

r_{12} -Dependent terms in the wave function as closed sums of partial wave amplitudes for large l

Werner Kutzelnigg

Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-4630 Bochum,
Federal Republic of Germany

(Received May 4/ Accepted May 17, 1985)

The ansatz $\Psi = (1 + \frac{1}{2}r_{12})\Phi + \chi$ with Φ the bare nuclear (or screened nuclear) wave function and χ expanded in products of one-electron functions is explored for second-order perturbation theory and for variational calculations of the ground state of Helium-like ions.

The energy increments $\tilde{E}_l^{(2)}$ corresponding to the partial wave expansion of χ go asymptotically as l^{-8} , while conventional partial wave increments go as l^{-4} . χ is coupled to Φ by a "residual" interaction U_{12} that has no singularity for $r_{12} = 0$. With the present ansatz it is sufficient to include l -values up to 5 in order to get the second-order energy accurate to one microhartree. For the same accuracy $l \leq 4$ is sufficient in a "CI with correlated reference function" while in conventional CI one must go to $l \sim 50$. The surprisingly faster convergence of the variational approach as compared to second-order perturbation theory is explained. The slow convergence of the traditional partial wave expansion is entirely due to the attempt to represent the quantity $1 = \langle \Phi | r_{12} r_{12}^{-1} | \Phi \rangle$ by its partial wave expansion. The best reference function Φ shows very little shielding and resembles closely the eigenstate of the bare nuclear Hamiltonian. The generalization to arbitrary systems is discussed and it is pointed out that the calculation of "difficult" integrals can be avoided without a significant loss in accuracy.

Key words: r_{12} -Dependent wave functions — Convergence of the CI-expansion — $(1/Z)$ -Expansion — Correlation cusp — He-isoelectronic series — Hylleraas-CI

1. Introduction

Configuration interaction (CI) is the most straightforward and most generally applicable method for an approximate solution of the n -electron Schrödinger-equation [1]. It is, “in principle”, exact, i.e. for a complete [2] basis there is convergence to the exact (nonrelativistic) wave function and energy. With a relatively small number of configurations one often recovers a large part of the correlation energy (say 80–90%). The convergence becomes rather slow if one strives at “chemical accuracy” – errors smaller than about 1 millihartree (mh) – such that for systems of medium size a huge number of configurations has to be included. “Spectroscopic accuracy” (errors of the order of one μ h) is strictly inaccessible to CI calculations, even for the He ground state the estimated number of configurations to get this accuracy is roughly 10 000, while the best published CI energy for He is in error by more than 20 μ h. (The difficulty to get accurate CI results for He does not come from the large dimension of the secular equation – much higher dimensions can easily be handled – but from numerical problems, essentially near linear dependencies that are unavoidable for very large basis sets).

One can reduce the length of a CI expansion by using natural orbitals or pair natural orbitals [3] (pseudonatural orbitals [4]) and attempts on these lines have been rather successful [3–6], but they do not eliminate the eventually slow convergence of all CI-expansions [7].

There are two sources of the slow convergence of the CI expansion.

- (a) The “combinatorial problem”. For an n -electron system and a basis of m spinfree one-electron functions the number of Slater determinants in a “full CI” goes as $\binom{2m}{n}$, which is a very large number unless both n and m are quite small.
- (b) The slow convergence of “the partial wave expansion” even for a two electron atom. If one truncates the one-electron basis at some angular quantum number l , then the error is of the order 1^{-3} [7–9].

It is by now more or less understood how to get along with the combinatorial problem. One should not perform a “brute force” full CI, but rather start from an optimum one-determinant (or few determinant) wave function and then apply CI limited to double (or possibly triple) substitutions, or preferentially use the coupled-cluster approach or one of its simplifications such as the coupled electron pair approximation (CEPA) [5].

We now propose a simple method to overcome the slow convergence of the partial wave expansion. The basic idea is that the partial wave expansion should have some regular pattern as function of l , at least asymptotically for large l [8] and that a closed summation of this asymptotic series should be possible. One finds that the behavior for large l is closely related to that for small r_{12} and that the slow convergence of the l -expansion is due to the attempt to represent the correlation cusp [10] as the limit of the sequence of functions that only contain even powers of r_{12} .

If one wants to improve the CI expansion in this sense one is inevitably led to the inclusion of linear r_{12} -dependent terms in the wave function. Having realized

this one may wonder whether one has gained anything, since r_{12} -dependent wave functions have been used for a long time. There seems to be general agreement that one can get very accurate wave functions and energies if one includes explicitly the interelectronic coordinates, but that one has then to deal with so complicated integrals, that applications are limited to systems with a very small number of electrons.

The main message of this paper is that one can avoid “difficult” integrals and get very good results nevertheless.

r_{12} -Dependent wave functions have so far been used mainly on three lines. (a) Very sophisticated r_{12} -dependent variational expressions in the spirit of Hylleraas [11], Kinoshita [12] and Pekeris [13] for systems with two or at most three electrons [14] which led to extremely accurate results, (b) the so-called Hylleraas-CI [15, 16] or superposition of correlated configuration (SCC) method [17], in which configuration state functions of traditional CI are multiplied by factors $(r_{ij})^\nu$. There is no *a priori* limitation to systems with two or three electrons, but the published (and quite accurate) results refer to systems with at most 4 electrons, (c) the use of so-called Gaussian geminals [18], in which the factors $(r_{ij})^\nu$ are replaced by $\exp[-\alpha(\vec{r}_i - \vec{r}_j)^2]$. All the necessary integrals can then be evaluated in closed form, but it is at least not obvious that Gaussian geminals are particularly appropriate for the description of the correlation cusp. Recent numerical results with carefully optimized nonlinear parameters have been quite encouraging [19].

The ansatz proposed in this paper is related to the Hylleraas-CI method, but it is presented in a different philosophy. We want to keep the spirit and the logistic of standard CI as long as possible. We want to simplify CI (i.e. get better accuracy for smaller expansion lengths) and not to complicate it. This paper will be organized as follows. In Sect. 2 the general theoretical background is outlined, in Sect. 3 the special ansatz to be used in this paper is presented in more detail. The application to second-order perturbation theory of the He isoelectronic series is described in Sect. 4, while Sect. 5 is devoted to the variational calculation of the He ground state and its isoelectronic ions. In Sect. 6 we discuss possible improvements and the generalization to larger systems. The numerical calculations of Sect. 4 and 5 have been performed on a basis of Slater type orbitals. We have also implemented our method on a Gaussian basis for molecular calculations. The results will be presented in a forthcoming paper.

The main interest of this paper is not in the numerical results but in their analysis and in the conclusions drawn from this analysis. Some items of the analysis can be traced back to previous work [7-10, 15, 20-23] but no previous analysis led to so simple and clearcut conclusions.

2. Theoretical background

The key to our approaches is the correlation cusp relation [10]

$$\lim_{r_{12} \rightarrow 0} \left(\frac{\partial \Psi}{\partial r_{12}} \right)_{av} = \frac{1}{2} \Psi(r_{12} = 0). \quad (2.1)$$

Where "av" stands for the spherical averaging. In the region of small r_{12} the wave function is linear in r_{12} .

The correlation cusp is responsible for the slow convergence of the partial wave expansion. However, the relation (2.1) also tells us how to avoid this slow convergence.

Let us divide the Hamiltonian of a n -electron system into three parts

$$H = K + T_{12} + V_{12} \quad (2.2)$$

where T_{12} is the sum of the kinetic energy operator for the first and second particle, V_{12} the interaction between the two particles and K everything else. We write the n -electron wave function as

$$\Psi = g\Phi \quad (2.3)$$

and choose Φ such that

$$\nabla_i\Phi \text{ and } \Delta_i\Phi \text{ continuous at } r_{12}=0 \text{ for } i=1, 2. \quad (2.4)$$

The factor g should hence take care of the correlation cusp.

The Schrödinger equation can be written as

$$\begin{aligned} (H - E)\Psi &= (K + T_{12} + V_{12} - E)g\Phi \\ &= (K - E)g\Phi + V_{12}g\Phi + [T_{12}, g]\Phi + gT_{12}\Phi = 0. \end{aligned} \quad (2.5)$$

At $r_{12}=0$ the term $V_{12} = r_{12}^{-1}$ becomes singular. This singularity must be compensated by the only other term in (2.5) that may become singular, hence

$$V_{12}g\Phi + [T_{12}, g]\Phi = \text{regular for } r_{12} = 0. \quad (2.6)$$

This is achieved if

$$g = 1 + \frac{1}{2}r_{12} + 0(r_{12}^2) \quad (2.7)$$

because

$$\begin{aligned} [T_{12}, \frac{1}{2}r_{12}] &= -\frac{1}{r_{12}} - \frac{1}{2} \text{grad}_1 r_{12} \cdot \nabla_1 - \frac{1}{2} \text{grad}_2 r_{12} \cdot \nabla_2 \\ &= -\frac{1}{r_{12}} - \frac{1}{2} \frac{(\vec{r}_1 - \vec{r}_2)}{r_{12}} \{\nabla_1 - \nabla_2\} = -V_{12} + \hat{U}_{12}. \end{aligned} \quad (2.8)$$

In (2.8) we have defined the operator \hat{U}_{12} which will play a central role in this paper.

The simplest possible g is obtained if one truncates (2.7) after the linear term. One can repeat the same argument for all electron pairs such that a possible choice for the overall g is

$$g = \prod_{i < j} (1 + \frac{1}{2}r_{ij}) \quad (2.9)$$

and that for the corresponding Φ (2.4) holds for all $r_{ij} = 0$. One can then expand

Φ in a Hilbert space basis of Slater determinants

$$\Phi = \sum_{\mu} c_{\mu} \varphi_{\mu} \tag{2.10}$$

and hope that the expansion (2.10) is rapidly convergent since Φ has no correlation cusp.

The ansatz (2.3, 9, 10) is by no means new [11, 22], which is not astonishing, since one is led to it almost automatically if one wants to separate the correlation cusp from the rest of the problem. Some authors, e.g. Grein and Tseng [23] have replaced the factor $\frac{1}{2}$ in (2.9) by some variational parameter, but if one wants an exact explicit summation of the partial wave amplitudes with large l , $\frac{1}{2}$ is the best choice.

One can insert (2.3) into the expectation value $\langle H \rangle$ of the Hamiltonian H and make $\langle H \rangle$ stationary with respect to variation of the c_{μ} . This leads to a matrix eigenvalue problem with the CI matrix $\bar{H}_{\mu\nu}$ and the overlap matrix $\bar{S}_{\mu\nu}$

$$\bar{H}_{\mu\nu} = \langle \varphi_{\mu} g | H | g \varphi_{\nu} \rangle \tag{2.11}$$

$$\bar{S}_{\mu\nu} = \langle \varphi_{\mu} g | g | \varphi_{\nu} \rangle. \tag{2.12}$$

The function g given by (2.9) is certainly unpleasant for a large number of electrons. An alternative choice is

$$g = 1 + \frac{1}{2} \sum_{i < j} r_{ij} \tag{2.13}$$

but it may lead to problems with size-consistency.

These problems can be avoided by limiting the products $r_{ij}r_{kl}$ etc. in the expansion of (2.10) to those with all subscripts different, i.e.

$$g = 1 + \frac{1}{2} \sum_{i < j} r_{ij} + \frac{1}{4} \sum_{\substack{i < j \\ i \neq k, l \\ j \neq k, l}} \sum_{k < l} r_{ij} r_{kl} + \dots \tag{2.14}$$

One may also choose

$$g = \exp\left(\frac{1}{2} \sum_{i < j} r_{ij}\right). \tag{2.15}$$

We shall briefly comment later on this possibility.

Often a single term Φ_0 will dominate in the expansion (2.10). One can then write

$$\Psi = g(\Phi_0 + \omega) = g\Phi_0 + \chi \tag{2.16}$$

and expand χ in a Hilbert space basis

$$\chi = \sum_{\mu} d_{\mu} \varphi_{\mu}. \tag{2.17}$$

Then χ is no longer expected to be free from a correlation cusp, but if χ is a small correction to Φ_0 it may not matter too much that the expansion (2.17) is somewhat slower convergent than that of ω . It should certainly be superior to an expansion of the full Ψ in the $\{\varphi_{\mu}\}$.

The numerical applications of this paper will be based on the ansatz (2.16) which we shall present in more detail in Sect. 3.

There is an interesting alternative to the choice (2.9) for g , namely (2.15); with this choice (2.5) becomes

$$g \left\{ K + T_{12} - E - \frac{1}{4} - \frac{1}{2} \frac{\vec{r}_{12}}{r_{12}} \cdot (\nabla_1 - \nabla_2) \right\} \Phi = 0. \quad (2.18)$$

One can obviously multiply from the left by g^{-1} , and using (2.8) one gets

$$\{K + T_{12} - E - \frac{1}{4} + \hat{U}_{12}\} \Phi = 0 \quad (2.19)$$

Eq. (2.19) is reminiscent of Boys' method of transcorrelated wave functions [24]. Its only disadvantage is that the "effective Hamiltonian" for Φ is non hermitean, since

$$[T_{12}, r_{12}]^+ = -[T_{12}, r_{12}]. \quad (2.20)$$

The factor (2.15) does not spoil the square integrability of Ψ provided that Φ decreases faster than

$$\exp \left\{ -\frac{1}{2}(r_1 + r_2) \right\} \quad (2.21)$$

for $r_1, r_2 \rightarrow \infty$.

3. CI with a correlated reference function

In the numerical applications of this paper we start from the ansatz (2.16). We hence write the wave function as

$$\Psi = c_0 N g \Phi + \sum_{\mu > 0} c_\mu \varphi_\mu \quad (3.1)$$

where Φ and the φ_μ are configuration state n -electron functions built up from a basis of (generally non orthogonal) orbitals (Φ may also be a multiconfiguration function), while g is a function that depends on the interelectronic coordinates r_{ij} . The normalization factor N is conveniently determined by the requirement

$$\|Ng\Phi\| = 1 \quad (3.2)$$

and the φ_μ are supposed to be normalized as well (which is not necessary).

The c_μ (and the energy E) are, as in the standard CI approach, obtained from a matrix eigenvalue equation

$$\sum_\nu \tilde{H}_{\mu\nu} c_\nu = E \sum_\nu \tilde{S}_{\mu\nu} c_\nu \quad (3.3)$$

with

$$\tilde{H}_{00} = N^2 \langle \Phi | g H g | \Phi \rangle; \quad \tilde{S}_{00} = 1 \quad (3.4a)$$

$$\tilde{H}_{0\mu} = \tilde{H}_{\mu 0} = N \langle \Phi | g H | \varphi_\mu \rangle \quad \text{for } \mu > 0 \quad (3.4b)$$

$$\tilde{S}_{0\mu} = \tilde{S}_{\mu 0} = N \langle \Phi | g | \varphi_\mu \rangle \quad \text{for } \mu > 0 \quad (3.5a)$$

$$\tilde{H}_{\mu\nu} = \langle \varphi_\mu | H | \varphi_\nu \rangle; \quad \tilde{S}_{\mu\nu} = \langle \varphi_\mu | \varphi_\nu \rangle \quad \text{for } \mu > 0, \nu > 0. \quad (3.5b)$$

The matrix \tilde{H} differs from that of conventional CI (to which it reduces for $g \equiv 1$) only in the elements of the first row and the first column (i.e. for $\mu = 0$ and/or $\nu = 0$).

If Φ satisfies

$$(H_0 - E_0)\Phi = 0; \quad H_0 = H - V \tag{3.6}$$

and if we introduce the operator (see Eq. 2.8)

$$\hat{U} = [H_0, g] + V \tag{3.7}$$

the matrix elements (3.4a) and (3.4b) can be rewritten as

$$\tilde{H}_{00} = E_0 + N^2 \langle \Phi | g \hat{U} + g V (g - 1) | \Phi \rangle \tag{3.8a}$$

$$\tilde{H}_{0\mu} = E_0 \tilde{S}_{0\mu} + N \langle \varphi_\mu | \hat{U} + (g - 1) V | \Phi \rangle. \tag{3.8b}$$

Note that

$$\langle \Phi | \hat{U} | \Phi \rangle = \langle \Phi | V | \Phi \rangle. \tag{3.9}$$

In this section we have not specified how we choose the (uncorrelated) reference function Φ , nor the correlation factor g . As to Φ it is convenient if it is an exact eigenfunction of some H_0 - otherwise the simplifications (3.8) don't hold. A possible choice of H_0 is the bare nuclear Hamiltonian.

We shall first study the two-electron case explicitly (Sect. 4 to 6) before we comment on the best choice of Φ and g , and on further improvements.

4. Perturbation theory of He-like atoms

The simplest problem to be treated by our method is that of the lowest orders of perturbation theory for the ground state of 2-electron atoms ($1/Z$ -expansion). Instead of an eigenvalue problem we have to solve the inhomogeneous equation (Φ being the zeroth order ground state function, which satisfies (3.6))

$$-(H_0 - E_0)\psi = (V - E_1)\Phi \tag{4.1}$$

$$H_0 = T_1 + T_2 - \frac{Z}{r_1} - \frac{Z}{r_2}; \quad V = \frac{1}{r_{12}}. \tag{4.2}$$

From the partial wave expansion of the perturbation V

$$V = \sum_{l=0}^{\infty} V_l P_l(\cos \vartheta_{12}); \quad V_l = \frac{r_{<}^l}{r_{>}^{l+1}} \tag{4.3}$$

we automatically get an expansion of ψ

$$\psi = \sum_{l=0}^{\infty} \psi_l P_l(\cos \vartheta_{12}) \tag{4.4}$$

by solving

$$-(H_0 - E_0)\psi_l = (V_l - \delta_{l0}E_1)\Phi. \tag{4.5}$$

It is known [8] that for large l the solution ψ_1 approaches asymptotically

$$\psi_l \rightarrow \frac{1}{2}(r_{12})_l \Phi; \quad (r_{12})_l = \frac{1}{2l+3} \frac{r_{<}^{l+2}}{r_{>}^{l+1}} - \frac{1}{2l-1} \frac{r_{<}^l}{r_{>}^{l-1}}. \quad (4.6)$$

We use this property to formulate the first-order perturbation function as

$$\psi = \frac{1}{2} r_{12} \Phi + \chi \quad (4.7)$$

With this choice ψ automatically has the correct asymptotic behavior for large l (and hence for small r_{12}) and the function χ serves mainly to correct $\frac{1}{2} r_{12} \Phi$ for small l (i.e. for large r_{12}). We expect that in the partial wave expansion of χ

$$\chi = \sum_{l=0}^{\infty} \chi_l P_l(\cos \vartheta_{12}) \quad (4.8)$$

large l -values contribute very little, i.e. that the expansion (4.8) will (at variance with (4.4)) be rapidly convergent.

Insertion of (4.7) into (4.1) leads to

$$\begin{aligned} -\frac{1}{2}(H_0 - E_0)r_{12}\Phi - (H_0 - E_0)\chi &= -\frac{1}{2}[T_1 + T_2, r_{12}]\Phi - (H_0 - E_0)\chi \\ &= (r_{12}^{-1} - E_1)\Phi \\ &= \{r_{12}^{-1} + \frac{1}{2}(\text{grad}_1 r_{12}) \cdot \nabla_1 \\ &\quad + \frac{1}{2}(\text{grad}_2 r_{12}) \cdot \nabla_2\}\Phi - (H_0 - E_0)\chi. \end{aligned} \quad (4.9)$$

One sees that the terms r_{12}^{-1} cancel. With the definition (2.8) of \hat{U}_{12} (4.9) becomes

$$-(H_0 - E_0)\chi = (\hat{U} - E_1)\Phi. \quad (4.10)$$

For the explicit choice

$$\Phi = N' \cdot e^{-Z(r_1+r_2)} \quad (4.11)$$

we get

$$\hat{U}\Phi = U\Phi; \quad U = Z \frac{r_1+r_2}{2r_{12}} (1 - \cos \vartheta_{12}) = Z \frac{r_1+r_2}{r_{12}} \sin^2 \frac{\vartheta_{12}}{2}, \quad (4.12)$$

i.e. the effect of the operator \hat{U} on Φ is equal to that of the multiplicative operator U , although \hat{U} is generally non-local (it is not even hermitean).

With \hat{U} replaced by U as given by (4.12) and Φ given by (4.11) inserted, (4.10) looks very much like (4.1) just with ψ replaced by χ and V replaced by U . The essential difference between U and V is that V has a strong short-range contribution - and a singularity for $r_{12}=0$ - while U has no appreciable short-range-contribution - it even vanishes for $r_{12}=0$. (U vanishes for $\vartheta_{12}=0$ and becomes $-Z \sin \vartheta_{12}/2$ for $r_1=r_2$). This is probably the key to an understanding of the improvement with respect to the traditional partial wave expansion.

The “residual potential” U has the l -expansion

$$U = \sum_{l=0}^{\infty} U_l P_l(\cos \vartheta_{12}) \quad (4.13a)$$

$$2U_l = \left\{ -\frac{l}{2l-1} \frac{r_{<}^{l-1}}{r_{>}^{l-1}} + \frac{l-1}{2l-1} \frac{r_{<}^l}{r_{>}^l} + \frac{l+2}{2l+3} \frac{r_{<}^{l+1}}{r_{>}^{l+1}} - \frac{l+1}{2l+3} \frac{r_{<}^{l+2}}{r_{>}^{l+2}} \right\} Z \quad (4.13b)$$

and (at variance with V) no discontinuity for $r_1 = r_2$.

We obtain an upper bound to the second order energy $E^{(2)}$ from the Hylleraas functional

$$F(\psi) = \langle \psi | H_0 - E_0 | \psi \rangle + 2 \operatorname{Re} \langle \psi | V - E_1 | \Phi \rangle = A^{(2)} + \sum_{l=0}^{\infty} \tilde{E}_l^{(2)} \geq E^{(2)} \quad (4.14)$$

$$A^{(2)} = \frac{1}{2} \langle \Phi | \Phi \rangle - E_1 \langle \Phi | r_{12} | \Phi \rangle + \frac{1}{2} \langle \Phi | r_{12} U | \Phi \rangle = -\frac{15}{128} h = -0.1171875 h \quad (4.15)$$

(h stands for “Hartree” = atomic energy unit)

$$\tilde{E}_l^{(2)} = 2 \operatorname{Re} \langle \chi_l | \hat{U} - E_1 | \Phi \rangle + \langle \chi_l | H_0 - E_0 | \chi_l \rangle. \quad (4.16)$$

The (upper bound to the) second order energy consists of an “overhead” $A^{(2)}$ and a sum over partial wave contributions $\tilde{E}_l^{(2)}$. In the conventional partial-wave expansion one gets instead of (4.14) a sum

$$\sum_{l=0}^{\infty} E_l^{(2)} \geq E^{(2)} \quad (4.17)$$

without an “overhead”. The individual partial wave contributions can in either case be obtained independently.

We have chosen two approaches to minimize the contributions (4.16): (a) a single term approximation

$$\chi_l = N_k r_1^l r_2^l e^{-\eta_l(r_1+r_2)} P_l(\cos \vartheta_{12}) \quad (4.18)$$

with the non-linear parameter η_l optimized, (b) an expansion

$$\chi_l = \sum_{n,m=1}^M c_{nm} r_1^{l+n-1} r_2^{l+m-1} e^{-\eta_l(r_1+r_2)} P_l(\cos \vartheta_{12}) \quad (4.19)$$

with η_l close to the optimum value for a single-term approximation and with M chosen such that increase of M does not change $\tilde{E}_l^{(2)}$ by more than 10^{-7} Hartree. $M = 10$ turned out to be sufficient. (For the column $\tilde{E}_l^{(2)}(l+\frac{1}{2})^8$ in Table 1 we have evaluated the $\tilde{E}_l^{(2)}$ to 4 significant digits). The results are displayed in Table 1. One sees:

(1) Almost 75% of $E^{(2)}$ comes from the overhead $A^{(2)}$. Among the partial wave contributions $\tilde{E}_l^{(2)}$ the term with $l = 1$ dominates by far, the sum of the terms with $l \neq 1$ gives a contribution of less than 1% of $E^{(2)}$.

(2) The one-term approximation (4.18) is usually quite good, it yields (with $A^{(2)}$) 99.8% of $E^{(2)}$. (An analogous statement does *not* hold for the conventional $E_l^{(2)}$,

Table I. Second order perturbation theory of He-like ions^a

<i>l</i>	$-\tilde{E}_l^{(2)}$ one term	$\frac{\eta_{\text{opt}}}{Z}$	$-\tilde{E}_l^{(2)}$ converged ^b	$-\tilde{E}_l^{(2)} \cdot (1 + \frac{1}{2})^8 \times 10^{-6}$	$\frac{1}{2}h^c$	$A_l^{(2)}$	$-E_l^{(2)} = -A_l^{(2)} - \tilde{E}_l^{(2)d}$	$-E_l^{(2)e}$	$-E_l^{(2)}$ exact ^f
Overhead	117 187.50		117 187.50						
0 <i>s</i>	649.61	1.0	743.73	562 500.0		-124 590.07	125 333.80	125 334	125 333.83
1 <i>p</i>	39 022.65	1.183	39 212.11	1.0050	-55 266.7	+12 717.18	26 494.93	26 495	26 495.13
2 <i>d</i>	464.25	1.62	482.04	0.7355	-5 084.6	-3 424.03	3 906.07	3 906	3 906.11
3 <i>f</i>	31.61	2.08	34.42	0.7751	-1 239.9	-1 043.29	1 077.71	1 077	1 077.72
4 <i>g</i>	4.23	2.55	4.85	0.8162	-443.1	-401.24	406.09	405	406.09
5 <i>h</i>	0.83	3.05	1.00	0.8449	-196.3	-183.98	184.99	183	185.00
6	0.23	3.5	0.27	0.8654	-100.0	-95.52	95.79	94	95.79
7	0.07	4.0	0.09	0.8794	-56.2	-54.31	54.40	53	54.40
8	0.02	4.5	0.03	0.8837	-34.0	-33.09	33.12	32	33.12
$\Sigma \geq 6$			0.50		-268.4	-262.09	262.6	256	262.7
$\Sigma \geq 11$			0.00		-44.0	-43.59	43.59	42	43.8 ^g
			157 666.14		500 000.0	117 187.50	157 666.14	157 656	157 666.43 ^h

^a All energies in μh (microhartree)

^b To 0.1 μh

^c Partial wave expansion (4.23) of $\frac{1}{2}$ Hartree

^d This paper

^e F. W. Byron and C. J. Joachain Phys. Rev. **157**, 1 (1967)

^f H. v. Hirschhausen, unpublished, private communication by H. M. Schmidt, supposedly convergent to the exact values

^g H. M. Schmidt and H. v. Hirschhausen, Ref. [21]

^h Supposedly exact value of Ref. [26]

there the one term approximation becomes increasingly poorer for increasing l . This is due to the singularity of V_l for $r_1 = r_2$. Only wave functions depending on $r_<$ and $r_>$ exhibit satisfactory convergence [21].

(3) The $\tilde{E}_l^{(2)}$ depend on l as $(l + \frac{1}{2})^{-8}$ while the $E_l^{(2)}$ of the conventional partial wave expansion have the l -dependence $(l + \frac{1}{2})^{-4}$ [8]. This is, of course, a substantial increase of the speed of convergence. In order to get a result convergent at the microhartree level it is sufficient in our expansion to include l -values up to $l = 5$, while in the conventional approach one would need to go to about $l = 50$. (This is quite prohibitive and one is in the conventional approach obliged to use some extrapolation techniques if one wants to get an accurate $E^{(2)}$).

It is possible to get the $E_l^{(2)}$ of the conventional l -expansion indirectly in the following very economic way. We expand the r_{12} -term in (4.7) formally in terms of $P_l(\cos \vartheta_{12})$ and use the expansion (4.3) for V . Then the overhead contribution $A^{(2)}$ given by (4.15) has the l -expansion:

$$A^{(2)} = \sum_{l=0}^{\infty} A_l^{(2)} \tag{4.20a}$$

$$A_l^{(2)} = \frac{1}{2} \langle \Phi | (r_{12})_l (r_{12}^{-1})_l | \Phi \rangle - \frac{1}{2} \langle \Phi | (r_{12})_l U_l | \Phi \rangle + \delta_{l0} E_1 \langle \Phi | (r_{12})_0 | \Phi \rangle. \tag{4.20b}$$

These $A_l^{(2)}$ are also displayed in Table 1, together with

$$E_l^{(2)} = \tilde{E}_l^{(2)} + A_l^{(2)}. \tag{4.21}$$

One sees that not only for large l the bulk of $E_l^{(2)}$ is in the $A_l^{(2)}$, but even so for $l = 0$. Only $A_l^{(2)}$ is a poor approximation to $E_l^{(2)}$, it has even the wrong sign. The asymptotic expression for A_l is [8]

$$A_l^{(2)} = -\frac{45}{256} (l + \frac{1}{2})^{-4} \{ 1 - \frac{5}{4} (l + \frac{1}{2})^{-2} \} + 0 \{ (l + \frac{1}{2})^{-8} \}. \tag{4.22}$$

The A_l decrease rather slowly with increasing l , and this slow decrease determines that of $E_l^{(2)}$, while the $\tilde{E}_l^{(2)}$ decrease as $(l + \frac{1}{2})^{-8}$.

It has probably been known to Schwartz [8] and been elaborated in more detail by Schmidt and v. Hirschhausen [21] that the asymptotic contributions to $E^{(2)}$ that go as $(l + \frac{1}{2})^{-4}$ and $(l + \frac{1}{2})^{-6}$ can be obtained in closed analytic form as given in (4.22) while the contributions $(l + \frac{1}{2})^{-8}$ and of higher order (only even orders appear to contribute) can only be obtained numerically.

It is noteworthy that *the contribution $0 \{ (l + \frac{1}{2})^{-4} \}$ to $A_l^{(2)}$ comes entirely from the expansion of*

$$\frac{1}{2} = \frac{1}{2} \sum_l \langle \Phi | (r_{12})_l (r_{12}^{-1})_l | \Phi \rangle = \frac{1}{2} \sum_l 1_l \tag{4.23a}$$

$$1_l = -\frac{45}{128} (l + \frac{1}{2})^{-4} + 0 \{ (l + \frac{1}{2})^{-6} \}, \tag{4.23b}$$

while the contribution of the 2nd term in (4.20b) is $0 \{ (l + \frac{1}{2})^{-6} \}$ and the 3rd term vanishes for $l \neq 0$.

One understands why the partial wave expansion converges so poorly: it is because one replaces the number 1 by its partial wave expansion (4.23).

In order to make this point very clear we have included in Table 1 the partial wave contributions (4.23) to the term $1/2$. The square of the norm of the (conventional) first-order perturbation function ψ has the following partial wave expansion

$$\langle \psi | \psi \rangle = \langle \frac{1}{2} r_{12} \Phi + \chi | \frac{1}{2} r_{12} \Phi + \chi \rangle = \sum_l \langle \psi | \psi \rangle_l \quad (4.24a)$$

$$\langle \psi | \psi \rangle_l = \frac{1575}{1024 \cdot Z^2} (l + \frac{1}{2})^{-6} + 0 \{ (l + \frac{1}{2})^{-8} \}. \quad (4.24b)$$

If one expands ψ in terms of normalized partial waves $\tilde{\psi}_l$

$$\psi = \sum_l \psi_l = \sum_l c_l \tilde{\psi}_l \quad (4.25)$$

the coefficients c_l go as $(l + \frac{1}{2})^{-3}$ while one finds the following l -dependence of the off-diagonal matrix elements and energy denominators

$$\langle \Phi | V - E_1 | \tilde{\psi}_l \rangle = 0 \{ (l + \frac{1}{2})^{-1} \} \quad (4.26a)$$

$$\langle \tilde{\psi}_l | H_0 - E_0 | \tilde{\psi}_l \rangle = 0 \{ (l + \frac{1}{2})^{-2} \} \quad (4.26b)$$

$$\langle \tilde{\psi}_l | U - E_1 | \Phi \rangle = 0 \{ (l + \frac{1}{2})^{-3} \} \quad (4.26c)$$

$$\psi = \frac{1}{2} r_{12} \Phi + \chi, \quad \chi = \sum_l d_l \tilde{\chi}_l; \quad d_l = 0 \{ (l + \frac{1}{2})^{-5} \} \quad (4.27)$$

The convergence of the expansion (4.26) is, on the whole, not very bad [7], it is only bad if one has to multiply ψ by a factor like r_{12}^{-1} which has a singularity at the origin and where the behavior of ψ for small r_{12} matters a lot.

From our λ one can also calculate the third order energy $E^{(3)}$ as well as two expectation values

$$E_{\text{exp}} = \frac{\langle \Phi + \lambda \psi | H | \Phi + \lambda \psi \rangle}{\langle \Phi + \lambda \psi | \Phi + \lambda \psi \rangle} \quad (4.28)$$

namely for $\lambda = 1$ and for the minimum as function of λ . We get for $Z = 2$ (i.e. for the He ground state).

$$E^{(3)} = 8702 \mu\text{h} \quad (\text{exact: } 8699 \mu\text{h} [25])$$

$$E_{\text{exp}}(\lambda = 1) = -2.894480 \text{ h}$$

$$E_{\text{exp}}(\lambda_{\text{opt}}) = -2.895449 \text{ h}$$

The "overhead" contribution to $E^{(3)}$ is $10986 \mu\text{h}$.

The above results agree well with the best results from the literature [25, 26].

5. Variational calculation of the He ground state (and the isoelectronic ions)

The considerations of Sects. 1-3 as well as the numerical results of Sect. 4 suggest use of the following variational ansatz for the ground state of two-electron atoms

$$\Psi = (1 + \frac{1}{2}r_{12})\Phi + \chi \tag{5.1}$$

with

$$(H_0 - E_0)\Phi = 0 \tag{5.2}$$

and with χ expanded in a given basis of two-electron functions

$$\chi = \sum_{\mu} d_{\mu} \varphi_{\mu}. \tag{5.3}$$

H_0 may be the bare-nuclear Hamiltonian, but in order to be somewhat more general we chose

$$H_0 = T_1 + T_2 - \frac{\alpha}{r_1} - \frac{\alpha}{r_2}; \quad \Phi = N e^{-\alpha(r_1+r_2)}, \tag{5.4}$$

where α need not be equal to Z . Then

$$H = H_0 + V; \quad E_0 = -\alpha^2; \quad V = \frac{\alpha - Z}{r_1} + \frac{\alpha - Z}{r_2} + \frac{1}{r_{12}} \tag{5.5}$$

$$U\Phi = \left\{ \alpha \frac{r_1 + r_2}{2r_{12}} (1 - \cos \vartheta_{12}) + (\alpha - Z) \left(\frac{1}{r_1} + \frac{1}{r_2} \right) \right\} \Phi. \tag{5.6}$$

The modified CI matrix \tilde{H} and the corresponding overlap matrix \tilde{S} are then

$$\begin{aligned} \tilde{H}_{00} &= \langle \Phi | \frac{1}{r_{12}} + (\alpha - Z)(1 + r_{12}) \left(\frac{1}{r_1} + \frac{1}{r_2} \right) + \frac{1}{2} + \frac{1}{4}r_{12} \\ &\quad + \frac{1}{4}\alpha(r_1 + r_2) + \frac{\alpha - Z}{4} \left(\frac{r_{12}^2}{r_1} + \frac{r_{12}^2}{r_2} \right) | \Phi \rangle / D + E_0 \\ &= \left\{ \frac{5}{8}\alpha + (\alpha - Z) \left(2\alpha + \frac{15}{4} \right) + \frac{1}{2} + \frac{35}{64\alpha} + \frac{3}{4} + 9 \frac{\alpha - Z}{4\alpha} \right\} / D + E_0 \end{aligned} \tag{5.7}$$

$$D = 1 + \frac{35}{16\alpha} + \frac{3}{2\alpha^2}; \quad \tilde{S}_{00} = 1 \tag{5.8}$$

$$\tilde{H}_{0\mu} = \tilde{H}_{\mu 0} = \langle \varphi_{\mu} | U + \frac{1}{2} + (1 + \frac{1}{2}r_{12})E_0 | \Phi \rangle / \sqrt{D}; \quad \mu > 0 \tag{5.9}$$

$$\tilde{S}_{0\mu} = \tilde{S}_{\mu 0} = \langle \varphi_{\mu} | 1 + \frac{1}{2}r_{12} | \Phi \rangle / \sqrt{D}; \quad \mu > 0 \tag{5.10}$$

$$\tilde{H}_{\mu\nu} = \langle \varphi_{\mu} | H | \varphi_{\nu} \rangle \quad \text{for } \mu, \nu > 0 \tag{5.11}$$

$$\tilde{S}_{\mu\nu} = \langle \varphi_{\mu} | \varphi_{\nu} \rangle \quad \text{for } \mu, \nu > 0. \tag{5.12}$$

This CI matrix differs from the traditional CI matrix only in the first row and the first column.

Table 2. The He ground state in conventional CI and in CI with correlated reference function^a

l	conventional CI ^b	new ^d ($\alpha = 2.0$)	l -decoupled ^e	overhead decomposition ($\alpha = 0$) ^f	new ^d ($\alpha = 1.885$)	new ^d ($\alpha = 1.6875$)
One-conf. ^g	2 750 000 ^c	2 750 000.0	—	—	2 808 650.0	2 847 656.3
Overhead ^h	—	126 582.3	—	(126 582)	80 067.5	5 124.8
s 0	129 027	26 368.0	26 368.0	105 600	14 342.3	49 534.5
p 1	21 486	727.6	6 350.0	18 028	642.5	1 332.2
d 2	2 252	36.8	85.6	2 046	19.2	61.2
f 3	554	6.9	6.3	526	2.2	9.8
g 4	198	2.2	0.9	189	0.46	2.8
h 5	87	0.85	0.18	83	0.15	1.1
$\Sigma \geq 6^i$	118	$\geq 0.5; \leq 2.3$	~ 0.1	~ 110	$\geq 0.06; \leq 0.3$	$\geq 0.7; \leq 3.0$
E_{best}^j		2 903 722.1		2 903 724.0	2 903 724.0	2 903 721.7
E_{exact}^k		2 903 724.4		2 903 724.4	2 903 724.4	2 903 724.4
Overhead/ E		82.3%			84.2%	9.1%
N_{conf}^l	8 000/1 000	200/70			80/40	250/90

^a All energies negative, in microhartree (μh)

^b From Ref. [7]

^c The l -limits of conventional CI are independent of α , we have chosen the reference functions with $\alpha = 2$ to make the comparison with the following columns easier

^d CI with correlated reference function described in Sect. 5

^e CI with g and a single X_l (for $\alpha = 2.0$)

^f As described in Sect. 5

^g $\langle \Phi | H | \Phi \rangle / \langle \Phi | \Phi \rangle$ with uncorrelated reference function Φ

^h $\langle \Phi | g H g | \Phi \rangle / \langle \Phi | g^2 | \Phi \rangle - \langle \Phi | H | \Phi \rangle / \langle \Phi | \Phi \rangle$ with $g = 1 + \frac{1}{2}r^{1,2}$

ⁱ The upper bounds given here are the differences between the exact values and our best calculations with l up to 5. These differences overestimate $\Sigma \geq 6$ because our basis is not fully saturated for the low l -values. The lower bounds are the sums of the increments for $l = 6$ and $l = 7$

^j Energy of a CI with STO basis functions from $l = 0$ to 7 and n from $l+1$ to 9 ($n = 8$ for $l = 0$) with the orbital exponents given in Table 5. This corresponds to 156 additional configurations or 49 natural configurations

^k Ref. [13], see also Ref. [26]

^l Number of configurations necessary for 1 μh and 10 μh accuracy respectively

It is not unexpected that the matrix element \tilde{H}_{00} is a better approximation to the eigenvalue E than the traditional H_{00} ; for $\alpha = Z = 2$ we get in hartrees

$$\begin{aligned} H_{00} &= E_0 + \langle \Phi | V | \Phi \rangle = -2.75 \\ \tilde{H}_{00} &= -2.876582 \\ E &= -2.903724. \end{aligned} \tag{5.13}$$

By variation of α one can minimize both H_{00} and \tilde{H}_{00}

$$\begin{aligned} H_{00}(\alpha = 1.6875) &= -2.8475 \\ \tilde{H}_{00}(\tilde{\alpha} = 1.885) &= -2.888718. \end{aligned} \tag{5.14}$$

Note that the optimum α for \tilde{H}_{00} is quite different from that for H_{00} .

We have used the STO expansion (4.19) for χ_l . The eigenvalue problems have been solved directly in the non-orthogonal basis by means of the “inverse-iteration” technique, that has turned out to be very stable and rather insensitive to near-linear dependencies.

The converged ground state energies can be expressed as sums of the energy of the uncorrelated reference function Φ , the “overhead” (difference between the expectation value of Φ and $g\Phi$) and increments for the various partial waves, where the s -increment is obtained from a calculation with s -AO’s only, the p -increment from a calculation with s - and p -AO’s etc. Increments for the traditional CI are constructed from Ref. [7] in the analogous way [27]. Both sets are displayed in Table 2, together with some other data to be explained below.

If one compares the CI increments \tilde{E}_l with the partial wave contributions $\tilde{E}_l^{(2)}$ to $\tilde{E}^{(2)}$ in Table 1 one sees that the convergence with our ansatz is faster for the CI energy than for $\tilde{E}^{(2)}$. In fact one has to include $l=5$ for 1 μ h accuracy of $E^{(2)}$, but only $l=4$ for CI - or $l=3$ for 10 μ h accuracy of $E^{(2)}$ but $l=2$ for CI.

The better convergence of the partial wave expansion for E_{CI} than for $\tilde{E}^{(2)}$ is really surprising, especially as it holds for the whole isoelectronic series (see Table 3), since the ansatz (5.1) has been rigorously justified for the wave functions to first-order of perturbation theory and only made plausible for the full wave function as an approximate treatment of the ansatz (2.16). There can be no doubt, however, that the ansatz $(1 + \frac{1}{2}r_{12})\Phi$ is a better approximation to the exact wave function Ψ than to the first order wave function $\Phi + \psi$ and that hence the “correction” χ is smaller for Ψ than for ψ .

This becomes obvious if one compares the ratios of the “overheads” to the exact energy improvements $E - H_{00}$ with the corresponding ratios $A^{(2)}/E^{(2)}$ in perturbation theory. In CI these ratios vary (depending on Z , excluding $Z = 1$) between 80% and 85% while the corresponding ratio in perturbation theory is 74%.

In order to understand what happens we must compare the Eqs. (4.10) or (4.16) of perturbation theory with the Schrödinger equation for the exact χ .

Table 3. The ground state of ions isoelectronic with He^a, for $\alpha = Z$ with the present ansatz^d

l	H ⁻	Li ⁺	Be ²⁺	C ⁴⁺	Ne ⁸⁺
One config. ^g	375 000.0	7 125 000.0	13 500 000.0	32 250 000.0	93 750 000.0
Overhead ^h	108 333.3	130 494.5	130 952.4	129 629.6	126 646.4
0 <i>s</i>	42 696.3	23 408.9	23 323.9	24 930.5	28 072.2
1 <i>p</i>	1 616.4	968.5	1 243.1	1 625.3	2 007.7
2 <i>d</i>	78.1	33.7	40.4	54.2	71.2
3 <i>f</i>	14.5	4.4	4.3	5.2	6.9
4 <i>g</i>	4.58	1.16	0.91	0.93	1.20
5 <i>h</i>	1.95	0.41	0.27	0.23	0.29
$\Sigma \geq 6^i$	$\geq 1.3; \leq 4.5$	$\geq 0.2; \leq 1.7$	$\geq 0.1; \leq 1.0$	$\geq 0.1; \leq 0.7$	$\geq 0.1; \leq 0.6$
E_{best}^j	527 746.5	7 279 911.7	13 655 565.3	32 406 246.0	93 906 806.0
E_{exact}^k	527 751.0	7 279 913.4	13 655 566.2	32 406 246.6	93 906 806.5
Overhead/ ΔE	70.9%	84.2%	84.2%	83.0%	80.8%

Footnotes as in Table 2

The counterpart of (4.10) for the exact χ is

$$(H - E)\chi = -(U - \Delta E + \frac{1}{2} - \frac{1}{2}\Delta E r_{12})\Phi. \quad (5.15)$$

We can now treat this equation by perturbation theory (which will not lead to the $1/Z$ expansion, since already in the overhead various orders in $1/Z$ are scrambled). The lowest order approximation to (5.15) is

$$(H_0 - E')\chi = -(U - \Delta E' + \frac{1}{2} - \frac{1}{2}\Delta E' r_{12})\Phi \quad (5.16)$$

and the corresponding Hylleraas functional

$$\langle \chi | H_0 - E' | \chi \rangle + 2 \langle \chi | U - \Delta E' + \frac{1}{2} - \frac{1}{2}\Delta E' r_{12} | \Phi \rangle. \quad (5.17)$$

In order to apply *RS* rather than *BW* perturbation theory one should identify E' and $\Delta E'$ in (5.17) as

$$E' = \langle \Phi | (1 + \frac{1}{2}r_{12}) H (1 + \frac{1}{2}r_{12}) | \Phi \rangle; \quad \Delta E' = E' - E_0. \quad (5.18)$$

The main difference between (4.10) and (5.16) is that the “effective potential” is no longer simply U , but $U - \frac{1}{2}\Delta E' \cdot r_{12}$. While U is repulsive, $-\Delta E' r_{12}$ is attractive (since E' is positive) such that there is some cancellation of the two contributions.

An intermediate step between (5.15) – where the χ_l for the different l couple through the V contained in H – and (5.16) – where the χ_l decouple – consists in switching off the coupling in (5.15) artificially by limiting χ in (5.15) to one χ_l . We have calculated these independent l -increments \tilde{E}_l (which do not add up to the total energy improvement) for various values of Z and included the results for $Z=2$ in Table 2. As the \tilde{E}_l in Table 1 the \tilde{E}_l show a perfect asymptotic $(l + \frac{1}{2})^{-8}$ dependence. The factors of the $(l + \frac{1}{2})^{-8}$ term depend sensitively on Z and can be roughly fitted to

$$\gamma(z) = \gamma^{(\infty)} \left(1 - \frac{1.83}{Z} + \frac{1.5}{Z^2} \right)^2 \quad (5.19)$$

where $\gamma^{(\infty)}$ is the coefficient for the $\tilde{E}_l^{(2)}$ of second-order perturbation theory. We conclude that the first term in the parenthesis in (5.19) comes from U , the second term ($\sim Z^{-1}$) from $E' \Delta r_{12}$ and that there is, in fact a strong cancellation, especially for $Z=1$ and $Z=2$.

For all Z that we have studied $|\gamma(Z)|$ is appreciably smaller than $|\gamma^{(\infty)}|$.

A further reduction of the absolute value of the increments to the actual \tilde{E}_l comes from the interaction of the different χ_l . This effect is very small (and almost not observable) for $Z=2$, and it changes its sign (i.e. raises the absolute values of the E_l) for $Z=1$.

It is obvious that as long as the χ_l are determined by the direct interaction with Φ the coupling (which arises formally to third-order in the perturbation treatment of (5.15)) must raise the energy (as a third-order energy is usually positive). However, if the matrix elements between Φ and χ_l become very small (e.g. for Z near 1) the χ_l are mainly determined through their interaction with χ_0 (this is

Table 4. The ground state of ions isoelectronic with He^a, for α_{opt} with the present ansatz^d

<i>l</i>	H ⁻	Li ⁺	Be ²⁺	C ⁴⁺	Ne ⁸⁺
α_{opt}	0.826	2.915	3.933	5.953	9.971
One config. ^g	453 474.0	3 170 900.9	13 537 386.0	32 277 166.0	93 767 284.0
Overhead ^h	55 305.0	91 438.5	97 906.8	104 632.2	110 214.8
<i>s</i> 0	18 781.1	16 487.2	18 891.2	22 694.8	271 887.2
<i>p</i> 1	183.6	1 050.8	1 334.1	1 689.4	2 038.3
<i>d</i> 2	5.5	32.2	42.6	57.0	73.0
<i>f</i> 3	0.8	3.2	4.1	5.4	7.0
<i>g</i> 4	0.26	0.61	0.75	0.94	1.22
<i>h</i> 5	0.14	0.17	0.19	0.24	0.32
$\Sigma \geq 6^i$	$\geq 0.1; \leq 0.6$	$\geq 0.1; \leq 0.8$	$\geq 0.1; \leq 0.5$	$\geq 0.1; \leq 0.6$	$\geq 0.1; \leq 0.6$
E_{best}^j	527 750.4	7 279 912.7	13 655 565.8	32 406 246.1	93 906 806.1
E_{exact}^k	527 751.0	7 279 913.4	13 655 566.2	32 406 246.6	93 906 806.5
Overhead/ <i>E</i>	74.4%	83.8%	88.5%	81.1%	79.0%

Footnotes as in Table 2

formally an effect of fourth and higher order in perturbation theory of (5.15) and this must lower the energy. These two effects do not depend on l as $(l + \frac{1}{2})^{-8}$ (the first of them probably as $(l + \frac{1}{2})^{-9}$, the second probably as $(l + \frac{1}{2})^{-4}$, but with a very small coefficient) and the total increments E_l do not have a simple leading asymptotic term.

It is noteworthy that the rate of convergence of the l -expansion can be changed by varying α . In the case of He the convergence for $\alpha = 1.885$ (which minimizes \tilde{H}_{00}) is significantly faster than for $\alpha = 2$, while for $\alpha = 1.6875$ (which minimizes H_{00}) it is poorer, as is seen by the number of configurations necessary to reach 1 μ h accuracy. It is not too surprising that one can speed up the convergence if one minimizes the energy with respect to the correlated reference function. The $\tilde{\alpha}$ which minimizes \tilde{H}_{00} can be expressed as

$$\tilde{\alpha}_{\text{opt}} = Z - 0.304/Z + 0.14/Z^2. \quad (5.20)$$

Note that in (5.20) no constant term is present, while the α that minimizes H_{00} is simply $\alpha_{\text{opt}} = Z - 0.3125$. For all Z , but very much so for large Z , $\tilde{\alpha}_{\text{opt}}$ according to (5.20) shows much less shielding than α_{opt} .

Since the correlation factor $(1 + \frac{1}{2}r_{12})$ somehow takes care of the mutual shielding of the electrons there is obviously little need for additional shielding.

As can be seen from Tables 3 and 4 the choice of α_{opt} rather than Z has an appreciable effect on the convergence of the l -expansion only for $Z = 2$ and - to a spectacular extent - for $Z = 1$. If one chooses α_{opt} , the convergence of the \tilde{E}_l of H^- is, by far, the fastest of those for all Z . One could think of determining Φ (not restricted to being a product of two simple Slater functions) by minimizing the expectation value with respect to $(1 + \frac{1}{2}r_{12})\Phi$. This could be called the "Hylleraas-SCF" method. So far we have made no numerical attempt in this direction.

We have seen that in perturbation theory the $E_l^{(2)}$ of the conventional partial wave expansion for large l are dominated by the contribution $A_l^{(2)}$ of the overhead. In order to show that the same holds for the E_l of CI, we have also evaluated the analogous partial wave contributions of the overhead $\tilde{H}_{00} - H_{00}$ in CI. We did so in the same way as one would do in a CI calculation, i.e. limiting the basis first to s -AO's, then to s and p (in Ref. [27] our quantities would have the subscript "ev" rather than "in").

These A_l are also included in Table 2. They go asymptotically (for $Z = \alpha = 2$) as

The leading term in $(l + \frac{1}{2})^{-4}$ comes again from the partial wave expansion (4.23) of the term $\frac{1}{2}$ in (5.7), but it is now divided by an energy denominator $D = 2.469$ which accounts for the reduction of the factor $\frac{45}{256} = 0.176$ to 0.071. Carrol et al. [7] find from a numerical fit to their l -increments of conventional CI

$$A_l = -0.071 (l + \frac{1}{2})^{-4} - 0.029 (l + \frac{1}{2})^{-5} + 0.016 (l + \frac{1}{2})^{-6} + 0\{(l + \frac{1}{2})^{-7}\} \quad (5.21)$$

$$E_l = -0.074 (l + \frac{1}{2})^{-4} - 0.031 (l + \frac{1}{2})^{-5}. \quad (5.22)$$

Table 5. Exponents of the STO's^a

$Z \backslash l$	0	1	2	3	4	5	6
1	0.7 (0.6)	1.0 (0.8)	1.8	2.1	2.5	3.0	3.5
2	2.5 (2.3)	3.2	4	5	6	7	8
3	4.25	4.8	6	7.5	9	10.5	12
4	5.3	6.5	8	10	12	14	16
6	8	10	12	15	18	21	24
10	15	15	20	25	30	35	40

^a Values in parentheses refer to calculations with $Z = \alpha_{\text{opt}}$

The closeness of the increments (5.21) and (5.22) shows again that the bulk of the conventional increments of the energy for large l is taken care of in our approach through the overhead.

One cannot expect, unlike in second-order perturbation theory, that the l -increments of the overhead add up with the \tilde{E}_l to the E_l of conventional CI.

The $(l + \frac{1}{2})^{-5}$ term in (5.21) – which is absent in the $A_l^{(2)}$ or the $E_l^{(2)}$, comes from the expansion of

$$\langle \Phi | r_{12} | \Phi \rangle = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \langle \Phi | (r_{12})_k \frac{1}{r_{12}} (r_{12})_l | \Phi \rangle = \sum_{l=0}^{\infty} \langle \Phi | r_{12} | \Phi \rangle_l \quad (5.23)$$

$$\langle \Phi | r_{12} | \Phi \rangle_l = \langle \Phi | (r_{12})_l \frac{1}{r_{12}} (r_{12})_l | \Phi \rangle + 2 \sum_{k=0}^{l-1} \langle \Phi | (r_{12})_k \frac{1}{r_{12}} (r_{12})_l | \Phi \rangle \quad (5.24)$$

The dependence of the denominator D (5.8) on $\alpha = Z$ explains easily the variations with Z of the energy increments E_1 of conventional CI observed previously [3].

If one is able to calculate energies to $1 \mu\text{h}$ accuracy, one realizes that certain effects are no longer negligible that one ignores in traditional quantum chemical calculations, such as mass polarization and relativistic corrections. We must hence state explicitly that in this paper we study the eigenvalues of the nonrelativistic Hamiltonian with clamped nuclei. (For He the mass polarization amounts to $22 \mu\text{h}$ and the relativistic correction to $39 \mu\text{h}$ [13]).

6. Possible improvements and generalization to larger systems

The ansatz that we have used in this paper is a special case of Hylleraas-CI. If one takes advantage of the greater flexibility of Hylleraas-CI, one gains about one more significant figure in the final result, but much of the simplicity of the present approach and of its analysis is lost, which we want to keep in view of a generalization to larger systems. A possible improvement that remains in the framework of our ansatz might be optimization of the reference function Φ . The possibility to minimize the expectation value of $g\Phi$ that has been referred to as

Hylleraas–Hartree–Fock in Sect. 5, probably leads to very complicated equations in the many electron case, especially if one uses the correlation factor (2.9).

Our ansatz (5.1) with Φ a single Slater determinant is certainly only good if one configuration Φ dominates in the CI expansion of Ψ . In cases of near-degeneracy, e.g. of the configurations $1\delta_g^2$ and $1\delta_u^2$ of H_2 for large R one should either choose a multiconfiguration reference function Φ , or as seems better, write

$$\Psi = g\{c_1\Phi_1 + c_2\Phi_2\} + \chi \tag{6.1}$$

and determine c_1 and c_2 together with the linear parameters in χ through the CI. The optimization of Φ_1 and Φ_2 is certainly different but probably less problematic than in conventional MS-SCF theory. Eigenfunctions of the bare nuclear Hamiltonian will probably do.

We believe that if one wants to take advantage of the results of this paper for arbitrary atoms or molecules one must cut a Gordian knot, this means that one must avoid the evaluation of many-electron integrals such as

$$\int \rho_1(1)\rho_2(2)\rho_3(3)r_{12}r_{23}^{-1} d\tau_1 d\tau_2 d\tau_3. \tag{6.2}$$

How can this be done?

The key idea is that traditional CI is not really bad, it only has difficulties to represent the wave function at those regions of configuration space where one interelectronic distance r_{ij} approaches zero. We must remedy this by performing a closed summation of partial waves for large l .

However, for many properties or intermediate expressions the traditional expansion is sufficiently well convergent and the r_{12} -dependent term is not needed – or rather it can simply be expanded (the norm of this expansion converges as l^{-6} , which is probably acceptable).

Let us regard CI as an approximation to our method in which the term $r_{12}\Phi$ is expanded in a basis $\Phi_{12}^{\mu\nu}$ of Slater determinants in which the spin-orbitals φ_1 and φ_2 are replaced by spin orbitals φ_μ and φ_ν

$$r_{12}\Phi = \sum_{\mu,\nu} c_{\mu\nu}^{12} \Phi_{12}^{\mu\nu} \tag{6.3}$$

The slow convergence of the CI comes mainly from the expansion of

$$\langle \Phi | \frac{1}{r_{12}} | \Phi \rangle = 1 = \sum_{\mu,\nu} c_{\mu\nu}^{12} \langle \Phi | \frac{1}{r_{12}} | \Phi_{12}^{\mu\nu} \rangle \tag{6.4}$$

The l -increments of (6.4) go as $(l + \frac{1}{2})^{-4}$, but a closed summation is, of course, trivial. Slowly convergent is also

$$\langle \Phi | r_{12} \frac{1}{r_{12}} | \Phi \rangle = \langle r_{12} \rangle = \sum_{\mu,\nu} \sum_{\rho,\delta} c_{\mu\nu}^{12} c_{\rho\delta}^{12} \langle \Phi_{12}^{\mu\nu} | \frac{1}{r_{12}} | \Phi_{12}^{\rho\delta} \rangle \tag{6.5}$$

– it goes as $(l + \frac{1}{2})^{-5}$ – but again a closed summation is trivial. What about terms with different r_{ij} factors? One sees easily that – at least in the atomic case – integrals such as

$$\langle \Phi | r_{12} r_{13}^{-1} | \Phi \rangle, \quad \langle \Phi | r_{12} r_{23}^{-1} r_{34} | \Phi \rangle \quad (6.6)$$

have a finite l -expansion. Hence we expect that the expansions analogous to (6.5) converge rapidly. Integrals like

$$\langle \Phi | r_{12} r_{23}^{-1} r_{31} | \Phi \rangle \quad (6.7)$$

have an infinite l -expansion but the l -increments go as $(l + \frac{1}{2})^{-6}$, which we regard as an acceptable rate of convergence. So the only expressions for which an expansion as (6.3) or (6.4) is slowly convergent are those which can be trivially evaluated in closed form. The conclusion is obvious, we can keep almost everything of traditional CI, but we must perform the summations (6.3) and (6.4) in closed form.

One can do this much in the spirit of the preceding sections of this paper in a way that can be implemented into direct CI; CEPA, coupled cluster and other methods, but it is best illustrated in the framework of Møller–Plesset perturbation theory (MB-PT). We limit ourselves to a closed-shell state. As usual we start with an SCF-calculation that leads to a reference Slater determinant Φ with the spin orbitals φ_i occupied. The first-order energy vanishes in view of the Brillouin theorem and one gets the second-order energy $E^{(2)}$ as the sum of pair energies ε_{ij} which are obtained by minimizing the Hylleraas functionals

$$\langle \omega_{ij} | F(1) + F(2) - e_i - e_j | \omega_{ij} \rangle + 2 \langle \omega_{ij} | r_{12}^{-1} | | \varphi_i(1) \varphi_j(2) \rangle \rangle \geq \varepsilon_{ij} \quad (6.8)$$

restricting the ω_{ij} to being strongly orthogonal to $|\varphi_i(1) \varphi_j(2)\rangle$. Alternatively, one can minimize a modified functional [19] without the strong-orthogonality constraint and remove the non-strongly-orthogonal contributions afterwards.

We now proceed on two parallel rails. We first construct ω_{ij} as traditionally in a basis expansion

$$\omega_{ij} = \sum_{\mu, \nu} |\varphi_\mu(1) \varphi_\nu(2)\rangle c_{\mu\nu} \quad (6.9)$$

then we use the alternative ansatz (now writing τ_{ij} for ω_{ij})

$$\tau_{ij} = \frac{1}{2} r_{12} |\varphi_i(1) \varphi_j(2)\rangle + \chi_{ij} \quad (6.10a)$$

$$\chi_{ij} = \sum_{\mu, \nu} |\varphi_\mu(1) \varphi_\nu(2)\rangle d_{\mu\nu} \quad (6.10b)$$

(using the same basis in (6.9) and (6.10)) which leads to a Hylleraas functional for χ_{ij} much like in Sect. 4 from which one gets a good approximation for the pair energy ε_{ij} . To second-order in the energy the pairs decouple and we can take full advantage of our ansatz (6.10) because no integrals involving products such as $r_{12} r_{23}$ etc. arise. The only difficulty is that integrals over products like $J(1) r_{12}$ or commutators like $[K(1) + K(2), r_{12}]$ appear, where J is a Coulomb operator and K is an exchange operator. These must be evaluated via matrix products

(i.e. by insertion of a completeness relation), somewhat like we have done successfully in some other contexts [28, 29].

If one wanted to compute the third-order-energy in terms of the τ_{ij} , integrals like (6.6) appear. However there is no objection against evaluating these contributions in terms of ω_{ij} since the basis set expansions are sufficiently well convergent.

Details will be given in a forthcoming paper.

There is an alternative to the insertion of completeness relations in two-electron space and this is suggested by the way that integrals over products of r_{ij} factors are usually calculated. One namely expresses

$$r_{ij}^{-1} = \frac{1}{\sqrt{\pi}} \int_0^\infty \exp[-\alpha r_{ij}^2] \alpha^{-1/2} d\alpha \quad (6.11)$$

so one can replace the factor r_{ij}^{-1} by $\exp[-\alpha r_{ij}^2]$ and r_{ij} by $(r_{ij})^2 \exp[-\alpha r_{ij}^2]$ such that the integrals over the particle coordinates can be done in closed form provided that the basis functions are Gaussians as well. Finally one performs the integration over α by a Gaussian quadrature. This amounts to replacing integrals over r_{ij} by sums of integrals over $(r_{ij})^2 \exp[-\alpha_k r_{ij}^2]$ via the replacement

$$r_{ij} \approx \sum_k c_k e^{-\alpha_k (r_{ij})^2} (r_{ij})^2. \quad (6.12)$$

In itself (6.12) is a very poor approximation, but we need this approximation only under integral signs. This explains somehow why ‘‘Gaussian geminals’’ [18, 19] work much better than one would have anticipated.

The best energy of the He ground state obtained with Gaussian geminals is, by the way, in error by $\sim 45 \mu\text{h}$ [19]. This relatively high accuracy (which is somewhat intermediate between those achievable by standard CI and by wave functions with linear r_{12} terms) has only been possible by very careful optimization of the non-linear parameters α_k .

7. Conclusions

The slow convergence of the partial wave expansion (and hence the CI expansion) is a direct consequence of the Coulomb-singularity of the electron interaction. The convergence is, on the whole, not so bad (the norm as well as the expectation values of most interesting operators converge as l^{-6}), but it is very poor in the neighborhood of the coalescence of the coordinates of two electrons, and this region is very critical if the wave functions has to be multiplied by r_{ij}^{-1} , such that the energy increments E_l only converge as l^{-4} .

The l -expansion has an asymptotic expansion in powers of $(l + \frac{1}{2})^{-1}$, of which the leading terms can be summed in closed form yielding a contribution $\frac{1}{2} r_{12} \Phi$ to the wave function. If one formulates the wave function as

$$\Psi = (1 + \frac{1}{2} r_{12}) \Phi + \chi \quad (7.1)$$

then the correction χ has a very rapidly converging l -expansion, since the effective interaction U_{12} by which χ is coupled to Φ (in perturbation theory) has no singularity at $r_{12} = 0$ (it even vanishes for $r_{12} = 0$). The residual energy increments \tilde{E}_l go asymptotically essentially as l^{-8} , such that χ serves mainly to correct for the partial waves with low l .

The leading term $(0(l+\frac{1}{2})^{-4})$ of the conventional l -expansion is entirely due to the attempt to express the number $\frac{1}{2}$ by its partial wave expansion

$$\frac{1}{2} = \frac{1}{2} \sum_{l=0}^{\infty} \langle (r_{12})_l (r_{12}^{-1})_l \rangle. \quad (7.2)$$

The smaller coefficient of this term in CI than in $E^{(2)}$ results from an energy denominator in the definition of the overhead contribution to the energy.

The energy increments \tilde{E}_l of a variational CI-like calculation with the ansatz (7.1) converge significantly faster than the corresponding increments $\tilde{E}_l^{(2)}$ of second-order perturbation theory. The main reason for this is that in CI the χ_l are coupled to Φ via the effective interaction $U_{12} - \frac{1}{2}\Delta E r_{12}$ in which two contributions cancel each other to a large extent.

The "best" uncorrelated reference function Φ is rather close to the eigenfunction of the bare nuclear Hamiltonian, with a little shielding, much less than in a Hartree-Fock like wave function. The factor $(1 + \frac{1}{2}r_{12})$ automatically takes care of most of the necessary shielding.

The slow convergence of the CI expansion is not only due to the $(l+\frac{1}{2})^{-4}$ dependence of the l -expansion, but for each individual l the convergence in terms of configuration states is very slow, due to a discontinuity of V_1 given by (4.3) for $r_1 = r_2$ (which does not exist in the full V and is an artifact of the l -expansion). It appears that in our approach we have also avoided this second source of the slow convergence of CI, and that we do need to use functions of $r_<$ and $r_>$ [21] in order to get well converged partial wave increments.

In order to generalize the results of the present analysis to larger systems one must keep in mind that the introduction of the term $\frac{1}{2}r_{12}$ serves to allow a closed summation of the slowly converging contributions to the partial wave expansion. For those contributions which have a rapidly converging l -expansion there is no objection against expanding $\frac{1}{2}r_{12}\Phi$ in a traditional Hilbert space basis.

Acknowledgements. The author thanks V. Staemmler for valuable comments, and H. M. Schmidt and H. v. Hirschhausen (Berlin) for their unpublished results quoted in Table 1. The numerical calculations of this paper were done on the CYBER 855 and 205 of the computer center of the Ruhr-Universität.

References and footnotes

1. For a review of the CI method see Shavitt, I. in *Modern Theoretical Chemistry, III, Methods of Electronic Structure Theory* (H. F. Schaefer III, Ed.), New York; Plenum, 1977. Many landmark papers on the CI method are reviewed in H. F. Schaefer III, *Quantum Chemistry*, Clarendon, Oxford 1984

2. It is not sufficient that the basis is complete with respect to an ordinary Hilbert space norm, but with respect to some Sobolev norm, the so-called convergence in energy space, as discussed by Klahn, B., Bingel, W. A.: *Theoret. Chim. Acta (Berl.)* **44**, 9, 27 (1977)
3. Kutzelnigg, W.: *Theoret. Chim. Acta* **1**, 327 343 (1963); Ahlrichs, R., Kutzelnigg, W.: *J. Chem. Phys.* **48**, 1819 (1968)
4. Edmiston, C., Krauss, M.: *J. Chem. Phys.* **45**, 1833 (1966)
5. For reviews see Meyer, W.: in *Modern Theoretical Chemistry III, Methods of Electronic Structure Theory* (H. F. Schaefer, III Ed., New York: Plenum 1977 and Kutzelnigg, W. *ibidem*)
6. Bender, C. F., Davidson, E. R.: *J. Chem. Phys.* **70**, 2675 (1966)
7. Carrol, D. P., Silverstone, H. J., Metzger, R. M.: *J. Chem. Phys.* **71**, 4142 (1979); see also Bunge, C. F.: *Theoret. Chim. Acta (Berl.)* **16**, 126 (1970); and Weiss, A.: *Phys. Rev.* **122**, 1826 (1961)
8. Schwartz, C.: *Phys. Rev.* **126**, 1015 (1962); *Meth. Comp. Phys.* **2**, 241 (1963)
9. Lakin, W.: *J. Chem. Phys.* **43**, 2954 (1965)
10. Kato, T.: *Commun. Pure Appl. Math.* **10**, 151 (1957); see also Pack, R. T., Byers-Brown, W.: *J. Chem. Phys.* **45**, 556 (1966)
11. Hylleraas, E. A.: *Z. Phys.* **54**, 347 (1929), *Z. Phys.* **65**, 209 (1930)
12. Kinoshita, T.: *Phys. Rev.* **105**, 1490 (1957)
13. Pekeris, C. L.: *Phys. Rev.* **112**, 1649 (1958); **115**, 1216 (1959); **126**, 1470 (1962); Frankowski, K., Pekeris, C. L.: *Phys. Rev.* **146**, 46 (1966)
14. Larson, S.: *Phys. Rev.* **169**, 49 (1968); **A6**, 1786 (1972); see also Perkins, J. F.: *Phys. Rev.* **A5**, 514 (1972) **A8**, 700 (1973)
15. Sims, J. S., Hagstrom, S.: *Phys. Rev.* **A4**, 908 (1971), **11**, 418 (1975); *Int. J. Quantum Chem.* **9**, 149 (1975); Sims, J. S., Hagstrom, S., Rumble, J. R.: *Int. J. Quantum Chem.* **10**, 853 (1976)
16. Clary, D. C., Handy, N. C.: *Chem. Phys. Letters* **51**, 483 (1977); Clary, D. C.: *Mol. Phys.* **34**, 793 (1977)
17. Muszyńska, J., Papierowska, D., Woźnicki, W.: *Chem. Phys. Letters* **76**, 136 (1980); Preiskorn, A., Woźnicki, W.: *Chem. Phys. Lett.* **86**, 369 (1982); *Mol. Phys.* **52**, 1291 (1984)
18. Longstaff, J. V., Singer, K.: *Theoret. Chim. Acta (Berl.)* **2**, 265 (1964); Lester, W. A., Krauss, M.: *J. Chem. Phys.* **41**, 1407 (1964); King, H. F.: *J. Chem. Phys.* **46**, 705 (1967); Pan, K. C., King, H. F.: *J. Chem. Phys.* **53**, 4397 (1970); **56**, 4667 (1972); Handy, N. C.: *Mol. Phys.* **26**, 169 (1973); Salmon, L., Poshusta, R. D.: *J. Chem. Phys.* **59**, 3497 (1973); Adamowitz, L., Sadlej, A.: *J. Chem. Phys.* **67**, 4298 (1977); **69**, 3992 (1978)
19. Jeziorski, B., Szalewicz, K.: *Phys. Rev.* **A19**, 2360 (1979); Szalewicz, K., Jeziorski, B., Monkhorst, H. J., Zabolitzky, J. G.: *J. Chem. Phys.* **78**, 1420 (1983); **79**, 5543 (1983); **81**, 368 (1984); *Chem. Phys. Letters* **91**, 169 (1982)
20. Green, L. C., Mulder, M. M., Milner, P. C.: *Phys. Rev.* **91**, 35 (1953)
21. Schmidt, H. M., v. Hirschhausen, H.: *Phys. Rev.* **A28**, 3179 (1983); see also Byron, F. W., Joachain, C. J.: *Phys. Rev.* **157**, 1 (1967)
22. Chandrasekhar, S.: *Astrophys. J.* **100**, 176 (1944)
23. Grein, F., Tseng, T. J.: *Chem. Phys. Lett.* **7**, 506 (1970); see also Hultgren, G. O., Kern, C. W.: *Chem. Phys. Lett.* **10**, 233 (1971); and Löwdin, P. O., Redei, L.: *Phys. Rev.* **114**, 752 (1959)
24. Boys, S. F.: *Proc. Roy. Soc.* **A309**, 195 (1968); Boys, S. F., Handy, N. C.: *Proc. Roy. Soc.* **A310**, 43, 63, **A311**, 309 (1969)
25. Scherr, C. W., Knight, R. E.: *Revs. Mod. Phys.* **35**, 436 (1963)
26. Midtal, J.: *Phys. Rev.* **138**, 1010 (1965)
27. A recent discussion on various definitions of the partial wave increments can be found in: Jankowski, K., Zaharewicz, D. W., Silverstone, H. J.: *J. Chem. Phys.* **82**, 1969 (1985)
28. Kutzelnigg, W.: *Isr. J. Chem.* **19**, 193 (1980); Schindler, M., Kutzelnigg, W.: *J. Chem. Phys.* **76**, 1919 (1982)
29. Wallmeier, H., Kutzelnigg, W.: *Chem. Phys. Letters* **78**, 341 (1981); Wallmeier, H.: *Phys. Rev.* **A29**, 2993 (1984)