atomic electrons and the relativistic hypervirial theorem (for molecular electrons),

(iii) to illustrate the effect of contamination from the negative-energy basis spinors

$$N_-\begin{pmatrix}0\\ \sigma\cdot pu\end{pmatrix},$$

and

(iv) to show that for a numerical test case (the He atom) the difference between the result of the coherent linear variation (that is, linear variation using coherent trial spinors) and that of the (random) linear variation using a "kinetically balanced" basis set is truly insignificant.

The first objective leads to the second one. From the second objective we see that if the relativistic virial theorem (hypervirial theorem in the molecular case) is satisfied, the Slater determinant lies almost in the subspace spanned by the (positive-energy) eigenfunctions of the Dirac-Hartree-Fock (DHF) Hamiltonian. The third objective is not trivial since in some cases the use of $\Lambda$ as a variable parameter will lead to a result which is visibly superior to the result obtainable from $\Lambda = (2c)^{-1}$. The fourth objective indicates that the (random) linear variation using a kinetically balanced basis set can be effectively substituted by the coherent linear variation method.

A substantial part of the work (the one related to the minimax treatment and one-electron calculations) was completed in 1984. The reports of Stanton and Havriliak (1984) on finite basis calculations and of Wood *et al* (1985) and Grant (1986) on the partitioning of the Dirac equation in matrix representation convince us of the merit of our approach which relies on the partitioning technique applied for each spinor individually so that the minimax viewpoint can be presented. The second part of this work (the part related to the many-electron treatment) was orally presented by one of us (SND) at the Adriatico Research Conference on relativistic many-body problems (Datta and Devaiah 1986). This paper is written in an integrated format. In § 2 we derive the minimax theorem, at first for one spinor only and then for a Slater determinant. In § 3 we obtain a sufficient condition for the validity of the relativistic virial theorem (or the relativistic hypervirial theorem). Numerical results indicating the superiority of the $\Lambda$-optimization procedure and the efficiency of the coherent linear variation method are discussed in § 4. Finally we have summarized the limitations of the present work.

## 2. The minimax theorem

### 2.1 *The one-electron case*

2.1a *Maximization*: We describe a one-electron system by the normalized four-component trial spinor

$$\psi = N\begin{pmatrix}u\\ l\end{pmatrix},$$

where $u$ is a normalized two-component function, $l = \hat{\Omega}u$ and

$$N = (u, [1 + \hat{\Omega}^\dagger\hat{\Omega}]u)^{-1/2}.$$

We write

$$\varepsilon(u,[1+\hat{\Omega}^{\dagger}\hat{\Omega}]u)=(u,[(c^2+V)+c\sigma\cdot p\hat{\Omega}+\hat{\Omega}^{\dagger}c\sigma\cdot p+\hat{\Omega}^{\dagger}(-c^2+V)\hat{\Omega}]u), \tag{1}$$

where $\varepsilon=\langle H_D\rangle$, and study the variational properties of $\varepsilon$.

Firstly, if we write $\psi=c_1\phi_+ + c_2\phi'_-$ where

$$\phi_+=\begin{pmatrix}u\\0\end{pmatrix} \quad \text{and} \quad \phi'_-=N'_-\begin{pmatrix}0\\\hat{\omega}u\end{pmatrix},$$

then $\hat{\Omega}=(c_2/c_1)N'_-\hat{\omega}$. The linear variation method gives two eigenvalues, $\varepsilon_+$ and $\varepsilon_-$; $\varepsilon_- < -c^2+(V)_{\hat{\omega}u}$ and $\varepsilon_+ > c^2+(V)_u$. The abbreviation $(X)_v$ denotes $(v,v)^{-1}(v,Xv)$. Now, $\varepsilon_-$ is the minimum of $\varepsilon$; since the two basis spinors are orthogonal to each other, it is not difficult to see that $\varepsilon_+$ is a maximum.

Secondly, in (1), if $\delta\varepsilon=0$ for all variations $\delta\hat{\Omega}^{\dagger}$ and arbitrary $u$, then $\hat{\Omega}=\hat{\Omega}^0=(\varepsilon^0+c^2-V)^{-1}c\sigma\cdot p$ and (1) reduces to

$$(u,[\varepsilon^0-c^2-V]u)=c^2(u,\sigma\cdot p[\varepsilon^0+c^2-V]^{-1}\sigma\cdot pu), \tag{2}$$

where $\varepsilon^0$ is the stationary value. Equation (2) was first derived by Rosicky and Mark (1975). The latter authors showed that in the free particle case (2) gives two decoupled equations for the positive- and negative-energy states separately, but they used Pauli approximation to treat the problem of one electron in a central field. Here we consider the case of an attractive potential, $V<0$ and assume that $(V)_u > -2c^2$. Then (2) can have real roots $\varepsilon^0_+ > c^2+(V)_u$; in addition, real roots $\varepsilon^0_- < -c^2$ may exist. It is easy to prove that for a specific $u$ there can be one and only one positive-energy root. One can also show, by *reductio ad absurdum*, that for a well-behaved and normalizable $u$ a real negative-energy root does not exist. Let the roots $\varepsilon^0_+$ and $\varepsilon^0_-$ coexist and let the $r_0$-surface be defined by $\varepsilon^0_- + c^2 - V(\mathbf{r}_0)=0$. If $\sigma\cdot pu$ is well behaved then

$$c^2(\sigma\cdot pu,[\varepsilon^0_- + c^2 - V]^{-1}[\varepsilon^0_+ + c^2 - V]^{-1}\sigma\cdot pu) = -1. \tag{2a}$$

The integral on the left hand side of (2a) does not exist since the integrand has a singularity at $\mathbf{r}=\mathbf{r}_0$; the Cauchy principal value, if it exists, is in general a complex number with a non-vanishing imaginary part. Hence (2a) does not hold good. Simultaneous negative- and positive-energy solutions are possible if $V=0$.

From (1), it is easy to show that

$$\delta^2\varepsilon|_{\hat{\Omega}=\hat{\Omega}^0_+}<0,$$

that is, the expectation value $\varepsilon^0_+$ is clearly a maximum. Besides, $\varepsilon_+=\varepsilon^0_+$ when $\hat{\omega}=\hat{\Omega}^0_+ = (\varepsilon^0_+ + c^2 - V)^{-1}c\sigma\cdot p$.

2.1b *Minimization*: Now, a change $\delta\varepsilon^0_+$ is created by changing $u$ to $u+\delta u$ in (2). If $\delta\varepsilon^0_+=0$ for all variations $\delta u^{\dagger}$ then $\varepsilon^0_+=\varepsilon_{0i}$, $u=u^0_i$ and $l=l^0_i$ where, apart from the normalization constant, $u^0_i$ and $l^0_i$ are the upper and lower components of $\psi^0_i$ which is one of the positive-energy eigenfunctions of $H_D$, and $\varepsilon_{0i}$ is the corresponding

eigenvalue. At this point

$$l_i^0 = c(\varepsilon_{0i} + c^2 - V)^{-1}\sigma\cdot p u_i^0$$

and

$$u_i^0 = c(\varepsilon_{0i} - c^2 - V)^{-1}\sigma\cdot p l_i^0.$$

When $\delta u^\dagger$ is not arbitrary, one usually obtains the upper component of an impure state. However, one can expand any $u$ as a linear combination of the $u_i^0$'s assuming that the latter form a complete set of upper-component functions; [$u_i^0$'s may not be orthogonal to each other, but $\psi_i^0$'s form an orthonormal set]. However, the corresponding linear combination of the $l_i^0$'s does not give the exact $l$. Still, from (2), one can show that

$$\delta_u \delta_{u^\dagger} \varepsilon_+^0 [u]|_{u=u_0^0} > 0$$

where $u_0^0$ is the upper component of the ground state. This inequality depends upon the observation that $\varepsilon_+^0 [\delta u] > \varepsilon_0$ where $\varepsilon_0$ is the ground state energy. Thus we have the variational principle

$$\min_u \varepsilon_+^0 [u] \geqslant \varepsilon_0, \quad \text{(all variations } \delta\hat{\Omega}\text{)}$$

and in particular, the minimax theorem

$$\min_u \max_{\hat{\Omega}} \langle H_D \rangle \geqslant \varepsilon_0, \quad \text{(all variations } \delta\hat{\Omega}\text{)}. \tag{3}$$

If the trial spinor

$$\psi_t = N_t \begin{pmatrix} u \\ \Lambda\sigma\cdot p u \end{pmatrix}$$

is chosen, variation of $\langle H_D \rangle$ with respect to $\Lambda$ is equivalent to a linear variation involving the basis spinors $\phi_+$ and $\phi_-$. Therefore, one obtains two stationary points—a minimum ($\varepsilon_-$) and a maximum ($\varepsilon_+$). The optimized parameter (corresponding to $\varepsilon_+$) is given by

$$\Lambda^0 = c[\varepsilon_+ + c^2 - \bar{V}]^{-1},$$

where

$$\bar{V} = (V)_{\sigma\cdot pu}$$

and

$$\varepsilon_+ = c^2 + (p^2)_u/2 + (V)_u + [(p^2)_u/4c^2][\bar{V} - (V)_u - (p^2)_u/2] + O(c^{-4}).$$

We find that $\varepsilon_+ [u]$ does not differ from $\varepsilon_+^0 [u]$ up to order $c^{-2}$, but it can be in error by an amount of order $c^{-4}$. We write, to order $c^{-2}$,

$$\max_\Lambda \langle H_D \rangle \geqslant \varepsilon_0.$$

If $\varepsilon_+ [u]$ is minimized by varying $u$, then, to order $c^{-2}$,

$$\min_u \max_\Lambda \langle H_D \rangle \geqslant \varepsilon_0. \tag{4}$$

The minimax technique creates an upper bound which can be in error by an amount of order $c^{-4}$.

For more than one upper component function, the minimax technique can be practised by writing the trial positive-energy spinors in the coherent form and preserving the mutual orthogonality of these spinors. The coherent linear variation may be viewed as the first step of the (random) linear variation using the "kinetically balanced" basis set. In the second step the (positive-energy) coherent spinors (eigenvectors from the first step) are mixed with the corresponding negative-energy vectors. The final positive-energy eigenvectors differ from the coherent solutions by coefficients of order $c^{-2}$; the final eigenvalues can differ from the eigenvalues of the first step by amounts of order $c^{-4}$ which is anyway the order of error of the same values for a given set of upper component basis functions.

## 2.2 *The Dirac-Hartree-Fock theory*

Although the Dirac-Coloumb Hamiltonian

$$H_{DC} = \sum_{k=1}^{N} H_D(k) + \sum_{k<k'} U(k,k')$$

has no normalizable eigenfunctions corresponding to bound states, it has been shown that the numerical HF calculation using the Hamiltonian $H_{DC}$ is equivalent to the HF calculations using the more correct projected Hamiltonians which can be derived within the framework of quantum electrodynamics (Sucher 1980; 1985; Mittleman 1981). Further, excellent numerical results have been obtained by using a projected Hamiltonian (Hess 1986). These observations indicate that the DHF theory can be formulated in an unambiguous way. In fact it has been suggested that a finite basis computation using kinetic balance may not suffer from catastrophic variational collapse (Mark and Schwarz 1982; Schwarz and Wallmeier 1982; Schwarz and Wechsel-Trakowski 1982). Here we give a formulation of the closed-shell DHF theory in steps representing a minimax procedure.

We use the set of orthonormal spinors:

$$\left\{ \psi_i = N_i \begin{pmatrix} u_i \\ \hat{\Omega}_i u_i \end{pmatrix}, i = 1, \ldots, N \right\}.$$

The average value calculated with the corresponding Slater determinant is:

$$\langle H_{DC} \rangle = \sum_{i=1}^{N} H_{D_{ii}} + (1/2) \sum_{i,j}^{N} (J_{ij} - K_{ij})$$

where $J_{ij}$ and $K_{ij}$ are the direct and exchange integrals respectively. At first, we consider variation of $\langle H_{DC} \rangle$ with respect to the $\hat{\Omega}_i$'s. If $\delta \langle H_{DC} \rangle = 0$ for all variations $\delta \hat{\Omega}_i^{\dagger}$ subject to the constraint that the orbitals remain orthonormal, then

$$h^l(\hat{\Omega}_i u_i) + c\sigma \cdot p u_i = c^2 \sum_{j=1}^{N} \gamma_{ij} (\hat{\Omega}_j u_j) \tag{5}$$

for $i = 1, \ldots, N$. In (5),

$$h^l = -c^2 + V + \sum_{j=1}^{N} (\hat{J}_j - \hat{K}_j).$$

and $\gamma_{ij}$'s are Lagrange's undetermined multipliers. The operators $\hat{J}_j$ and $\hat{K}_j$ are defined by

$$\hat{J}_j(1) = N_j^2(u_j(2), [U(1,2) + \hat{\Omega}_j^\dagger(2)U(1,2)\hat{\Omega}_j(2)]u_j(2)),$$
$$\hat{K}_j(1)u_i(1) = N_j^2(u_j(2), [U(1,2) + \hat{\Omega}_j^\dagger(2)U(1,2)\hat{\Omega}_i(2)]u_i(2))u_j(1),$$

and

$$\hat{K}_j(1)\hat{\Omega}_i(1)u_i(1) = N_j^2(u_j(2), [U(1,2) + \hat{\Omega}_j^\dagger(2)U(1,2)\hat{\Omega}_i(2)]u_i(2))\hat{\Omega}_j(1)u_j(1).$$

We consider a unitary transformation of the orbitals such that for the transformed operator $h^l$ the transformed matrix $\gamma$ is diagonal with elements $\gamma_{ii} = c^{-2}\varepsilon_i$ for $i = 1, 2, \ldots, N$. Then,

$$h^l(\hat{\Omega}_i^0 u_i) + c\sigma \cdot p u_i = \varepsilon_i(\hat{\Omega}_i^0 u_i) \quad (i = 1, 2, \ldots, N), \tag{6}$$

where $\hat{\Omega}_i^0$ has been written in lieu of the operator $\hat{\Omega}_i$, and $u_i$ is the upper component of the $i$th transformed orbital. Equation (6) is the HF equation in terms of the lower components, and has the solution

$$\hat{\Omega}_i^0 = c(\varepsilon_i - h^l)^{-1}\sigma \cdot p$$

for $i = 1, 2, \ldots, N$. The operator $\hat{\Omega}_i^0$ is the HF equivalent of the ideal one-electron operator $\hat{\Omega}^0$ and gives a kinetically balanced orbital. The apparent pseudoeigenvalue $\varepsilon_i[u]$ has to be determined from the equation

$$c^{-}(\eta_i, [\varepsilon_i - h^l]^{-1}\eta_i) = (u_i, [\varepsilon_i - h^u]u_i) \quad (i = 1, 2, \ldots, N), \tag{7}$$

where $\eta_i = \sigma \cdot p u_i$ and $h^u = c^2 + V + \sum_{j=1}^{N}(\hat{J}_j - \hat{K}_j)$. The operators $\hat{J}_j$ and $\hat{K}_j$ are now defined in terms of the transformed spinors. The stationary value of $\langle H_{DC} \rangle$ is a functional of the upper components, written as $E[\{u\}]$. Again, we assume that $(V)_{u_i} > -2c^2$ for every $i$. Also, $(h^u)_{u_i} > c^2 + (V)_{u_i}$. Hence for well-behaved and normalizable $u_i$'s, $\varepsilon_i > (h^u)_{u_i}$. Considering that

$$\left(V + \sum_{j(\neq i)} \{\hat{J}_j - \hat{K}_j\}\right)_{(\delta\Omega_i)u_i} < 0$$

we find:

$$\delta_{\hat{\Omega}_i}\delta_{\hat{\Omega}_i^\dagger}\langle H_{DC} \rangle|_{\hat{\Omega}_i = \hat{\Omega}_i^0} < 0$$

for every $i$. Therefore, for a fixed set of upper component functions the stationary point $E[\{u\}]$ is a local maximum, and we write

$$\max_{\{\hat{\Omega}\}} \langle H_{DC} \rangle = E[\{u\}], \quad \text{(all variations } \delta\hat{\Omega}_i,\ i = 1, \ldots, N).$$

In the second step, one works with the set of orbitals

$$\left\{\psi_i = N_i\begin{pmatrix} u_i \\ \Omega_i^0 u_i \end{pmatrix}, i = 1,\ldots,N\right\}$$

and extremizes $\langle H_{DC}\rangle$ (that is, $E$) by varying the upper component functions subject to the constraint that the orbitals remain orthonormal. If $\delta E = 0$ for all variations $\delta u_i^\dagger$, then, using kinetic balance,

$$h^u u_i^0 + c\sigma\cdot p(\hat{\Omega}_i^0 u_i^0) = \varepsilon_i^0 u_i^0 \quad (\text{for } i = 1,2,\ldots,N), \tag{8a}$$

where $\hat{\Omega}_i^0 = (\varepsilon_i^0 - h^l)^{-1} c\sigma\cdot p$ and $u_i^0$ is the optimal trial upper component function. Of course, (8a) can be directly obtained from (7) when $\delta\varepsilon_i = 0$ for all variations $\delta u_i^\dagger$, $i = 1,\ldots,N$. Now, (6) is rewritten as

$$c\sigma\cdot p u_i^0 + h^l(\hat{\Omega}_i^0 u_i^0) = \varepsilon_i^0(\hat{\Omega}_i^0 u_i^0) \quad (i = 1,2,\ldots,N). \tag{8b}$$

Equations (8a) and (8b) together represent the complete Dirac-Hartree-Fock equation for a closed-shell system. The Fock operator

$$F = H_D + \sum_j \begin{pmatrix} \hat{J}_j - \hat{K}_j & 0 \\ 0 & \hat{J}_j - \hat{K}_j \end{pmatrix}$$

has the pseudoeigenvalues $\varepsilon_i^0$ corresponding to the eigenfunctions

$$\psi_i^0 = N_i^0\begin{pmatrix} u_i^0 \\ \Omega_i^0 u_i^0 \end{pmatrix}.$$

The optimized energy is written as $E_{DHF}$ (the Dirac-Hartree-Fock value). By making $E$ stationary for all variations of $u_i$, $i = 1,\ldots,N$, we have really made $E$ stationary with respect to all variations in the *kinetically balanced Slater determinant* $\Phi$. This does not tell us whether $E$ is maximized or minimized by varying $\Phi$. Normally, however, the stationary point will be a minimum:

$$\min_{\Phi} E[\{u\}] = E_{DHF} \quad (\text{all variations } \delta\Phi).$$

But $E_{DHF} > E_0$ where $E_0$ is the correct ground state energy (minimum of the expectation value of the configuration-space Hamiltonian which was derived by Mittleman (1981)). Therefore,

$$\begin{aligned} \min_{\Phi}\max_{\{\hat{\Omega}\}} \langle H_{DC}\rangle = E_{DHF} > E_0 \quad & (\text{all variations } \delta\hat{\Omega}_i, i = 1,\ldots,N) \\ & (\text{all variations } \delta\Phi) \end{aligned} \tag{9}$$

which is the relevant minimax theorem.

The preceding analysis reveals the two-step nature of the DHF method. We now discuss the implication of using the trial operators $\hat{\Omega}_{ti}$'s. It is easily seen that (for the same set of upper component functions) orbital energies and the one- and two-electron integrals calculated with $\Lambda_i^0$'s do not differ, up to order $c^{-2}$, from the corresponding quantities obtained from the use of $\hat{\Omega}_i^0$'s. Therefore, the maximum value ($\tilde{E}$) will be correct through order $c^{-2}$. This conclusion compares well with that drawn by Stanton

and Havriliak (1984) on the HF calculation using a "kinetically balanced" basis set. We write:

$$\max_{\{\Lambda\}} \langle H_{DC} \rangle = \tilde{E}[\{u\}]. \tag{10}$$

The Slater determinant corresponding to the maximum point is written as $\tilde{\Phi}$. If $\delta\tilde{E} = 0$ for all variations $\delta u_i^{\dagger}$, $i = 1, \ldots, N$, then

$$\min_{\tilde{\Phi}} \max_{\{\Lambda\}} \langle H_{DC} \rangle = E_{\Lambda HF}, \tag{11}$$

where $(E_{\Lambda HF} - E_{DHF})$ is at most of order $c^{-4}$.

## 3. The relativistic virial theorem

The non-relativistic HF treatment yields the virial theorem when the expectation value of the Hamiltonian is extremized through variation with respect to a scale factor. For example, the Eckart-Kellner function for the He atom ($Z = 2$) gives the virial relationship if $\partial \langle H \rangle / \partial \zeta = 0$ where $\zeta$ is the exponent of the 1s orbital. The optimum value of $\zeta$ is $(Z - 5/16)$. In the relativistic case, scaling of an orbital involves the scaling of not only the upper component but also the lower one. Therefore, within the one-electron approximation, the relativistic virial theorem may not be obtained from arguments based on the scaling of the upper component functions only.

Rosicky and Mark (1975) presented the derivation of the relativistic virial theorem for one electron. Later, we have investigated conditions for the validity of the generalized relativistic virial theorem for the electron in a central field (Datta 1984). Here, we show that the derivation of the virial theorem (the hypervirial theorem in the molecular case) within the framework of the DHF theory requires that $\langle H_{DC} \rangle$ is also stationary for variation of the scale factor in $\hat{\Omega}_i$, $i = 1, \ldots, N$. Thus the derivation of the relativistic virial theorem for electrons can serve as an example of the minimax technique.

A scaling $u_i(\mathbf{r}) \rightarrow \lambda^{3/2} u_i(\lambda \mathbf{r})$ is performed for all $i$ and as suggested by Fock (1978) we require that

$$\partial \langle H_{DC} \rangle_{\{u(\lambda)\}} / \partial \lambda = 0$$

at $\lambda = 1$. This gives

$$\langle T \rangle + \langle V_T \rangle + \sum_{n}^{\text{nuclei}} R_n \cdot \langle \nabla_n V_T \rangle + \sum_i \langle \psi_i | \psi_i \rangle^{-1}$$
$$(u_i, [\{\partial \hat{\Omega}_i^{\dagger}(1/\lambda)/\partial \lambda\}_{\lambda=1} \{c\sigma \cdot p - (F_{ii} - h^I)\hat{\Omega}_i\} + \text{h.c.}] u_i) = 0 \tag{12}$$

where

$$T = c \sum_{k=1}^{N} \alpha_k \cdot p_k,$$

$$V_T = \sum_{k=1}^{N} V(k) + \sum_{k<k'}^{N} U(k, k') + V_{nn}(\{R\})$$

and $F_{ii}$ is the diagonal element of the Fock matrix. The third term on the left hand side

of (12) is absent for an atomic system and appears for a molecular system when the latter is treated within Born-Oppenheimer approximation; $\{R\}$ represents the set of nuclear coordinates and $V_{nn}(\{R\})$ is the potential energy of internuclear repulsion. From the form of (12) it is obvious that the relativistic (molecular) hypervirial theorem

$$\langle T\rangle + \langle V_T\rangle + \sum_n R_n\cdot\langle \nabla_n V_T\rangle = 0 \tag{13}$$

is obtained if $\hat{\Omega}_i = \hat{\Omega}_i^0$ for every $i, i = 1, 2, \ldots, N$. For a specific form of the $\hat{\Omega}_i$'s, the fourth term in the left hand side of (12) vanishes if $\delta_{\hat{\Omega}_i^\dagger}\langle H_{DC}\rangle = 0$, $i = 1, \ldots, N$, while the variation $\delta\hat{\Omega}_i^\dagger$ is represented by the change $-\delta\lambda[\partial\hat{\Omega}_i^\dagger(1/\lambda)/\partial\lambda]_{\lambda=1}$. Thus the variation of $\langle H_{DC}\rangle_{\{u(\lambda)\}}$ with respect to the scale factor $\lambda$ need not necessarily give (13). However, if

$$[\partial\langle H_{DC}\rangle_{\hat{\Omega}_i(1/\lambda)}/\partial\lambda]_{\lambda=1} = 0 \quad \text{for } i = 1, \ldots, N,$$

then the condition

$$[\partial\langle H_{DC}\rangle_{\{u_i(\lambda)\}}/\partial\lambda]_{\lambda=1} = 0$$

is sufficient for the validity of (13).

If $\delta\langle H_{DC}\rangle = 0$ for all variations $\delta\hat{\Omega}_i^\dagger$, $i = 1, \ldots, N$, then $\hat{\Omega}_i = \hat{\Omega}_i^0$ for every $i$ and (12) reduces to (13). If $\hat{\Omega}_i = \hat{\Omega}_{ti} = \Lambda_i\sigma\cdot p$ then (13) will be obtained [from (12)] if $\langle H_{DC}\rangle$ is stationary for variation of the $\Lambda_i$'s. Actually, one can directly derive (13) from variation of $\langle H_{DC}\rangle_{\{u(\lambda)\}}$ with respect to $\lambda$. The trick is to assume that $\Lambda_i = \lambda^{-1}\Lambda_{0i}$ where $\Lambda_{0i}$ is independent of $\lambda$ such that $\partial\hat{\Omega}_{ti}(1/\lambda)/\partial\lambda = 0$. However, the optimal trial function is not obtained unless $\langle H_{DC}\rangle$ is extremised by varying the $\Lambda_{0i}$'s.

The above discussion can be easily exemplified. We use the $1s$ STO (with exponent $\zeta$) and form the upper component function of the $1s_{\pm 1/2}$ spinors of a two-electron atom (nuclear charge $= Z|e|$).

Case (i). $\hat{\Omega}_{1s} = \Lambda\sigma\cdot p$.

In this case the variation of $\langle H_{DC}\rangle$ with respect to $\Lambda$ gives the maximum:

$$\tilde{E} = 2c^2(1 + c^{-2}\zeta^2)^{1/2} - 2(Z - 5/16)\zeta,$$

and the optimized parameter:

$$\Lambda^0 = -c\zeta^{-2} + c\zeta^{-2}(1 + c^{-2}\zeta^2)^{1/2}.$$

At this point,

$$\langle T\rangle + \langle V_T\rangle = 2\zeta^2(1 + c^{-2}\zeta^2)^{-1/2} - 2(Z - 5/16)\zeta,$$

which vanishes if

$$\zeta = (Z - 5/16)/\gamma,$$

where

$$\gamma = [1 - c^{-2}(Z - 5/16)^2]^{1/2}.$$

In fact the virial theorem is obtained when $\tilde{E}$ is minimized by varying $\zeta$:

$$\min \tilde{E} = 2c^2\gamma.$$

However, the direct minimization of $\langle H_{DC}\rangle$ by varying $\zeta$ does not give $\langle T\rangle + \langle V_T\rangle = 0$ unless $\Lambda = \Lambda^0$.

Case (ii). $\hat{\Omega}_{1s} = \Lambda_0 \zeta^{-1} \sigma \cdot p$.

This is essentially the previous choice, and yields the compact expression

$$\langle H_{DC}\rangle = 2\{c^2\gamma_1 + \zeta(Z_{eff} - Z)\} + 5\zeta/8$$

where

$$\gamma_1 = (1 - \Lambda_0^2)/(1 + \Lambda_0^2)$$

and

$$Z_{eff} = 2c\Lambda_0/(1 + \Lambda_0^2).$$

The virial theorem is obtained with $Z_{eff} = Z - 5/16$ *irrespective of* the value of $\zeta$. Indeed, this result is obtained from $\partial\langle H_{DC}\rangle/\partial\zeta = 0$. However, a unique value of $\zeta$ is not obtained unless $\langle H_{DC}\rangle$ is stationary with respect to $\Lambda_0$:

$$\partial\langle H_{DC}\rangle/\partial\Lambda_0 = 0 \Rightarrow \Lambda_0 = \zeta\Lambda^0$$

such that $\zeta = Z_{eff}/\gamma$.

If the Hellman-Feynman theorem holds, the hypervirial theorem reduces to

$$\langle T\rangle + \langle V_T\rangle + \sum_n R_n \cdot \nabla_n E_{tot} = 0, \tag{14}$$

which is the relativistic virial theorem in the molecular case. $E_{tot}$ is the calculated total energy, $\langle H_{DC}\rangle + V_{nn}$. If the set of trial functions is invariant to rotation, translation and inversion of the electrons, $E_{tot}$ can depend only on the bond lengths $\mathscr{R}$ and bond angles $\theta$. Finally, one can rewrite (14) as (Epstein 1974)

$$\langle T\rangle + \langle V_T\rangle + \sum_{\mathscr{R}} \mathscr{R}(dE_{tot}/d\mathscr{R}) = 0.$$

At the minimum point of the potential energy surface,

$$\langle T\rangle + \langle V_T\rangle = 0$$

and

$$E_{tot} = c^2 \sum_{i=1}^{N} \langle\beta\rangle_{ii}.$$

The virial theorem is traditionally used for checking the quality of wavefunctions. From arguments based on quantum electrodynamics, Mittleman (1981) derived the configuration-space Hamiltonian

$$H_M = \sum_{k=1}^{N} m_+(k) H_D(k) m_+(k) + \sum_{k<k'}^{N} m_+(k) m_+(k') U(k,k') m_+(k) m_+(k'),$$

where $m_+$ is the projection operator onto the subspace $S_F$ spanned by the positive-energy eigenfunctions of the Fock operator. The Mittleman Hamiltonian clearly describes the DHF calculation. But somehow one has to bring trial spinors inside the subspace $S_F$ (Datta 1980), a task which is exactly performed when $\langle H_{DC}\rangle$ is made

stationary for variation of all occupied orbitals, that is, when $\delta\langle H_{DC}\rangle = 0$ for all variations $\delta\psi_i, i = 1, \ldots, N$, subject to the orthogonality constraint. However, if $\langle H_{DC}\rangle$ is maximized by varying $\hat{\Omega}_i$ for every $i$, then the relativistic orbitals are brought almost in $S_F$, that is, the portions lying outside $S_F$ can be represented by coefficients (of linear expansion) which are at most of order $c^{-2}$. This inference supports Kim's suggestion (Kim 1967) that the virial theorem value may be used in finding the correct solution of the Dirac-Hartree-Fock-Roothaan treatment.

## 4. Results and discussion

### 4.1 *Illustration of the minimax technique*

The minimax technique can be easily illustrated by the results of a single-zeta calculation on $H_2^+$. We have used the non-relativistic 1s STO's (with exponent $\zeta$) to write the upper component function of the $1s\sigma$ molecular spinor as

$$u = [2(1 + S_{ab})]^{-1/2}\begin{pmatrix} 1s_a + 1s_b \\ 0 \end{pmatrix}, \tag{16a}$$

where $S_{ab}$ is the overlap integral, $\int d^3r\, 1s_a(\mathbf{r})\, 1s_b(\mathbf{r})$. We choose

$$\hat{\Omega} = \Lambda\sigma\cdot p \equiv \zeta^{-1}\Lambda_0\sigma\cdot p \tag{16b}$$

where

$$\Lambda_0 = [(1-\gamma)/(1+\gamma)]^{1/2}$$

and

$$\gamma = (1 - c^{-2}\kappa^{-2}Z^2)^{1/2}.$$

With this spinor, the previously obtained atomic results (Datta 1984) are reproduced in the united atom and the separated atom limits.

The electronic energy of $H_2^+$ is calculated using the trial spinor. Partial variation of $\langle H_D\rangle$ with respect to $\kappa$ leads to the maximum value $\varepsilon_+$. As shown in figure 1, the trial state for each $\zeta$ collapses on both sides of the stationary point and variation of $\varepsilon_+(\zeta)$ with respect to $\zeta$ leads to a saddle point. For $R = 2{\cdot}0$ a.u. the saddle point corresponds to $\zeta = 1{\cdot}23873$, $\kappa = 0{\cdot}80731$ and $\varepsilon_s = -1{\cdot}08651435$ a.u. The difference between $\varepsilon_s$ and the (nonrelativistic) LCAO energy ($\varepsilon_{nrel}$) equals $-0{\cdot}835 \times 10^{-5}$ a.u. The accurate relativistic correction is $-0{\cdot}736 \times 10^{-5}$ a.u. (Laaksonen and Grant 1984; Mark and Schwarz 1982). This shows that a minimal basis calculation may not be sufficient for computing relativistic corrections in the molecular case. To see the advantage of the minimax technique using $\hat{\Omega}_t$ over the direct minimization of energy while the fixed operator $\hat{\Omega}_a$ is in use, we note that

$$\varepsilon_+ - \varepsilon_+^0 = -(1/8)c^{-4}(u,u)^{-1}(\sigma\cdot pu, \{V - \bar{V}\}^2\sigma\cdot pu) + O(c^{-6}).$$

Since $\varepsilon_+$ is the maximum obtained from variation of $\langle H_D\rangle$ with respect to $\Lambda, \Lambda^0$ gives the closest approximation to the kinetically balanced spinor. The bounds obtainable from $\Lambda = \Lambda^0$ and $\Lambda = (2c)^{-1}$ differ from each other by an amount that increases very rapidly as $Z$ increases. The trial 1s upper component function leads to the correct ground state energy of a one-electron atom when $\Lambda = \Lambda^0$, whereas for $\Lambda = (2c)^{-1}$ the "bound" ($\varepsilon_a$) is in error by the amount—$(1/32)c^{-4}Z^6 + \cdots$. For $Z = 68$, this error is

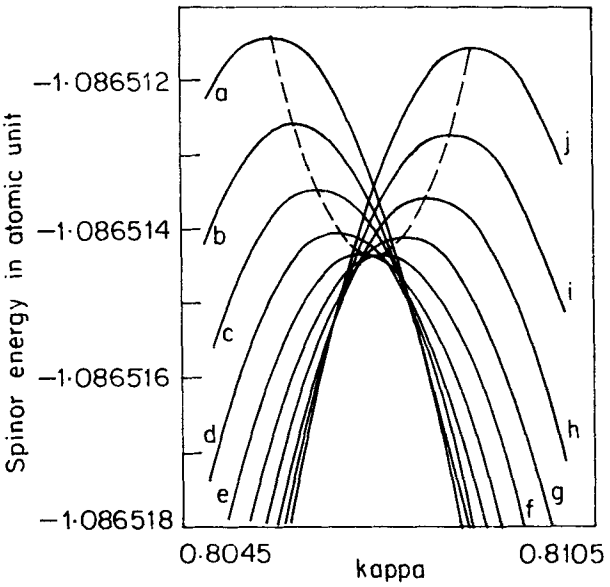


**Figure 1.** Energy of the $1s\sigma$ spinor versus $\kappa$: the $H_2^+$ molecule ion at $R = 2{\cdot}0$ a.u.; $\zeta =$ (**a**) 1·2410, (**b**) 1·2405, (**c**) 1·2400, (**d**) 1·2395, (**e**) 1·2390, (**f**) 1·2385, (**g**) 1·2380, (**h**) 1·2375, (**i**) 1·2370, (**j**) 1·2365.

about 6% of the relativistic correction to the $1s$ orbital energy. In order to reduce the error, one must selectively mix other basis functions and discard the spurious solutions. Thus for systems which are considerably relativistic, the use of $\Lambda^0$ will give better economy of basis functions. The same argument holds for the molecular spinors. A test calculation of the $1s\sigma$ orbital energy of the (hypothetical) molecular ion $K_2^{+37}$ at $R = 2{\cdot}0$ a.u. yields $\varepsilon_s - \varepsilon_a = 0{\cdot}004335$ a.u. (figure 2).

To see the immediate effect of the minimax technique on the orbital energies and total energy obtainable from a minimal basis calculation on a many-electron molecule, we choose the STO-3G basis and $\hat{\Omega} = \Lambda\sigma \cdot p$ for the relativistic $1s\sigma$ orbitals of the $H_2$ molecule. A scale parameter $\lambda^2$ is included in the exponents of the gaussian primitives such that the effective exponent ($\zeta$) of the STO can be treated as a variational parameter. We have maximized $\langle H_{DC} \rangle$ by varying $\Lambda$ while $\lambda$ is kept fixed, and minimized the maximum value $\tilde{E}(\lambda)$ by varying the scale factor ($\lambda$). The optimized parameter $\Lambda_{1s\sigma}$ for $H_2$ is slightly greater than that for $H_2^+$ at the same bond length (1·4 a.u.). Even though we have used only two parameters ($\Lambda$ and $\lambda$) the difference ($E_{rel} - E_{nrel}$) turns out to be rather good; the calculated value accounts for 90% of the accurate value (table 1). However, ($\varepsilon_{rel} - \varepsilon_{nrel}$) sharply differs from the corresponding quantity computed by the numerical integration technique (Laaksonen and Grant 1984). This difficulty is traced to the lack of the correct $r \to 0$ behaviour of the basis spinors chosen by us. One can always improve the energy values ($\varepsilon$ and $E$) by using a larger basis; however, the difference ($\varepsilon_{rel} - \varepsilon_{nrel}$) may not improve, and can actually worsen, unless enlarging the basis improves the quality of the molecular spinor near the nuclei (table 1).

### 4.2 *The coherent linear variation*

Here we have chosen the He atom as the numerical test case. If total energy is calculated with the basis spinors $1s_{1/2}$ and $1s_{-1/2}$ which are built of the $1s$ STO and $\hat{\Omega}_\eta$, the

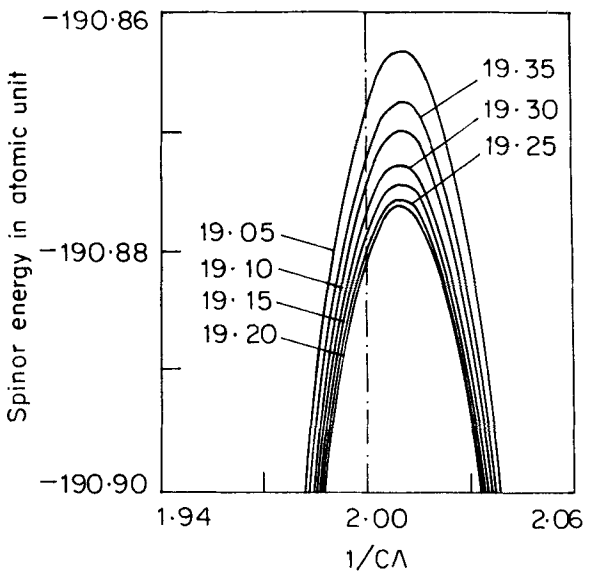


**Figure 2.** Energy of the $1s\sigma$ spinor versus $(c\Lambda)^{-1}$: $Z = 19$ and $R = 2{\cdot}0$ a.u. Values of the exponent are indicated. Optimum values: $\zeta = 19{\cdot}1853$ and $c\Lambda = 0{\cdot}497574$.

**Table 1.** Dirac-Hartree-Fock results for the ground state of $H_2$ at $R = 1{\cdot}4$ a.u.

| Property | Method | Value relativistic | Value non-relativistic | Difference |
|---|---|---|---|---|
| Total energy (a.u.) | (a) Present work | −1·119206 | −1·119193 | $-1{\cdot}3 \times 10^{-5}$ |
| | (b) DHFR | −1·129791 | −1·129777 | $-1{\cdot}3 \times 10^{-5}$ |
| | (c) Numerical | −1·133644 | −1·133630 | $-1{\cdot}44 \times 10^{-5}$ |
| Orbital energy (a.u.) | (a) Present work | −0·590271 | −0·590284 | $+1{\cdot}3 \times 10^{-5}$ |
| | (b) DHFR | −0·59418 | −0·59458 | |
| | (c) Numerical | −0·594662 | −0·594658 | $-0{\cdot}4 \times 10^{-5}$ |

(a) This work (minimal basis calculations using the STO-3G basis set): $\zeta$ (non-relativistic) = 1·19015; $\zeta$ (relativistic) = 1·19034 and $2c\Lambda = 0{\cdot}99999$.
(b) Dirac-Hartree-Fock-Roothaan calculation by Matsuoka *et al* (1980).
(c) Laaksonen and Grant (1984).

minimax technique yields the optimum parameters $\Lambda^0 = 0{\cdot}499981c^{-1}$ and $\zeta = 1{\cdot}68763$ (see § 3). For the non-relativistic calculation, the optimum value of $\zeta$ is 1·6875. The calculated value, $(E_{rel} - E_{nrel}) = -10{\cdot}8 \times 10^{-5}$ a.u., is in qualitative agreement with the DHF value, $-12{\cdot}5 \times 10^{-5}$ a.u. Again, the lack of the appropriate boundary condition (for $r \to 0$) gives a totally wrong estimate of $(\varepsilon_{rel} - \varepsilon_{nrel})$: calculated value is $-1{\cdot}4 \times 10^{-5}$ a.u. whereas the DHF value is $-3{\cdot}4 \times 10^{-5}$ a.u.

The effect of increasing the basis size is investigated with $\zeta = 2{\cdot}0$. The upper component basis functions are built up from the 1*s*, 2*s*, 3*s* and 4*s* STO's. We keep $\Lambda$ fixed at the optimum value for the 1*s* basis function ($c\Lambda_{1s} = 0{\cdot}4999734$) and minimize $\langle H_{DC} \rangle$

by varying the coefficients $c_i$'s. In principle, as the coefficients vary, the optimum value of $\Lambda$ should change. Therefore, one expects that the coherent linear variation will yield different coefficients along with an optimized $\Lambda$ different from $\Lambda_{1s}$. In reality, for any basis size, the minimal basis value ($\Lambda_{MB}$) is very close to the optimum value of $\Lambda$ as long as $Z/c \ll 1$. Therefore, fixation of $\Lambda$ at $\Lambda_{MB}$ is effectively equivalent to the optimization of $\Lambda$ when $Z/c$ is small. In the present case we have used $\Lambda_{coh} = \Lambda_{1s}$. For atoms with more electrons (but still in the nonrelativistic limit), one may assign: $\Lambda_{coh,\,1s} = \Lambda_{1s}$, $\Lambda_{coh,\,2s} = \Lambda_{2s}$, etc. The main point is that a judicious choice of $\Lambda$ can save a lot of computational efforts. A better way (which is applicable to all systems) will be to design a set of coherent basis spinors and use these in the linear expansion technique.

At present we make three observations. First, the difference between the result of linear variation using $\Lambda_{1s}$ for the coherent spinors and that obtained from a (random) linear variation using the "kinetically balanced" basis set is negligible (table 2). Secondly, the calculated wavefunctions, both relativistic and nonrelativistic, improve with the increase in basis size; this is evidenced by the trends in orbital energies and total energy (table 3). Thirdly, the linear expansion technique quickly improves the difference ($\varepsilon_{rel} - \varepsilon_{nrel}$) to a value ($-3{\cdot}1 \times 10^{-5}$ a.u.) which is comparable to (but again greater than) the corresponding quantity ($-3{\cdot}4 \times 10^{-5}$ a.u.) found from the numerical integration technique. However, unless the relativistic wavefunction is correctly described near the nucleus, ($E_{rel} - E_{nrel}$) converges to a value ($-13{\cdot}8 \times 10^{-5}$ a.u. for $\zeta = 2{\cdot}0$; $-10{\cdot}1 \times 10^{-5}$ a.u. for $\zeta = 1{\cdot}6875$) that can be quite different from the result ($-12{\cdot}5 \times 10^{-5}$ a.u.) of the numerical DHF calculation.

### 4.3 *Reproducibility of the nonrelativistic results*

The nonrelativistic molecular results have been verified by using $c = 137{,}032{\cdot}5$ a.u. (instead of 137·0325 a.u.) in the relativistic calculations; (the optimum value of $\zeta_{nrel}$ differs from that of $\zeta_{rel}$).

### 4.4 *Limitations of the present work*

The calculations reported here utilize very small basis sets and serve as test calculations only. The main import of these exemplary computations is that the smallest matrix

**Table 2.** Comparison of the result of the coherent linear variation with that of the random linear variation (for $Z = 2$).

| Basis | $(E_{coherent} - E_{random})$ in $10^{-9}$ a.u. | $(\varepsilon_{1s}^{coherent} - \varepsilon_{1s}^{random})$ in $10^{-9}$ a.u. |
|---|---|---|
| $\zeta = 2{\cdot}0,\ 2c\Lambda = 0{\cdot}9999468$ | | |
| $1s$ | 0·00 | 0·00 |
| $1s, 2s$ | −1·61 | 1·11 |
| $1s, 2s, 3s$ | −2·30 | 1·84 |
| $1s, 2s, 3s, 4s$ | −2·64 | −0·76 |
| $\zeta = 1{\cdot}6875,\ 2c\Lambda = 0{\cdot}999962$ | | |
| $1s$ | 0·00 | 0·00 |
| $1s, 2s$ | −0·67 | 9·38 |
| $1s, 2s, 3s$ | −1·01 | 8·05 |
| $1s, 2s, 3s, 4s$ | −1·19 | 7·32 |

**Table 3.** Results of the Coherent linear variation employing STO's in the upper components of basis spinors: the He atom.

| Basis | $(\varepsilon_{rel}^{1s} - \varepsilon_{nrel}^{1s})$ in $10^{-5}$ a.u. | $(E_{rel} - E_{nrel})$ in $10^{-5}$ a.u. | $\varepsilon_{rel}^{1s}$ in a.u. | $E_{rel}$ in a.u. |
|---|---|---|---|---|
| $\zeta = 2{\cdot}0, 2c\Lambda = 0{\cdot}9999467$(a) | | | | |
| $1s$ | −10·65 | −21·30 | −0·750107 | −2·750213 |
| $1s, 2s$ | −3.57 | −14·93 | −0·903858 | −2·842936 |
| $1s, 2s, 3s$ | −3·27 | −14·17 | −0·911623 | −2·851673 |
| $1s, 2s, 3s, 4s$ | −3·16 | −13·85 | −0·913932 | −2·854690 |
| $\zeta = 1{\cdot}6875$, $2c\Lambda = 0{\cdot}999962$ (a) | | | | |
| $1s$ | −5·40 | −10·80 | −0·896538 | −2·847764 |
| $1s, 2s$ | −3·43 | −10·17 | −0·914309 | −2·851017 |
| $1s, 2s, 3s$ | −3·13 | −10·11 | −0·915221 | −2·851104 |
| $1s, 2s, 3s, 4s$ | −3·00 | −10·10 | −0·915264 | −2·851105 |
| Numerical HF and DHF (b) | −3·4 | −6·5 | −0·917991 | −2·86175 |

a. $c = 137{\cdot}0325$ a.u.; b. Desclaux (1973).: $E_{rel}$ includes the energy of magnetic interaction ($6 \times 10^{-5}$ a.u.).

eigenvalue approaches the exact value from above. This suggests that a set of coherent basis spinors can be used without fear of variational collapse.

The following limitations of the present work are identified:

(i) The trial spinors chosen by us do not give an accurate description of the relativistic wavefunction near the nucleus. The situation can be rectified by using the multiple-zeta basis (Lee and McLean 1982) or the non-integral STO's (Pavlik and Blinder 1967; Datta and Ewig 1982; Malli 1984), or even by considering a finite size of the nucleus (Ishikawa 1986).

(ii) We have not illustrated the minimax technique for the excited state. It is well known that the influence of enlarging the space of linear variation upon the distribution of eigenvalues can be described by the separation theorem (Epstein 1974). The minimax technique can be considered as the first step towards obtaining the separation theorem. For calculations based on the Dirac equation in the algebraic approximation the separation theorem has been discussed by Wood *et al* (1985) and by Grant (1986).

(iii) Unless the wavefunction is the exact DHF solution, positive-energy eigenvectors of the Fock matrix may not be completely free from the negative-energy disease. For instance, an arbitrary spinor

$$\psi = N\begin{pmatrix} u \\ \hat{\Omega}u \end{pmatrix}$$

can be expanded as

$$\psi = \sum_i d_i \psi_i^0 + \sum_i d_i N_i^0 \begin{pmatrix} 0 \\ \{\hat{\Omega} - \hat{\Omega}_i^0\} u_i^0 \end{pmatrix}$$

where $\psi_i^0$'s are the positive-energy eigenfunctions of the Fock operator. It is obvious that the residual part vanishes only if $\psi$ is identical with one of the

$\psi_i^0$'s. Therefore, just because the virial theorem holds one cannot say with certainty that the calculated relativistic orbitals lie inside $S_F$. It is equally obvious that the deviation from $S_F$ can be reduced by (a) relying on the appropriate boundary conditions, (b) enlarging the basis and (c) using a kinetically balanced basis set consisting of the coherent basis spinors and the associated negative-energy vectors.

## Acknowledgements

Financial support by CSIR and DST is gratefully acknowledged. Thanks are due to D Mukherjee who had suggested in 1983 the interpretation of our data in terms of the minimax technique, to S H Patil, N F Stepanov and S Koneru for valuable discussion and to J Thakur who has critically read the manuscript and offered many suggestions. A part of the work was orally presented at the Adriatico Research Conferernce on relativistic many-body problems, and SND thanks the organizers, DST and IIT Bombay.

## References

Datta S N 1980 *Chem. Phys. Lett.* **74** 568
Datta S N 1984 *Pramana – J. Phys.* **23** L275
Datta S N 1987 *Pramana – J. Phys.* **28** 633
Datta S N and Devaiah G 1986 in: Abstract of the Conference on relativistic many-body problems, Trieste (Italy)
Datta S N and Ewig C S 1982 *Chem. Phys. Lett.* **85** 443
Datta S N and Jagannathan S 1984 *Pramana – J. Phys.* **23** 467
Desclaux J P 1973 *At. Data Nucl. Data Tables* **12** 311
Drake G W F and Goldman S P 1981 *Phys. Rev.* **A23** 2093
Dyall K G, Grant I P and Wilson S 1984 *J. Phys.* **B17** L45, 439
Epstein S T 1974 *The variation method in quantum chemistry* (New York: Academic Press)
Fock V A 1978 *Fundamentals of quantum mechanics* (Moscow: Mir Publishers)
Gazdy B 1983 *Chem. Phys. Lett.* **99** 41
Gazdy B and Ladanyi K 1984 *J. Chem. Phys.* **80** 4333
Goldman S P 1985 *Phys. Rev.* **A31** 3541
Grant I P 1986 *J. Phys.* **B19** 3187
Hegarty D 1986 *Theor. Chim. Acta* **70** 351
Hess B A 1986 *Phys. Rev.* **A33** 3742
Ishikawa Y 1986 in: Abstract of the Conference on relativistic many-body problems, Trieste (Italy)
Ishikawa Y, Binning R C and Sando K M 1983 *Chem. Phys. Lett.* **101** 111
Ketley I J and Moss R E 1983 *Mol. Phys.* **48** 1131
Kim Y K 1967 *Phys. Rev.* **154** 17
Kutzelnigg W 1984 *Int. J. Quantum Chem.* **25** 107
Laaksonen L and Grant I P 1984a *Chem. Phys. Lett.* **109** 485
Laaksonen L and Grant I P 1984b *Chem. Phys. Lett.* **112** 157
Ladik J, Cizek J and Mukherjee P K 1983 *Relativistic effects in atoms, molecules and solids* (ed.) G Malli (New York: Plenum Press) p. 305
Lee Y S and Mclean A D 1982 *J. Chem. Phys.* **76** 735
Malli G 1984 *J. Chem. Phys.* **80** 2060
Mark F and Schwarz W H E 1982 *Phys. Rev. Lett.* **48** 673
Matsuoka O, Suzuki N, Aoyama T and Malli G 1980 *J. Chem. Phys.* **73** 535
Mittleman M H 1981 *Phys. Rev.* **A24** 1167
Morrison J D and Moss R E 1980 *Mol. Phys.* **41** 491

Oreg J and Malli G 1976 *J. Chem. Phys.* **65** 1663
Pavlik P I and Blinder S M 1967 *J. Chem. Phys.* **46** 2749
Rosicky F and Mark F 1975 *J. Phys.* **B8** 2581
Rutkowski A 1986 in: Abstract of the Conference on relativistic many-body problems, Trieste (Italy)
Schwarz W H E 1987 *Phys. Scr.* **36** 403
Schwarz W H E and Wallmeier H 1982 *Mol. Phys.* **46** 1045
Schwarz W H E and Wechsel-Trakowski E 1982 *Chem. Phys. Lett.* **85** 94
Sepp W D and Fricke B 1985 *AIP Conf. Proc.* **136** 20
Stanton R E and Havriliak S 1984 *J. Chem. Phys.* **81** 1910
Sucher J 1980 *Phys. Rev.* **A22** 348
Sucher J 1985 *AIP Conf. Proc.* **136** 1
Wood J, Grant I P and Wilson S 1985 *J. Phys.* **B18** 3027