Theoret. chim. Acta (Berl.) 2, 265—280 (1964)

Department of Chemistry Royal Holloway College
(University of London)
Englefield Green Survey

A Survey of Gaussian Two-electron Functions

I. The Ground State of the Helium Atom

By

J. V. L. Longstaff and K. Singer

The energy of the ground state of helium is calculated according to the variation method with trial functions formed from a number of different types of linear combinations of Gaussian functions. These functions are compared with respect to a) their convergence towards the exact wave function with increasing number of linear terms, b) the effectiveness of different kinds of correlation functions c) numerical labour.

Die Energie des Grundzustands des Heliumatoms wurde gemäß der Variationsmethode mit Hilfe einer Anzahl verschiedener Typen von Linearkombinationen Gaussscher Funktionen berechnet. Diese Funktionen werden bezüglich a) ihrer Konvergenz zur exakten Wellenfunktion mit zunehmender Zahl von linearen Gliedern, b) der Wirksamkeit verschiedener Korrelationsfunktionen, c) des numerischen Aufwands verglichen.

L'énergie de l'état fondamental de l'atome d'hélium a été calculée selon la méthode de variations avec divers types de combinaisons linéaires des fonctions gaussiennes. Des comparaisons sont faites concernant a) la convergrence de ces fonctions avec le nombre de membres linéaires vers l'exacte fonction ondulatoire, b) l'efficacité de diverses fonctions de corrélation, c) le travail numérique.

1. Introduction

It is well known that the evaluation of quantum mechanical molecular integrals is greatly facilitated by the use of linear combinations of Gaussian functions (l. c. G. f.) or of Gaussian functions multiplied by polynomials [*1*]; and it has been shown that this is true even when these functions contain exponential correlation terms of the form $\lambda\, \boldsymbol{r}_i \cdot \boldsymbol{r}_j$ or $\lambda\, r_{ij}^2$ ($\boldsymbol{r}_i$, $\boldsymbol{r}_j$ being the position vectors of the i-th and j-th electrons) [*2*, *12*].

The expansion of atomic orbitals in terms of l. c. G. f. 's is of interest for several reasons: 1. molecular orbitals are usually formed from atomic orbitals, 2. a standard method for the evaluation of molecular integrals is based on such expansions (cf. [*10*]), 3. the use of l. c. G. f. 's in atomic problems can serve as model experiments for applications in molecular problems.

Some information about the expansion of hydrogen-like atomic orbitals into l. c. G. f. 's is available [*7*, *9*, *14*] and some work on helium and beryllium, involving Gaussian functions has been reported [*9*]. The usefulness of Gaussian basis functions in molecular problems has been demonstrated by the excellent results obtained for methane [*6*].

The classical papers of Hylleraas [*3*, *4*] and of James and Coolidge [*5*] have shown that accurate calculations of electronic energies almost certainly require the use of 'correlated' pair electron functions; and there has been in recent years a revival of interest in possible applications of such functions in many-electron systems (cf. [8]). This paper reports the first part of a fairly extensive survey of various types of Gaussian two-electron wave functions; the work was undertaken as a preliminary study to the possible use of such functions in larger molecular systems. Some encouraging calculations for the ground state of the hydrogen molecule by means of Gaussian two-electron functions have been reported [7] and an extension of this work (unpublished) has led to fairly accurate wave functions ($D_{\text{H-H}} = 4.72$ eV). This paper deals with the ground state of helium; a similar survey for the ground state of the hydrogen molecule will be published in due course.

2. The choice of functions

Variational calculations, *i. e.* minimisations of the energy with respect to all parameters were carried out for functions belonging to the classes listed in Tab. 1. The selection of these functions was based on the following considerations: type I a contains the largest number of exponential parameters (a_i, b_i, g_i) for a given number of linear terms, or, conversely, the smallest number of linear terms for a given number of exponential parameters. Type II a is the other extreme, *i. e.* the largest possible number of linear terms (excluding polynomial factors) is formed from a given number of exponential parameters (a_i or a_i and g_i). Calculations with a function of this type have been reported by REEVES [9].

The optimisation of exponential parameters is much more laborious than that of the linear parameters; one would therefore expect that I a might be computationally advantageous when the optimum values of the exponential parameters are known while II a might be superior when they have to be optimised. Types I b and II b are derived from I a and II a by replacement of the exponential correla-

Table 1

I a (nC, na, nb, ng)	$\sum_{k=1}^{n} C_k (1 + O_{12}) \exp(-a_k r_1^2 - b_k r_2^2 - 2g_k \boldsymbol{r}_1 \cdot \boldsymbol{r}_2)$
I b (nC, nG, na, nb)	$\sum_{k=1}^{n} (C_k + G_k \boldsymbol{r}_1 \cdot \boldsymbol{r}_2)(1 + O_{12}) \exp(-a_k r_1^2 - b_k r_2^2)$
I c (nC, nG, na, nb)	$\sum_{k=1}^{n} (C_k + G_k r_{12}^2)(1 + O_{12}) \exp(-a_k r_1^2 - b_k r_2^2)$
II a (NC, ma, Ng)	$\sum_{k=1}^{m} \sum_{l \geqslant k}^{m} C_{kl} (1 + O_{12}) \exp(-a_k r_1^2 - a_l r_2^2 - 2g_{kl} \boldsymbol{r}_1 \cdot \boldsymbol{r}_2)$
II b (NC, NG, ma)	$\sum_{k=1}^{m} \sum_{l \geqslant k}^{m} (C_{kl} + G_{kl} \boldsymbol{r}_1 \cdot \boldsymbol{r}_2)(1 + O_{12}) \exp(-a_k r_1^2 - a_l r_2^2)$
III a (m, m, s)	$(1 + O_{12}) \left[\sum_{k=1}^{m} C_k \exp(-a_k r_1^2)\right] \left[\sum_{l=1}^{m} D_l \exp(-b_l r_2^2)\right] \left[\sum_{i=1}^{s} G_i \exp(-2g_i \boldsymbol{r}_1 \cdot \boldsymbol{r}_2)\right]$

$O_{12} =$ operator exchanging the coordinates of electrons 1 and 2

$N = m(m + 1)/2$

tion factors exp $(-2g\,\boldsymbol{r}_1\cdot\boldsymbol{r}_2)$ by polynomial correlation terms $1+\lambda\,\boldsymbol{r}_1\cdot\boldsymbol{r}_2$ $(\lambda = G/C)$, involving a change from exponential to linear correlation parameters. I b and II b are linear combinations of products of one-electron functions which can be transformed into pair functions of the form $\sum_i \varkappa_i\,\chi_i(\boldsymbol{r}_1)\,\chi_i(\boldsymbol{r}_2)$ where the χ_i are mutually orthogonal. The fact that the types I a, II a (and III a) are not sums of products of one-electron functions is likely to limit their application in many-electron systems on account of the non-orthogonality problem.

Type I c differs from I b by having correlation factors of the form $1+\lambda\,r_{12}^2$ instead of $1+\lambda\,\boldsymbol{r}_1\cdot\boldsymbol{r}_2$. III a was (rather incompletely) examined because of its analogy to the ordinary $(1s'\,1s'')$-function for helium. Here the parameters of the l. c. G. f. approximations to the H-atom $1s$-function should, when correctly scaled, be close to the optimum parameter values of the helium III a-function.

Although correlation factors of the form exp $(-\lambda\,\boldsymbol{r}_1\cdot\boldsymbol{r}_2)$ and $1+\lambda\,\boldsymbol{r}_1\cdot\boldsymbol{r}_2$ are known to be less effective than either exp $(-\lambda\,r_{12})$ or $1+\lambda\,r_{12}$, functions incorporating the latter types of correlation factors were not examined because their use in molecular problems would lead to serious numerical difficulties.

To facilitate reference in the text, any particular function belonging to the classes of Tab. 1 is denoted by its type (e. g. I a, II b, etc.) followed by a bracket listing the numbers of each kind of parameter. Thus the function

$$\sum_{i=1}^{12} C_i\,(1+O_{12})\,\exp\,(-a_i\,r_1^2 - b_i\,r_2^2)$$

will be referred to as I a $(12C,\ 12a,\ 12b)$ or as I a $(12C,\ 12a,\ 12b,\ 0g)$. Similarly, I b $(6C,\ 6G,\ 6a,\ 6b)$ is the function

$$\sum_{i=1}^{6} (C_i + G_i\,\boldsymbol{r}_1\cdot\boldsymbol{r}_2)\,(1+O_{12})\,\exp\,(-a_i\,r_1^2 - b_i\,r_2^2)$$

and II a $(21C,\ 6a,\ 1g)$ refers to

$$\sum_{k=1}^{6}\sum_{l\geqslant k}^{6} C_{kl}\,(1+O_{12})\,\exp\,(-a_k\,r_1^2 - a_l\,r_2^2 - 2g\,\boldsymbol{r}_1\cdot\boldsymbol{r}_2)$$

"$2g$" in II a $(21C,\ 6a,\ 2g)$ implies that $g_{kl} = g_1$ or g_2 depending on the value of the index kl. (This specification is incomplete.) Reference will also be made to I a $(0g)$, *i. e.* the I a-series without correlation parameters, or to I a (ng) the I a-series with the maximum number of n correlation parameters, etc.

3. Results and discussion

The results obtained after minimising the energy with respect to all parameters (cf. section 4) are listed in Tab. 2. Column 2 contains the energies in modified atomic units $(R_{\mathrm{He}}hc = 0.5$ a. u.), column 3 the error $\Delta = E - E_{\mathrm{exp}}$. Columns 4,5 and 6 contain data which are relevant to the numerical labour required to obtain the results (see section 3d).

a) Convergence towards the exact wave function

The convergence of the energies (or errors) to limiting values, E_∞ (or Δ_∞) with increasing numbers of linear terms (N) is illustrated in Fig. 1. The data are sufficient to indicate that Δ_∞ for the maximally correlated I a – functions

[*Ia* (*ng*)] is very small – certainly less than 0.001 a. u.: the best calculated energy for $N = n = 16$ corresponds to $\Delta = 0.0014$ and Fig. 1 shows that there is no sign of a change of slope of the linear dependence of log Δ on log$N$ at this point. The apparent absence of a finite limiting error is a little surprising since Gaussian functions have an analytically incorrect behaviour, *i. e.* there is no cusp at $r = 0$. For the I a- and II a-functions without correlation parameters [I a (0*g*) and II a (0*g*)] Δ_∞ appears to be close to 0.025; the best calculated energies, –2.87835 for I a (16*C*, 16*a*, 16*b*, 0*g*) and –2.87833 for II a (21*C*, 6*a*, 0*g*) lie within 0.0007 a. u. of the "*S*-limit", *i. e.* the limiting energy obtained by superposition of hydrogen-like open shell functions of spherical symmetry by Shull and Löwdin [*11*].

Table 2

function	E (a. u.)	$\Delta = E\text{-}E_{\exp}$	e. m. e.	r. e. m. e.	t. r./c
I a (4*C*, 4*a*, 4*b*, 0*g*)	–2.86327	0.04045	10	4	32
I a (8*C*, 8*a*, 8*b*, 0*g*)	–2.87651	0.02721	36	8	128
I a (12*C*, 12*a*, 12*b*, 0*g*)	–2.87767	0.02605	78	12	288
I a (16*C*, 16*a*, 16*b*, 0*g*)	–2.87835	0.02537	136	16	512
I a (12*C*, 12*a*, 12*b*, 1*g*)	–2.88921	0.01451	78	12	300
I a (12*C*, 12*a*, 12*b*, 2*g*)	–2.89375	0.00997	78	12	312
I a (3*C*, 3*a*, 3*b*, 3*g*)	–2.86422	0.03950	6	3	27
I a (4*C*, 4*a*, 4*b*, 4*g*)	–2.88073	0.02299	10	4	48
I a (8*C*, 8*a*, 8*b*, 8*g*)	–2.89808	0.00564	36	8	192
I a (12*C*, 12*a*, 12*b*, 12*g*)	–2.90059	0.00313	78	12	432
I a (16*C*, 16*a*, 16*b*, 16*g*)	–2.90233	0.00139	136	16	768
I b (3*C*, 3*G*, 3*a*, 3*b*)	–2.86516	0.03856	21	11	66
I b (6*C*, 6*G*, 6*a*, 6*b*)	–2.89463	0.00909	78	23	276
I b (12*C*, 12*G*, 12*a*, 12*b*)	–2.89959	0.00415	300	47	1128
I c (6*C*, 6*G*, 6*a*, 6*b*)	–2.89521	0.00851	78	23	276
I c (12*C*, 12*G*, 12*a*, 12*b*)	–2.89976	0.00396	300	47	1128
II a (6*C*, 3*a*, 0*g*)	–2.85121	0.05251	13.5	10	30
II a (10*C*, 4*a*, 0*g*)	–2.87185	0.03187	38	24.5	98
II a (15*C*, 5*a*, 0*g*)	–2.87675	0.02697	87.5	49.5	247.5
II a (21*C*, 6*a*, 0*g*)	–2.87833	0.02539	175.5	88	528
II a (6*C*, 3*a*, 1*g*)	–2.86339	0.04033	13.5	10	40
II a (10*C*, 4*a*, 1*g*)	–2.88364	0.02008	38	24.5	122.5
II a (15*C*, 5*a*, 1*g*)	–2.88850	0.01522	87.5	49.5	297
II a (21*C*, 6*a*, 1*g*)	–2.88991	0.01381	175.5	88	616
II a (6*C*, 3*a*, 2*g*)	–2.86916	0.03456	13.5	10	50
II a (10*C*, 4*a*, 2*g*)	–2.88681	0.01691	38	24.5	147
II a (15*C*, 5*a*, 2*g*)	–2.89279	0.01093	87.5	49.5	346.5
II a (21*C*, 6*a*, 2*g*)	–2.89546	0.00826	175.5	88	704
II a (6*C*, 3*a*, 6*g*)	–2.86936	0.03436	13.5	10	90
II a (10*C*, 4*a*, 10*g*)	–2.89038	0.01334	38	24.5	343
II a (15*C*, 5*a*, 15*g*)	–2.89550	0.00822	87.5	49.5	990
II b (6*C*, 1*G*, 3*a*)	–2.86386	0.03986	49.5	37.5	112.5
II b (6*C*, 6*G*, 3*a*)	–2.87201	0.03171	49.5	37.5	112.5
II b (10*C*, 10*G*, 4*a*)	–2.89293	0.01079	144	94.5	378
III a (4, 4, 0)	–2.86809	0.03563	136	64	512
III a (4, 4, 1)	–2.88442	0.01930	136	64	576

$E_{\exp} = -2.90372$; e. m. e. = number of equivalent matrix elements; r. e. m. e. = number of equivalent matrix elements recalculated when an exponential parameter is varied; t. r./c = number of equivalent matrix elements recalculated during one cycle.

Although no reliable extrapolation can be carried out for the series with linear correlation parameters [I b (nC, nG, na, nb), II b (NC, NG, ma)] it is clear from Fig. 1 that I b (nG) does not converge as well as I a (ng) and that $\varDelta_\infty$ for I b (nG) is probably less than 0.0025 ($E \leq -2.9012$). I b (nG) is slightly better than I a (ng) at $n = 3$ and slightly worse at $n = 12$, the differences being $\approx$ 0.001 in both cases.

As to the II a-functions with correlation parameters, $\varDelta_\infty$ for II a ($1g$) appears to be approx. 0.013 ($E = -2.891$). Graphical estimation of $\varDelta_\infty$ for II a ($2g$) from Fig. 1 is not possible; the best calculated energy [II a ($21C$, $6a$, $2g$)] is -2.89546, *i. e.* $\varDelta = 0.00825$ and $\varDelta_\infty$ is probably less than 0.006.

For the II a (NC, ma, Ng) function ($N = m(m+1)/2$, maximally correlated) the decrease of $\log \varDelta$ with $\log N$ is — as one would expect — slower than for I a (ng). Although $\varDelta_\infty$ cannot be estimated from the diagram, it is not likely that its value is larger than that for I a (ng); the energies for I a ($0g$) and II a ($0g$) tend to the same limit (see above); and since for II a (Ng) the ratio of correlation parameters to other parameters is larger than for I a (ng), it must be assumed that the g-correlation energy in the former will be at least as effectively allowed for as in the latter. The limiting behaviour of II a (Ng) functions is, however, of little practical interest since accurate wave functions of this type are disadvantageous on account of numerical labour (cf. section 4).

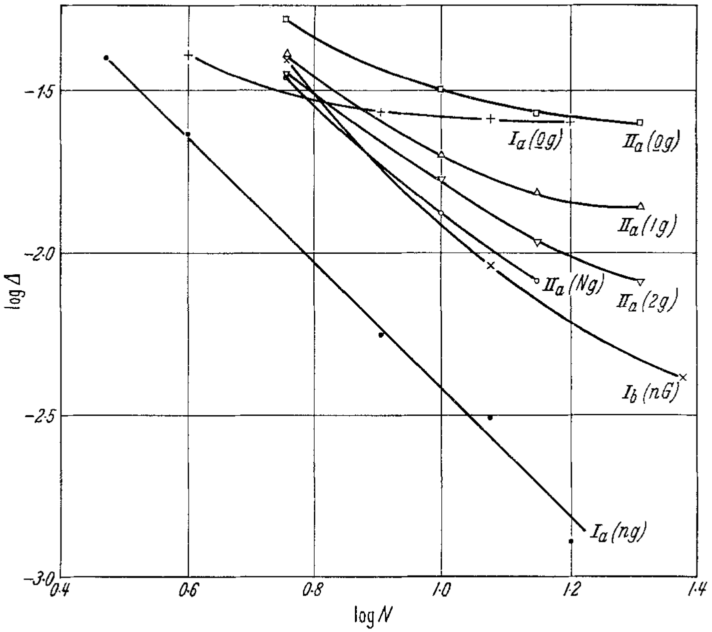

Fig. 1. Relationship between error ($\varDelta$) and the number of independent linear terms (N)

b) Correlation

According to the definition $E_{\text{corr.}} = E$ (Hartree-Fock) $- E_{\text{exp}}$ the correlation energy for the ground state of helium is $-2.86168 + 2.90372 = 0.04204$ a. u. The work of TAYLOR and PARR [*13*] and its extension by other workers (cf. [*8*] ch. 3) has shown that this quantity can be regarded as the sum of "radial" and "angular" correlation energy (approx. 40% and 60% respectively). The radial part can be accounted for by a superposition of configurations of the type (ns' ms'), the angular part by addition of $(2p')^2$-, $(3d')^2$-, and $(4f')^2$- terms, or, alternatively, by addition of (np' mp')- and similar terms for $l = 2, 3$, etc. (the primes indicate a variable scaling factor; m, n are integers).

In the present work the functions without exponential or linear correlation parameters, *i. e.* I a ($0g$), II a ($0g$) correspond to a superposition of split Gaussian orbitals of $1s$-symmetry: $\sum\limits_i \sum\limits_j (1s_i\, 1s_j)$; the contribution to the correlation energy arising from the use of such functions might be called $(1s_i\, 1s_j)$-correlation. In I b and II b terms of the type $(2p_i\, 2p_j)$ are added to the former. The exponential correlation factors in I a and II a (and III a) may, in view of the expansion

$$\exp\left(-\lambda\, \boldsymbol{r}_1 \cdot \boldsymbol{r}_2\right) = \sum_n c_n\, (\boldsymbol{r}_1 \cdot \boldsymbol{r}_2)^n \tag{1}$$

be regarded as equivalent to a superposition of $(1s_i\, 1s_j)$ with $(2p_i\, 2p_j)$, $(3d_i\, 3d_j)$, $\cdots$, *etc.* with fixed expansion coefficients. Since $\boldsymbol{r}_1 \cdot \boldsymbol{r}_2$ depends not only on the angle between the radius vectors but also on their magnitude, factors of the type $\exp(-\lambda\, \boldsymbol{r}_1 \cdot \boldsymbol{r}_2)$ or $1 + \lambda\, \boldsymbol{r}_1 \cdot \boldsymbol{r}_2$ contribute to both angular and radial correlation.

It is of some interest that the $(1s_i\, 1s_j)$-correlation in the simple Gaussian functions

$$\text{I a } (0g){:} \qquad \sum_{k=1}^{n} C_k\, (1 + O_{12}) \exp\left(-a_k r_1^2 - b_k r_2^2\right)$$

$$\text{or II a } (0g){:} \qquad \sum_{k=1}^{m} \sum_{l \geqslant k}^{m} C_{kl}\, (1 + O_{12}) \exp\left(-a_k r_1^2 - a_l r_2^2\right)$$

accounts for nearly as much of the radial correlation (*i. e.* 96%) as superpositions of exponential functions multiplied into Laguerre polynomials [*11*].

Concerning the comparison of polynomial and exponential correlation factors $(1 + \lambda\, \boldsymbol{r}_1 \cdot \boldsymbol{r}_2)$ and $\exp(-\lambda\, \boldsymbol{r}_1 \cdot \boldsymbol{r}_2)$ the following points are relevant: E [I a (ng)] $-$ E [I b (nG)] $= 0.0009$ for $n = 3$ and -0.001 for $n = 12$. Graphical interpolation indicates that the energies for I b (nG) and for I a (ng) are about equal for $n = 7$ ($E = -2.897$) and that the superiority of I a increases with larger values of n. (cf. section 3a). The differences between the energies of corresponding II a (Ng) and II b (Ng) [$N = m\,(m + 1)/2$] is rather larger (0.00275 for $N = 6$, 0.00255 for $N = 10$) and in favour of II b. These results suggest the tentative generalisation that the correlation factors $\exp(-\lambda\, \boldsymbol{r}_1 \cdot \boldsymbol{r}_2)$ contribute more effectively towards angular correlation and less effectively towards radial correlation than the factors $1 + \lambda\, \boldsymbol{r}_1 \cdot \boldsymbol{r}_2$. Thus, where the radial correlation is almost completely accounted for by $(1s_i\, 1s_j)$-correlation without correlation parameters, exponential correlation parameters are more effective; polynomial correlation parameters are better where a significant contribution to the radial correlation remains to be made. The latter situation applies to the less accurate functions ($E > -2.897$).

The energy obtained with I b ($12C$, $12G$, $12a$, $12b$), $E = -2.8996$, is slightly better than TAYLOR and PARR's [*13*] superposition $(1s'\, 1s'') + (2p')^2 + (3d')^2 + (4f')^2$, $E = -2.8975$. This raises the question whether the inclusion of configurations with $l > 1$ to account for angular correlation is as effective as superpositions of the type

$$\sum_i \sum_j (1s_i\, 1s_j) + (2p_i\, 2p_j).$$

Comparison of I b (nG) and I c (nG) shows that correlation factors $1 + \lambda\, r_{12}^2$ are only very slightly better than the computationally simpler $1 + \lambda\, \boldsymbol{r}_1 \cdot \boldsymbol{r}_2$; the energy differences are 0.0006 for $n = 6$ and 0.0002 for $n = 12$.

Since it is desirable to reduce numerical labour trial functions containing less than the maximumn number of correlation parameters are of interest. When the number of g-parameters in II a is changed form 0 to 1 *i. e.* $g_{kl} = g$ instead of 0 for all k, l), the energy is reduced by 0.012 a. u. This improvement is almost constant between $m = 3$ and 6 amounts to a little less than half the residual correlation energy after allowance is made for ($1s_i\, 1s_j$)-correlation. The same improvement (within the limits of error) results when one g-parameter is introduced in I a ($12C$, $12a$, $12b$, $0g$). The corresponding effect of one linear correlation parameter changing II b ($6C$, $0G$, $3a$) to II b ($6C$, $1G$, $3a$), which is equivalent to a multiplication of the former function by $1 + \lambda\, \boldsymbol{r}_1 \cdot \boldsymbol{r}_2$, is only slightly larger (0.0127).

A change in the number of g-parameters in II a from one to two produced a much smaller improvement which increased almost linearly from 0.002 for $m = 3$ to 0.006 for $m = 6$.

c) *Parameter values*

Tables of parameter values are in the appendix. In view of the numerical labour required for the optimisation of exponential parameters any method for the approximate prediction of optimal parameter values would be valuable. No useful regularities have, unfortunately, been hitherto detected in the parameter values of the type I series.

The parameters of the II a series, on the other hand, are a little more predictable. **Reeves** [9] obtained a good estimate for the a-parameters of II a ($6C$, $3a$, $0g$) by appropriate scaling of the exponential parameters of the three-term Gaussian approximation to the $1s$-(H-atom) orbital. This method appears to be inapplicable to the higher members of the II a series ($m = 4, 5, 6$); the a-parameters of the helium-function diverge from the scaled values of the corresponding $1s$-parameters, the former spreading over a wider range than the latter (Tab. 3). It may be noted, in passing, that the range of the helium a-parameters decreases progressively as one and two g-parameters are introduced: the inability of the electrons to avoid each other by being preferentially on opposite sides of the nucleus apperas to lead to a tendency

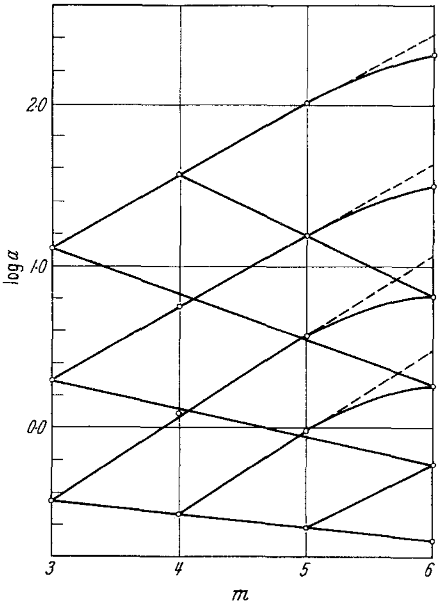

Fig. 2. The variation of the exponential parameters (a) with m for the series II a (NC, ma, $0g$) [$N = m\,(m + 1)\,/2$]. (Ordinate: $\log a_i$, $i = 1$ to m, abscissa m)

towards greater radial (anti-)correlation through a greater separation between the spherical shells.

The fairly regular variation of the a-parameters in II a (NC, ma) with m is illustrated in Fig. 2*. A check calculation with interpolated parameters for II a ($10C$, $4a$, $0g$) gave $E = -2.87082$; a similar calculation with a-parameters determined by *linear* extrapolation (along the dotted lines in Fig. 2) gave 2.87696 for II a ($21C$, $6a$, $0g$). These values differ from the best energies obtained for these functions by, respectively, 0.00103 and 0.00137.

No regularities in the values of the exponential correlation parameters were detected in I a. A surprisingly large number of these parameters is negative *i. e.* give rise to a positive angular correlation**. This feature becomes less paradoxical if one remembers that the matrix elements H_{mn} or S_{mn} depend on the values of g_m and g_n so that there may be compensations between positive and negative $g's$ in different terms; further, the kinetic energy integral contains terms which increase with $g_m + g_n$ and with $g_m\, g_n$ so that the corresponding matrix elements may be appreciably reduced if g_m and g_n have opposite signs.

Table 3

*Comparison of the ratios of a-parameters for Gaussian approximations to the H-atom 1s orbital** $\sum_{i=1}^{n} C_i \exp(-a_i r^2)$ *and for the He-atom* II a*-functions*

	a_2/a_1	a_3/a_1	a_4/a_1	a_5/a_1	a_6/a_1
$H, n = 3$	4.87	36.1			
II a ($6C$, $3a$, $0g$)	5.23	35.0			
II a ($6C$, $3a$, $2g$)	5.18	35.5			
$H, n = 4$	3.43	14.7	98.2		
II a ($10C$, $4a$, $0g$)	4.21	19.2	127		
II a ($10C$, $4a$, $2g$)	4.13	18.5	121		
$H, n = 5$	3.01	9.45	35.6	220	
II a ($15C$, $5a$, $0g$)	3.82	15.4	69.8	453	
II a ($15C$, $5a$, $1g$)	3.66	14.2	64.0	423	
II a ($15C$, $5a$, $2g$)	3.10	10.9	46.7	310	
$H, n = 6$	2.89	6.60	17.2	57.3	242
II a ($21C$, $6a$, $0g$)	2.91	9.19	33.5	150	987
II a ($21C$, $6a$, $1g$)	2.95	9.23	33.1	149	977
II a ($21C$, $6a$, $2g$)	2.81	9.14	32.2	136	875

* LONGSTAFF and SINGER [7]

The correlation parameters of II a and of II b exhibit qualitative features which one might expect: G_{kl} is usually strongly negative when a_k and a_l are both fairly large (> 1) and of the same order of magnitude; *i. e.* if the two electrons are

---

* similar diagrams can be obtained for IIa ($1g$), IIa ($2g$) and IIa (Ng)

** the negative values persist or reappear even after repeated attempts to obtain positive optimum values in fairly small strongly overlapping orbitals. G_{kl} is usually small – positive or negative – when a_k and a_l are small (< 0.4) or when a_k and a_l are of different order of magnitude *i. e.* when the mean distance between the electrons is large even in the absence of $(2p_i\, 2p_j)$-correlation. The exponential correlation parameters follow the same pattern (except for the difference in sign).

d) Numerical labour

A comparison of the practical utility of the functions surveyed in this paper will depend on the answers to the following to questions: A. Which type of function will yield a certain degree of accuracy with the least expenditure of computing time when all parameters are unknown? B. Which type of function will yield a certain degree of accuracy with the least expenditure of computing time when the optimum values of the exponential parameters are known?

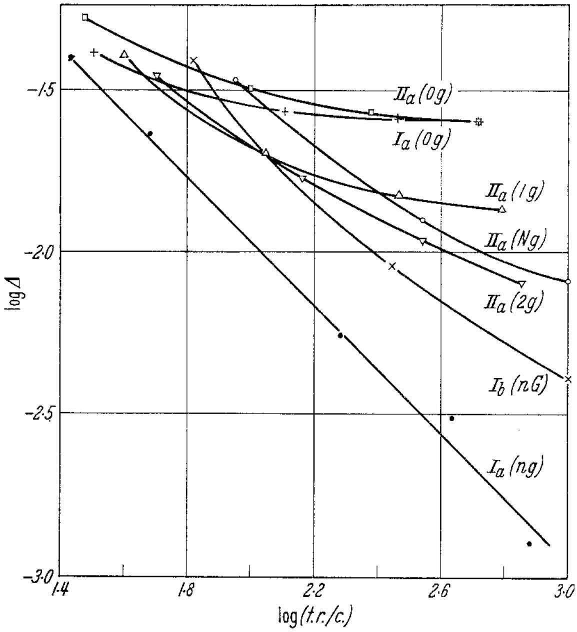

Fig. 3. Relationship between the error (Δ) and the numerical labour required to optimise all parameters (t.r./c)

Column 4 of Tab. 2 contains the "equivalent" numbers of H- or S-matrix elements which have to be calculated to determine the energy. "Equivalent" means that allowance has been made for the fact that for some matrix elements the numerical labour is halved because $\langle 1,2\,|\,H\,|\,1,2\rangle = \langle 1,2\,|\,H\,|\,2,1\rangle$. The entries in this column are clearly relevant to question B (computation when exponential parameters are known). When all exponential parameters have to be optimised by minimisation of the energy with respect to each parameter in turn, it is only necessary to recalculate those matrix elements which contain the parameter which is being varied. The numbers of equivalent matrix elements which have to be recalculated at each step of the minimisation (r. e. m. e.) is listed in column 5 of Tab. 2. If these numbers are multiplied by the number of exponential parameters, one obtains – apart from a common factor of 3 – the total number of matrix elements which have to be recalculated during one minimising cycle (t. r./c) (column 6 of Tab. 2). These entries provide the information required to answer question A (computation when all parameters are unknown) if the following assumptions are made: 1. the time required to minimise the energy with respect to the linear parameters is relatively small; 2. the time required to evaluate an equivalent matrix element and to store the matrix element in its appropriate matrix position does not depend on the type of function or the number of its linear terms; 3. the number of minimising cycles required to attain the final energy is independent of the type of function and of the number of its linear terms. Although these assumptions are certainly incorrect, it is unlikely that a more elaborate analysis would alter the qualitative conclusions drawn from the entries in Tab. 2. Actually observed computing times could have been listed but this would have given undue emphasis to the characteristics of the computer and the computer programmes used in this work.

Fig. 3, where $\log \varDelta$ is plotted against log (t. r./c), shows that question A can be answered in this sense: If the energy is to be calculated with an error of $\varDelta <$ 0.04 a. u. I a (*ng*) is superior to all other types listed in Tab. 1; this is so in spite of the relatively large number of exponential parameters which have to be optimised in this type of function. For $\varDelta < 0.02$ I b (*nG*) comes second (although it is possible that II a (*Ng*) may be better at $\varDelta < 0.00025$). When a smaller accuracy, *i. e.* $\varDelta > 0.02$ is sufficient, II a (*1g*) and II a (*2g*) are superior to I b (*nG*), and at even larger errors ($\varDelta > 0.03$) *Ia* (*0g*) is second best.

For $\varDelta < 0.03$ II a (*2g*) becomes increasingly better than II a (*1g*) but the maximally correlated II a (*Ng*) does not overtake II a (*2g*) until $\varDelta \leq 0.008$.

A plot of $\log \varDelta$ against log (e. m. e.) (not reproduced in this paper) shows that the answer to question B is similar to that to question A; the superiority of Ia (*ng*) is enhanced in this case because the I a-functions contain the relatively smallest number of linear parameters.

III a is the least favourable type of trial function. For III a (4, 4, 1), $E = -2.884$; this is well above the (interpolated value of) $E = -2.892$ for I a (*5C*, *5a*, *5b*, *5g*). The numerical labour required for III a (4, 4, 1) a) to calculate the initial energy, b) to complete one minimising cycle, exceeds the corresponding quantities for *Ia* (*5g*) by factors of 9 and 7.5 respectively. The analogous factors when III a (4, 4, 1) is compared with II a (*10C*, *4a*, *2g*) ($E = -2.887$) are 4 and 3.9.

It should be remembered that the above comparisons apply only to two-electron systems and that the additional features arising in many-electron systems strongly favour I b and II b against, respectively, I a and II a (cf. p. 4 above).

4. Computational methods

The evaluation of all integrals is elementary. Optimisation of the values of the exponential parameters was effected by minimisation of the energy with respect to each parameter in turn; from the E-values corresponding to three adjacent equidistant parameter values the minimum was calculated by parabolic interpolation or extrapolation. When this had been done for all exponential parameters an extrapolation procedure was employed: let values of the parameters determine a many dimensional vector $\boldsymbol{a}$; at the end of a cycle $\boldsymbol{a}$ has changed to $\boldsymbol{a} + \delta \boldsymbol{a}$; the energy is now calculated for parameter values $\boldsymbol{a} + 2\,\delta \boldsymbol{a}$, if E has further decreased, it is calculated for $\boldsymbol{a} + 3\,\delta\,\boldsymbol{a}$ *etc.* until j is found such that $\boldsymbol{a} + j\,\delta\,\boldsymbol{a}$ is a minimum. Another minimising cycle is then carried out followed by extrapolation and this sequence is continued until further optimisation has become unprofitable. In most cases the criterion adopted was that an additional cycle and extrapolation should not reduce E by more than 10^{-5} a. u. The extrapolation procedure greatly improved the rate of convergence towords the final energy.

Linear parameters were in most cases minimised by matrix methods; a) first

E was adjusted so as to make the determinant $\| H - ES \| = 0$; b) the linear parameters were determined as the eigenvector of the matrix HS^{-1}. This was done initially and a number of times during each cycle and whenever the energy was calculated during extrapolation.

In some of the computer programmes the linear parameters were optimised in the same manner as the exponential parameters; check calculations in some such cases showed that the final energy did not decrease further by a matrix minimisation with respect to the linear parameters.

In some of the programmes the exponential parameters were automatically scaled by setting $E = -V^2/(4\ TS)$ (V potential energy, T kinetic energy, S overlap sum). The final energy does not appear to depend on whether or not the scaling procedure is used, though the rate of convergence to the final value may be affected.

There remains the important question whether the energies listed in Tab. 2 are close to the absolute minima. To check this by carrying out minimisations starting from widely different sets of initial parameter values requires a great deal of computing time and was only done in one or two cases. It appears that the final energy in the Ia and IIa series is insensitive to the initial choice of parameters, but that this may not be so for Ib and IIb, which contain polynomial factors. Work on more general types of polynomial Gaussian functions for helium (unpublished) indicates that subsidiary minima may prove rather troublesome. The fairly smooth variation of the energy with the number of linear terms for a given type of function shown in Fig. 1 is reassuring inasmuch as it indicates the absence of appreciable fluctuations in the deviations from the "true" energies from one member of the series to another.

Most of the numerical work was done on a Ferranti "Mercury" computer, the remainder on a Ferranti "Atlas" computer.

Acknowledgements. We wish to thank the Department of Industrial and Scientific Research for a grant during the early stages of this work and the University Grants Committee for continued support of the project. We are grateful to the director and staff of the University of London Computer Unit for generous allocation of computing time and for advice concerning programming techniques.

Appendix

Parameter values

	a	b	g	C	G
Ia (*4C*, *4a*, *4b*, *0g*)	.264	.892		1	
	.244	6.56		0.966	
	2.98	.835		2.98	
	.812	29.8		2.67	
Ia (*8C*, *8a*, *8b*, *0g*)	.643	.534		—1	
	.196	.598		1.126	
	.642	113.2		4.58	
	.464	1.109		5.09	
	.224	3.63		2.82	
	.941	3.63		8.01	
	.290	18.1		3.11	
	1.66	14.21		8.61	

	a	b	g	C	G
Ia ($12C$, $12a$, $12b$, $0g$)	.590	.269		1	
	.169	.528		3.94	
	.187	2.59		12.45	
	.254	16.7		17.5	
	.420	156.7		17.8	
	.643	.900		—30.1	
	.527	1.034		55.5	
	.625	5.05		42.3	
	1.51	32.7		47.3	
	3.72	2.87		—206	
	2.71	4.44		214	
	2.08	1.97		38.7	
Ia ($16C$, $16a$, $16b$, $0g$)	.408	.331		1	
	.161	.542		2.24	
	.447	1.213		1.49	
	.179	2.33		7.24	
	.239	12.74		10.25	
	.282	83.5		6.93	
	.609	.815		—19.9	
	.525	1.004		30.8	
	.637	4.58		29.9	
	1.198	32.0		26.2	
	.947	385		13.34	
	1.715	2.10		11.48	
	4.00	4.14		—8.26	
	.503	.769		4.26	
	1.108	1.499		1.330	
	2.57	8.99		30.3	
Ia ($3C$, $3a$, $3b$, $3g$)	.270	.831	.033	1	
	3.09	.631	.1238	3.93	
	.551	19.2	—.016	2.84	
Ia ($4C$, $4a$, $4b$, $4g$)	.274	.892	.033	1	
	.279	5.80	—.044	1.193	
	2.87	.960	.259	2.69	
	.784	29.1	.092	2.59	
Ia ($8C$, $8a$, $8b$, $8g$)	1.138	.946	—.302	—1	
	.206	.617	.0219	.287	
	.437	119.9	—.0040	.713	
	.654	1.303	—.0598	2.09	
	.213	3.20	.0163	.577	
	1.267	4.72	.319	2.16	
	.364	14.50	—.0418	1.021	
	2.40	31.2	.691	1.73	
Ia ($12C$, $12a$, $12b$, $12g$)	.509	.340	—.039	1	
	.1648	.532	.0186	1.78	
	.175	2.57	.0108	5.21	
	.259	16.7	.0087	8.51	
	.439	153	—.0160	9.34	
	.653	.922	—.142	—17.4	
	.539	1.045	—.0358	29.8	

	a	b	g	C	G
	.585	5.02	.0531	20.1	
	1.54	30.5	.456	23.9	
	3.73	2.86	—.766	—91.2	
	2.74	4.38	—.557	95.4	
	1.96	1.94	—.0686	19.5	
I a (16*C*, 16*a*, 16*b*, 16*g*)	.409	.332	.0451	1	
	.161	.542	.0089	1.517	
	.447	1.213	.0234	5.84	
	.179	2.28	—.0048	4.95	
	.246	12.25	.00380	7.63	
	.284	81.3	—.00400	5.11	
	.606	.812	—.0358	—2.74	
	.524	1.002	.0427	10.39	
	.626	4.61	.0431	20.4	
	1.196	30.5	.0772	19.8	
	.950	374	.0050	9.93	
	1.715	2.02	.0377	13.78	
	3.72	3.92	—2.32	—7.89	
	.493	.770	.288	—.370	
	1.041	1.455	—.725	—2.07	
	2.93	8.79	.385	21.5	
I a (12*C*, 12*a*, 12*b*, 1*g*)	.610	.269	.0503	1	
	.168	.530	,,	1.08	
	.196	2.56	,,	4.70	
	.260	16.5	,,	6.49	
	.429	158	,,	6.71	
	.641	.901	,,	—9.44	
	.525	1.035	,,	18.4	
	.609	5.01	,,	15.4	
	1.470	32.5	,,	17.5	
	3.72	2.86	,,	—81.9	
	2.72	4.43	,,	84.6	
	2.05	1.99	,,	15.5	
I b (3*C*, 3*G*, 3*a*, 3*b*)	.273	.861		1	—.0694
	3.23	.642		3.67	—.988
	.531	20.1		2.51	.153
I b (6*C*, 6*G*, 6*a*, 6*b*)	.696	.236		1	—.0459
	.259	3.42		1.76	.161
	.350	16.2		2.05	.193
	1.80	.704		3.24	—1.247
	.737	89.8		2.77	.161
	7.88	1.66		5.07	—6.23
I b (12*C*, 12*G*, 12*a*, 12*b*)	.202	.560		1	—.0339
	.239	2.06		2.77	.0644
	.181	8.24		1.375	—.1092
	.230	76.2		1.207	—.1044
	.995	.598		2.61	—.647
	.777	2.02		2.42	—1.004
	.648	448		2.41	—.546
	5.08	.775		7.01	.569

	a	b	g	C	G
	2.83	2.44		1.72	—7.74
	3.01	13.27		5.41	—17.1
	20.3	.525		4.08	.242
	65.6	1.91		5.04	2.37
I c ($6C$, $6G$, $6a$, $6b$)	.738	.266		1	.0357
	.237	3.49		1.77	—.0389
	.339	17.1		1.98	—.0204
	2.01	.891		2.98	.815
	.730	90.8		2.67	—.0571
	8.42	2.19		5.00	3.64
I c ($12C$, $12G$, $12a$, $12b$)	.220	.569		1	.0264
	.232	2.03		3.16	—.0174
	.237	8.45		1.398	.1345
	.232	81.2		1.432	—.0021
	1.008	.659		2.312	.373
	.900	1.92		2.14	.758
	.857	441		2.95	.479
	4.93	.701		8.20	—.338
	3.17	2.43		1.59	4.30
	4.10	11.9		7.19	13.30
	20.4	.495		4.69	—.1036
	59.1	1.80		6.05	—1.013

For the functions II a and II b the parameters C_{kl}, g_{kl} or G_{kl} are listed in the order $kl =$ 11, 12, 13, · ·, $1m$, 22, 23, · ·, $2m$, etc.

II a ($6C$, $3a$, $0g$)

a .373 1.964 13.19
C .313 1.022 .808 1.091 1.244 —.0883

II a ($10C$, $4a$, $0g$)

a .2907 1.224 5.591 36.89
C .201 1.021 1.089 .636 1.58 2.18 1.392 .154 1.058 .01016

II a ($15C$, $5a$, $0g$)

a .2400 .9157 3.688 16.76 108.84
C .0973 .785 1.010 .635 .360 1.68 2.64 1.74
.986 .338 1.59 .806 —.0158 .234 —.0225

II a ($21C$, $6a$, $0g$)

a .2033 .5914 1.869 6.803 30.55 200.7
C 1 12.58 24.33 21.4 13.80 7.29 28.2 73.5
64.8 39.8 22.5 30.1 86.6 45.1 26.8 —27.97
58.8 13.28 —56.2 32.9 —42.4

II a ($6C$, $3a$, $1g$)

a .3846 1.987 13.32
g .0499
C .617 1.954 1.55 1.99 2.27 .0198

II a (10C, 4a, 1g)

a .3024 1.245 5.707 37.69
g .0488
C .214 1.053 1.111 .650 1.54 2.10 1.332 .190 1.041 −.005

II a (15C, 5a, 1g)

a .2577 .9439 3.664 16.49 108.9
g .0477
C .116 .818 1.053 .692 .404 1.62 2.40 1.72
.870 .452 1.48 .912 −.0235 .0748 −.0360

II a (21C, 6a, 1g)

a .2168 .6397 2.000 7.187 32.26 211.7
g .0494
C 1.262 17.9 31.8 27.0 18.1 9.11 40.7
89.6 79.6 48.8 26.8 37.5 96.5 51.2
31.9 −36.0 70.4 12.7 −72.0 41.5 −52.1

II a (6C, 3a, 2g)

a .3756 1.966 13.12
g for kl = 11, 12, 13 0.0358, for kl = 22, 23, 33 .5015
C .315 1.013 .815 1.160 1.239 −.0339

II a (10C, 4a, 2g)

a .2959 1.197 5.253 33.97
g for kl = 11 to 22 .0401; for 23 to 44 .4248
C .205 1.014 1.104 .695 1.506 2.15 1.400 .431 1.094 −.0373

II a (15C, 5a, 2g)

a .3352 1.037 3.654 15.67 104.6
g for kl = 11 to 24 0.0591; 25 to 55 1.296
C .0434 .450 .922 .583 .469 1.239 1.893
2.256 .744 2.128 1.59 1.92 −.0323 .348 −.00033

II a (21C, 6a, 2g)

a .1827 .5137 1.672 5.885 24.78 160
g for kl = 11 to 26 0.0327; for 33 to 66 0.547
C .0372 9.94 23.03 18.8 13.27 7.75 33.5
78.8 84.0 50.0 29.0 80.6 96.1 74.5
41.5 −8.82 67.4 20.5 −60.2 38.9 −56.0

II a (6C, 3a, 6g)

a .3766 1.969 13.15
g .0321 0.0443 −0.0376 .479 .654 −.015
C .315 1.006 .809 1.146 1.234 −.086

II a (10C, 4a, 10g)

a .2941 1.213 5.439 35.60
g .0238 .0289 −.0283 −.00438 .232 .349
.018 2.051 3.50 2.90
C .204 .999 1.117 .639 1.62 2.04 1.492 .629 .951 −.225

II a (15*C*, 5*a*, 15*g*)

a .2352 .7939 2.796 12.03 80.33
g .0155 .0218 −.0145 −.0119 .161 .120
.194 −.0281 −.292 .804 1.508 −.157
.792 2.32 .990
C .0622 .457 .756 .511 .329 1.069 1.66
1.608 .764 .977 1.215 .956 −.0216 .469 −.00408

II b (6 *C*, 1*G*, 3*a*)

a .393 2.061 13.85
C .174 1.043 .807 .514 1.175 −.0729
G −.1049

II b (6*C*, 6*G*, 3*a*)

a .3757 1.941 13.01
C 1.00 6.44 5.17 3.57 7.92 .138
G −.0747 −.383 .1186 −5.10 −3.79 −52.7

II b (10*C*, 10*G*, 4*a*)

a .2772 1.112 5.014 33.22
C 1.00 11.28 13.14 7.89 10.02 28.3 18.4
2.52 16.4 −4.09
G −.0609 −.318 .134 −.382 −6.16 −6.58
.432 −65.9 1.84 −451

III a (4, 4, 0)

a .450 .834 4.315 40.52
C .771 1.179 .694 1.649
b .2499 1.260 3.657 15.57
D 1.95 .755 2.04 1.58

III a (4, 4, 1)

a .2728 1.002 4.228 52.18
C .1078 .732 1.253 .630
b .260 1.064 6.47 13.15
D 1.316 1.714 −.431 2.29
g .0440
G 1.00

References

[1] Boys, S. F.: Proc. Roy. Soc. (London) **A 200**, 542 (1950).
[2] — Proc. Roy. Soc. (London) **A 258**, 402 (1960).
[3] Hylleraas, E.: Z. f. Physik **54**, 347 (1929).
[4] — Z. f. Physik **65**, 209 (1930).
[5] James, H. M., and A. S. Coolidge, J. chem. Physics **1**, 825 (1933).
[6] Krauss, M.: J. chem. Physics **38**, 564 (1963).
[7] Longstaff, J. V. L., and K. Singer: Proc. Roy. Soc. (London) **A 258**, 421 (1960).
[8] Parr, R. G.: "The Quantum Theory of Molecular Electronic Structure", New York, W. A. Benjamin, (1963).
[9] Reeves, C. M.: J. chem. Physics **39**, 1 (1963).
[10] Shavitt, I. "Methods in Computational Physics", Vol. 2 p.29, New York, Acad. Press (1963).
[11] Shull, H., and P.-O. Löwdin: J. chem. Physics **30**, 617 (1959).
[12] Singer, K.: Proc. Roy. Soc. (London) **A 258**, 412 (1960).
[13] Taylor, G. R., and G. R. Parr: Proc. Nat. Acad. Sci. **38**, 182 (1952).
[14] Whitten, J. L.: J. chem. Physics **39**, 349 (1963).

(Received June 5, 1964)