On the Error of Approximations in Quantum Mechanics II. Some Particular Applications

T. A. Hoffmann

Phil. Trans. R. Soc. Lond. A 1965 **257**, 327-349
doi: 10.1098/rsta.1965.0008

ON THE ERROR OF APPROXIMATIONS IN QUANTUM MECHANICS
II. SOME PARTICULAR APPLICATIONS

By T. A. HOFFMANN*
Aktiebolaget Atomenergi, Studsvik, Sweden

(*Communicated by C. A. Coulson, F.R.S.—Received* 19 *March* 1964)

CONTENTS

In the preceding paper (Hoffmann 1965) the author has developed the general basis of the estimation of the error in a quantum-mechanical calculation. The method based on the properties of the common trace-forming procedure in matrix calculus allowed the determination of the next approximating steps in a steepest descent process. The present paper supplies the results of applying the method to some cases usually applied in quantum-mechanical and quantum-chemical calculations.

1. Introduction

In the preceding paper (Hoffmann 1965), to be denoted in the following by I, the author defined the error of an approximation of a quantum-mechanical calculation. In the most general case defining a Hamiltonian for an n-electron problem, the n-electron density matrix $\mathbf{R}$ has to fulfil equation (I 29)† the equation corresponding to the Schrödinger equation, equation (I 30), the equation assuring the idempotency of $\mathbf{R}$ and equation (I 31), the equation determining the number of states taken into account in the system.

In the following we shall present the explicit form of these equations and the way of approximating them in some special cases.

Let us first investigate how to obtain an explicit form for the Hamiltonian in some simple cases.

---

* On leave of absence from the Research Institute for Telecommunication, Budapest and from the Central Research Institute for Physics of the Hungarian Academy of Sciences, Budapest.

† (In) denotes equation (n) of paper I.

2. INDEPENDENT ONE-ELECTRON EIGENFUNCTIONS

If there are n-electrons without any interaction between them, the Hamiltonian can be obtained as the sum of the one-electron Hamiltonians, say $\mathbf{H}(j)$ for the jth electron, containing the co-ordinates of the jth electron only. So in this case

$$\mathbf{H} = \sum_{j=1}^{n} \mathbf{H}(j). \tag{1}$$

Similarly the kernel function defined in (I 9) can also be obtained as the sum of one-electron functions

$$H(x,y) = \delta(y-x) \sum_{j=1}^{n} H_0(y_j) + \delta'(y-x) \sum_{j=1}^{n} H_1(y_j) + \delta''(y-x) \sum_{j=1}^{n} H_2(y_j), \tag{2}$$

where the symbols x and y stand for the same *set* of *all* co-ordinates as in I, but the variables y_j already denote the specified values of the co-ordinates of the jth electron in the set y, etc.

The eigenfunction specified by a total energy E_i (or in the case of degeneracy one of the orthogonalized eigenfunctions) is built up as the product of n one-electron eigenfunctions having the sum of their eigenvalues just E_i, i.e.

$$\psi_i(y) = \prod_{j=1}^{n} \phi_{k_j}(y_j), \tag{3}$$

where the set of indices, k_j, indicates the one-electron states which the various electrons occupy in the system belonging to energy E_i. If for each electron various states s are allowed, there are altogether s^n distinct states to be considered. Denoting the energy of an electron in state k_j by ϵ_{k_j} we have

$$\mathbf{H}(j)\,\phi_{k_j}(j) \equiv \mathbf{H}(j)\,\phi_{k_j}(y_j) = \epsilon_{k_j}\,\phi_{k_j}(y_j) \equiv \epsilon_{k_j}\,\phi_{k_j}(j), \tag{4}$$

$$E_i = \sum_{j=1}^{n} \epsilon_{k_j}. \tag{5}$$

$\phi_{k_j}(y_j)$ denotes the one-electron eigenfunction of state k_j and of the electron with the co-ordinates y_j. The ϕ one-electron eigenfunctions are regarded as normalized to unity.

Let us now discuss the meaning of the Dirac functions in (2) somewhat more in detail. If we assume the space of one electron to be subdivided in the way explained preceding equation (I 4) into m parts (m is a large number generally and it may tend to infinity), the same subdivision being valid for all electrons, then the number of rows in the matrix (I 4) will be m^n. The Dirac functions occurring in (2) in their symbolical form are explicitly

$$\delta(y-x) = \prod_{j=1}^{n} \delta(y_j - x_j), \tag{6}$$

where x_j and y_j denote the co-ordinates of the jth electron only. Similar representation is also valid for the derivatives of the Dirac function. We may arrange the sequence of the rows of the matrix (I 4) in such a way that the first m rows shall belong to the same fixed points of the electrons $y_2, y_3, \ldots$, but the co-ordinates of the first electron, y_1, take up all their m values in turn. In the second m rows y_1 runs through the same values again as in the first m rows; however, the value of y_2 is now changed to a new constant. Therefore we may locate all the various values of y_1 and y_2—with their proper combinations in the first m^2 rows. The next m^2 rows repeat these values, only the value of y_3 is now another one (but constant) than that in the first m^2 rows. The arrangement is now to be continued in the same manner for all of the electrons.

Now let us define the *inner product* of two matrices in the following way. The inner product of a matrix of the size $p \times r$, say $\mathbf{M}$, with another matrix of the size $q \times s$, say $\mathbf{N}$, is defined as a matrix of the size $pq \times rs$, where each element, m_{ij}, of matrix $\mathbf{M}$ (a number in the usual sense) is replaced by a matrix $m_{ij}\mathbf{N}$. This inner product will be denoted by $\mathbf{M} \circ (\mathbf{N})$.

It is very easy to see, and we shall here give without proof the following properties of the inner product.

1. Let $\mathbf{M}_1$ and $\mathbf{M}_2$ be such matrices that they can be multiplied together and $\mathbf{N}_1$ and $\mathbf{N}_2$ similarly. Let us denote the (common) matrix product of $\mathbf{M}_1$ and $\mathbf{M}_2$ by $\mathbf{M}$ and of $\mathbf{N}_1$ and $\mathbf{N}_2$ by $\mathbf{N}$:

$$\mathbf{M} = \mathbf{M}_1 \mathbf{M}_2 \tag{7}$$

and

$$\mathbf{N} = \mathbf{N}_1 \mathbf{N}_2. \tag{8}$$

Then we have

$$\mathbf{M} \circ (\mathbf{N}) = [\mathbf{M}_1 \circ (\mathbf{N}_1)] \,.\, [\mathbf{M}_2 \circ (\mathbf{N}_2)]. \tag{9}$$

2. If

$$\mathbf{M} = \mathbf{M}_1 + \mathbf{M}_2 \tag{10}$$

and

$$\mathbf{N} = \mathbf{N}_1 + \mathbf{N}_2, \tag{11}$$

then

$$\mathbf{M} \circ (\mathbf{N}) = \mathbf{M}_1 \circ (\mathbf{N}_1) + \mathbf{M}_1 \circ (\mathbf{N}_2) + \mathbf{M}_2 \circ (\mathbf{N}_1) + \mathbf{M}_2 \circ (\mathbf{N}_2). \tag{12}$$

3. If *both* $\mathbf{M}_1$ and $\mathbf{M}_2$, and in addition $\mathbf{N}_1$ and $\mathbf{N}_2$ are commutable, then so is the inner product in the sense, that if

$$\mathbf{M}_1 \mathbf{M}_2 = \mathbf{M}_2 \mathbf{M}_1 \tag{13}$$

and

$$\mathbf{N}_1 \mathbf{N}_2 = \mathbf{N}_2 \mathbf{N}_1, \tag{14}$$

then, e.g.

$$\mathbf{M}_1 \mathbf{M}_2 \circ (\mathbf{N}_1 \mathbf{N}_2) = \mathbf{M}_2 \mathbf{M}_1 \circ (\mathbf{N}_2 \mathbf{N}_1) \tag{15}$$

but not in the case if (13) or (14) is not valid.

4. The trace of the inner product is the product of the traces of the factors i.e.

$$\operatorname{tr}[\mathbf{M} \circ (\mathbf{N})] = \operatorname{tr}\mathbf{M} \,.\, \operatorname{tr}\mathbf{N}. \tag{16}$$

Naturally the inner multiplication is in general a non-commutative operation.

5. If none of the matrices $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{N}_1$ and $\mathbf{N}_2$ is $\mathbf{0}$, and further

$$\mathbf{M}_1 \circ (\mathbf{N}_1) = \mathbf{M}_2 \circ (\mathbf{N}_2) \tag{17}$$

and the sizes of $\mathbf{M}_1$ and $\mathbf{M}_2$ are the same, which already assures that the sizes of $\mathbf{N}_1$ and $\mathbf{N}_2$ are also the same, then it follows that

$$\mathbf{M}_1 = a\mathbf{M}_2 \tag{18}$$

and

$$a\mathbf{N}_1 = \mathbf{N}_2, \tag{19}$$

where a is some number.

6. A matrix $\mathbf{M}$ of size $k \times l$ can always be extended to a matrix of size $pk \times pl$ as follows

$$\mathbf{M}_{\text{ext.}} = \mathbf{M} \circ (\mathbf{I}_{(p)}), \quad \text{or} \quad \mathbf{M}'_{\text{ext.}} = \mathbf{I}_{(p)} \circ (\mathbf{M}), \tag{20}$$

where $\mathbf{I}_{(p)}$ denotes the unit matrix of the size $p \times p$. Naturally, even in this case with the multiplication by the unit matrix, the inner multiplication is not commutative. It is important that the extension performed on the equations should be made in the same way on both sides of it.

7. If
$$\mathbf{M} \circ (\mathbf{N}) = \mathbf{0}, \tag{21}$$
then either
$$\mathbf{M} = \mathbf{0} \tag{22}$$
or
$$\mathbf{N} = \mathbf{0} \tag{23}$$
or both. (This conclusion is *not true* in the case of the *ordinary* matrix multiplication.)

8. The transposition rule is for the inner multiplication
$$[\mathbf{M} \circ (\mathbf{N})]^+ = \mathbf{M}^+ \circ (\mathbf{N}^+). \tag{24}$$

Returning now to our physical problem we see from (3) that the eigenfunction matrix for the system built up of non-interacting electrons can be written in the form
$$\boldsymbol{\psi} = \boldsymbol{\phi}(n) \circ (\boldsymbol{\phi}(n-1) \circ (\boldsymbol{\phi}(n-2) \circ \ldots \circ (\boldsymbol{\phi}(1))\ldots)), \tag{25}$$
where the one-electron matrix $\boldsymbol{\phi}(j)$ denotes the matrix of the size $m \times s$, having in each column the various one-electron eigenfunctions at the various points.

Similarly the matrix with elements (6) can be written in the form
$$\mathbf{I}_{(m^n)} = \mathbf{I}_{(m)} \circ (\mathbf{I}_{(m)} \circ (\mathbf{I}_{(m)} \circ \ldots \circ (\mathbf{I}_{(m)})\ldots)), \tag{26}$$
where we have replaced the symbolical notation with discrete Kronecker symbols by Dirac functions for convenience of more simple notation.

Using now (25), (26), (1) and (6) equation (I 6) may be written by successive applications of rules 1 and 6 given above:
$$\begin{aligned}\sum_{j=1}^{n} \boldsymbol{\phi}(n) \circ (\boldsymbol{\phi}(n-1) \circ \ldots \circ (\mathbf{H}(j)\,\boldsymbol{\phi}(j) \circ (\boldsymbol{\phi}(j-1) \circ \ldots))\ldots) \\ = [\boldsymbol{\phi}(n) \circ (\boldsymbol{\phi}(n-1) \circ \ldots \circ (\boldsymbol{\phi}(l) \circ (\boldsymbol{\phi}(l-1) \circ \ldots))\ldots)]\,.\mathbf{E} \\ = \sum_{j=1}^{n} \boldsymbol{\phi}(n) \circ (\boldsymbol{\phi}(n-1) \circ \ldots \circ (\boldsymbol{\phi}\boldsymbol{\epsilon}(j) \circ (\boldsymbol{\phi}(j-1) \circ \ldots))\ldots),\end{aligned} \tag{27}$$
where s^n is the number of states taken into account, i.e. the matrix $\mathbf{E}$ has the dimension $s^n \times s^n$, and
$$\mathbf{E} = \sum_{j=1}^{n} \overset{1}{\mathbf{I}}_{(s)} \circ (\overset{2}{\mathbf{I}}_{(s)} \circ \ldots \circ (\overset{j-1}{\mathbf{I}}_{(s)} \circ (\overset{j}{\boldsymbol{\epsilon}} \circ (\overset{j+1}{\mathbf{I}}_{(s)} \circ \ldots \circ (\overset{n}{\mathbf{I}}_{(s)})\ldots)))\ldots) \tag{28}$$
is analogous to equation (5) in the sense that the matrix $\boldsymbol{\epsilon}$ of size $s \times s$ is a diagonal matrix having in the diagonals the s one-electron energies.

Since equation (4) holds for all k_j in its matrix form too, i.e.
$$\mathbf{H}(j)\,\boldsymbol{\phi}(j) = \boldsymbol{\phi}(j)\,\boldsymbol{\epsilon}, \tag{29}$$
equation (27) is really fulfilled. It can also be shown, but we do not give here the proof, that equation (27) supplies, with the aid of rules 1 to 8 given above, all the equations (4).

Using rule 8 we have by (25)
$$\boldsymbol{\psi}^{+*} = \boldsymbol{\phi}^{+*}(n) \circ (\boldsymbol{\phi}^{+*}(n-1) \circ (\boldsymbol{\phi}^{+*}(n-2) \circ \ldots \circ (\boldsymbol{\phi}^{+*}(1))\ldots)) \tag{30}$$

and so by definition (I 20) and rule 1

$$\mathbf{R} = \psi\psi^{+*} = \boldsymbol{\phi}(n)\,\boldsymbol{\phi}^{+*}(n) \circ (\boldsymbol{\phi}(n-1)\,\boldsymbol{\phi}^{+*}(n-1) \circ \ldots \circ (\boldsymbol{\phi}(1)\,\boldsymbol{\phi}^{+*}(1))\ldots). \tag{31}$$

Let us now introduce the one-electron density matrix by

$$\mathbf{R}(j) = \boldsymbol{\phi}(j)\,\boldsymbol{\phi}^{+*}(j), \tag{32}$$

where the summation involved in the matrix multiplication extends to the s states of the one-electron matrices. In this manner (31) can be written in the form

$$\mathbf{R} = \mathbf{R}(n) \circ (\mathbf{R}(n-1) \circ \ldots \circ (\mathbf{R}(1))\ldots). \tag{33}$$

Equation (I 29) can now be rewritten by the use of (33), (1), (26) and the rules for the inner multiplication

$$\sum_{j=1}^{n} \mathbf{R}(n) \circ (\mathbf{R}(n-1) \circ \ldots \circ ([\mathbf{H}(j)\,\mathbf{R}(j) - \mathbf{R}(j)\,\mathbf{H}(j)] \circ (\mathbf{R}(j-1) \circ \ldots))\ldots) = \mathbf{0}. \tag{34}$$

Let us take out the last term with $j = n$ from the summation, so that we can write equation (34) in the form

$$[\mathbf{H}(n)\,\mathbf{R}(n) - \mathbf{R}(n)\,\mathbf{H}(n)] \circ (\mathbf{R}(n-1) \circ \ldots \circ ((\mathbf{R}(1))\ldots)$$
$$= -\mathbf{R}(n) \circ \left(\sum_{j=1}^{n-1} \mathbf{R}(n-1) \circ (\mathbf{R}(n-2) \circ \ldots \circ ([\mathbf{H}(j)\,\mathbf{R}(j) - \mathbf{R}(j)\,\mathbf{H}(j)] \circ (\mathbf{R}(j-1) \circ \ldots))\ldots)\right). \tag{35}$$

Rule 5 can be applied now, since the dimensions of the matrices fulfil its conditions. So we obtain

$$\mathbf{H}(n)\,\mathbf{R}(n) - \mathbf{R}(n)\,\mathbf{H}(n) = -a\mathbf{R}(n) \tag{36}$$

where a is a number, and at the same time

$$a\mathbf{R}(n-1) \circ (\mathbf{R}(n-2) \circ \ldots \circ (\mathbf{R}(1))\ldots)$$
$$= \sum_{j=1}^{n-1} \mathbf{R}(n-1) \circ (\mathbf{R}(n-2) \circ \ldots \circ ([\mathbf{H}(j)\,\mathbf{R}(j) - \mathbf{R}(j)\,\mathbf{H}(j)] \circ (\mathbf{R}(j-1) \circ \ldots))\ldots). \tag{37}$$

Now since the matrix $\boldsymbol{\phi}(j)$ forms an orthonormal matrix (partly based on the assumption made after equation (5), partly based on the orthogonality of eigenfunctions belonging to different eigenvalues and partly because in the case of degenerate eigenvalues we prescribe an orthonormalization process),

$$\boldsymbol{\phi}^{+*}(j)\,\boldsymbol{\phi}(j) = \mathbf{I}_{(s)}, \tag{38}$$

quite analogously to equation (I 23). Therefore from (32) we have for all j

$$\mathbf{R}^2(j) = \boldsymbol{\phi}(j)\,\boldsymbol{\phi}^{+*}(j)\,\boldsymbol{\phi}(j)\,\boldsymbol{\phi}^{+*}(j) = \boldsymbol{\phi}(j)\,\boldsymbol{\phi}^{+*}(j) = \mathbf{R}(j). \tag{39}$$

Now let us multiply both sides of equation (36) by $\mathbf{R}(n)$ once from the left and once from the right and use equation (39)

$$\mathbf{R}(n)\,\mathbf{H}(n)\,\mathbf{R}(n) - \mathbf{R}(n)\,\mathbf{H}(n) = -a\mathbf{R}(n) \tag{40}$$

and

$$\mathbf{H}(n)\,\mathbf{R}(n) - \mathbf{R}(n)\,\mathbf{H}(n)\,\mathbf{R}(n) = -a\mathbf{R}(n). \tag{41}$$

Adding (40) and (41) we obtain

$$\mathbf{H}(n)\,\mathbf{R}(n) - \mathbf{R}(n)\,\mathbf{H}(n) = -2a\mathbf{R}(n), \tag{42}$$

from which a comparison with equation (36) gives immediately

$$a\mathbf{R}(n) = 2a\mathbf{R}(n). \tag{43}$$

Since $\mathbf{R}(n)$ does not vanish identically we conclude from this

$$a = 0 \tag{44}$$

and so from (36) we have finally

$$\mathbf{H}(n)\,\mathbf{R}(n)-\mathbf{R}(n)\,\mathbf{H}(n) = \mathbf{0} \tag{45}$$

and from (37)

$$\sum_{j=1}^{n-1}\mathbf{R}(n-1)\circ(\mathbf{R}(n-2)\circ\ldots\circ([\mathbf{H}(j)\,\mathbf{R}(j)-\mathbf{R}(j)\,\mathbf{H}(j)]\circ(\mathbf{R}(j-1)\circ\ldots))\ldots) = \mathbf{0}. \tag{46}$$

Equation (46) is quite analogous to equation (34) except that the nth electron is taken away. The process can be continued and so we obtain finally for *all* electrons

$$\mathbf{H}(j)\,\mathbf{R}(j)-\mathbf{R}(j)\,\mathbf{H}(j) = \mathbf{0}. \tag{47}$$

The original equations (I 29) and (I 30) are thus reduced to the set of independent equations (47) and (39) for all j.

Quite analogous to equation (I 31) we may also obtain an equation for the trace. Taking into account that s in equation (I 31) is replaced now by s^n and using rule 4, i.e. equation (16), we obtain by (33) from (I 31)

$$\prod_{j=1}^{n}\operatorname{tr}\mathbf{R}(j) = [\operatorname{tr}\mathbf{R}(j)]^n = s^n, \tag{48}$$

where we have considered that all $\mathbf{R}(j)$ have the same traces and therefore

$$\operatorname{tr}\mathbf{R}(j) = s. \tag{49}$$

Equation (49) can also be obtained more rigorously by using the elementary orthonormality properties of the matrix $\boldsymbol{\phi}$.

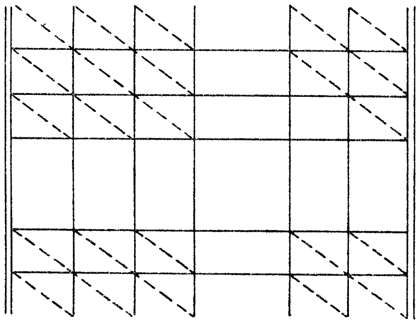
So equations (I 29) to (I 31) are replaced by equations (39), (47) and (49).

The matrix $\mathbf{R}(j)$ is a square matrix of the size $m \times m$, i.e. the dimension of which depends on the subdivision of the space of electron j in smaller volumes.

Now we introduce a new matrix operation to be called *partial trace-forming*.

Let us take a square matrix, $\mathbf{A}$, of the size $k \times k$, where k is a positive integer. Let l be any *divisor* of k. Then the matrix $\mathbf{A}$ can be divided into small square matrices of the size $l \times l$:

$\mathbf{A} =$



(50)

Let us take the traces of these small square matrices. In this way the sum of the first l of the diagonal elements should be A_{11}, the sum of the next l of the diagonal elements A_{22}, and so on. The sum of the diagonal elements of the second square matrix on the top of the matrix should be A_{12}, etc. Then we obtain a square matrix of the size $k/l \times k/l$, which will be called the lth partial trace of $\mathbf{A}$ and denoted by $\mathbf{tr}_l\,\mathbf{A}$, i.e.

$$\mathbf{tr}_l\,\mathbf{A} = \left\| \begin{matrix} A_{11} & A_{12} & A_{13} & \dots & A_{1,k/l} \\ A_{21} & A_{22} & A_{23} & \dots & A_{2,k/l} \\ A_{31} & A_{32} & A_{33} & \dots & A_{3,k/l} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ A_{k/l,1} & A_{k/l,2} & A_{k/l,3} & \dots & A_{k/l,k/l} \end{matrix} \right\|. \tag{51}$$

From this definition it is evident that the trace of the matrix (51) gives the common trace of $\mathbf{A}$, i.e.

$$\mathrm{tr}\,(\mathbf{tr}_l\,\mathbf{A}) = \mathrm{tr}\,\mathbf{A} \tag{52}$$

for any l, which is a divisor of k, the size of the matrix $\mathbf{A}$. Similarly as an extension of (52), if k has a divisor m, and m has as divisor l, we easily can see that

$$\mathbf{tr}_{m/l}(\mathbf{tr}_l\,\mathbf{A}) = \mathbf{tr}_m\,\mathbf{A}; \tag{53}$$

finally, it is self-evident that for a matrix $\mathbf{A}$ of size $k \times k$

$$\mathbf{tr}_k\,\mathbf{A} = \mathrm{tr}\,\mathbf{A}. \tag{54}$$

The operation of partial trace-forming is—in contrast with the normal trace-forming—in general not commutative, i.e.

$$\mathbf{tr}_l(\mathbf{A}\,.\,\mathbf{B}) \neq \mathbf{tr}_l(\mathbf{B}\,.\,\mathbf{A}), \tag{55}$$

as one can easily see by evaluating the elements of the product matrices. However, if $\mathbf{A}$ and $\mathbf{B}$ are commutable themselves, naturally so is the partial trace too.

Let $\mathbf{A}$ be a square matrix of size $k \times k$ and $\mathbf{B}$ be a square matrix of the size $l \times l$. Then we have by the definition of the inner product and the partial trace, the following very useful equation

$$\mathbf{tr}_l[\mathbf{A} \circ (\mathbf{B})] = (\mathrm{tr}\,\mathbf{B})\,\mathbf{A}, \tag{56}$$

where $\mathrm{tr}\,\mathbf{B}$ is the normal trace of $\mathbf{B}$, a number. As the size of the matrix $\mathbf{A} \circ (\mathbf{B})$ is $kl \times kl$, the condition of divisibility for the sizes is here automatically fulfilled.

One may check also very simply that the partial trace formation is additive, i.e. if $\mathbf{A}$ and $\mathbf{B}$ are two matrices of the same size,

$$\mathbf{tr}_l(\mathbf{A}+\mathbf{B}) = \mathbf{tr}_l\,\mathbf{A} + \mathbf{tr}_l\,\mathbf{B}. \tag{57}$$

Therefore if we have a matrix equation, it is allowed to compose an equation by forming the partial traces on both sides of the matrix equation in the same way.

If we have n electrons and the subdivision of the space of each electron is into m volume elements, it is easy to see that the mth partial trace-forming denotes a sort of integration for the total space of one electron. The m^2th partial trace-forming denotes integration for the total space of *two* electrons, etc. The m^{n-1}th partial trace-forming denotes integration over all electrons except one. The partial trace-forming as a size-reducing operation is in some respect the inverse operation of the inner product.

Let us now choose a fixed state and let us investigate the electron distribution in this state. If only one state is taken into account, we have

$$s = 1 \tag{58}$$

and so from (49)

$$\operatorname{tr}\mathbf{R}(j) = 1. \tag{59}$$

The one-electron density matrix, the spatial distribution of one electron (say electron n), $\boldsymbol{\rho}(n)$, is obtained by integrating for all other electrons, which is equivalent as mentioned before to the m^{n-1}th partial trace-forming. So using (33), (56), (16) and (59) we obtain

$$\boldsymbol{\rho}(n) = \mathbf{tr}_{m^{n-1}}\mathbf{R} = \mathbf{R}(n)\prod_{j=1}^{n-1}\operatorname{tr}\mathbf{R}(j) = \mathbf{R}(n), \tag{60}$$

i.e. the $\mathbf{R}(n)$ one-electron density matrices actually supply the density matrices obtained by the common rules of quantum mechanics. The same is true for all $\boldsymbol{\rho}(j)$.

The total density matrix, $\boldsymbol{\rho}$, is therefore

$$\boldsymbol{\rho} = \sum_{j=1}^{n}\boldsymbol{\rho}(j) = \sum_{j=1}^{n}\mathbf{R}(j). \tag{61}$$

Using equations (59) we find

$$\operatorname{tr}\boldsymbol{\rho} = n, \tag{62}$$

as one would expect.

Taking into account the orthogonality of the various functions $\boldsymbol{\phi}_{kj}(j)$ we obtain after some calculation, using (32) and (39)

$$\boldsymbol{\rho}^2 = \sum_{j=1}^{n}\sum_{k=1}^{n}\mathbf{R}(j)\,\mathbf{R}(k) = \sum_{j=1}^{n}\mathbf{R}^2(j) = \sum_{j=1}^{n}\mathbf{R}(j) = \boldsymbol{\rho}. \tag{63}$$

Using equations (1), (61), (47) and the fact that a Hamiltonian for an electron j is commutable every time with the density matrix of another electron, k, we obtain

$$\mathbf{H}\boldsymbol{\rho} - \boldsymbol{\rho}\mathbf{H} = \sum_j\sum_k[\mathbf{H}(j)\,\mathbf{R}(k) - \mathbf{R}(k)\,\mathbf{H}(j)] = \mathbf{0}. \tag{64}$$

We may summarize therefore finally, that in the independent one-electron approximation the quantum-mechanical system is defined by the equations (62), (63) and (64).

3. Independent one-electron eigenfunctions taking into account spin and the Pauli principle

Let us make the same assumptions as in § 2 but additively requiring that no more than two electrons are allowed to be in the same one-electron state and attributing spins of one direction to some of the electrons and spins of the other direction to the others.

All the equations of § 2 are likewise valid in this case, because there we have not used the possibility that more than one electron may be placed in some of the states. To investigate the ground state in this case we make use of the fact that the occupied states are filled once with an electron with a spin in one direction and once with an electron with a spin in the other one. The space parts of the eigenfunctions of these electrons are, however, exactly the same and therefore all the equations used in § 2 remain valid. However, to take advantage of the fact of existence of doubly filled states, some of the equations could be slightly rewritten.

So the definition of the total electron density matrix in the ordinary three-dimensional space is changed from (61) to

$$\boldsymbol{\rho} = 2\sum_{j=1}^{\frac{1}{2}n} \mathbf{R}(j), \tag{65}$$

where we have assumed that n is an even number, so that each one-electron state is filled by two electrons. Later we shall extend the formulas to the case of n odd, when one of the one-electron states is filled by one electron only, but we write here the formulas only for the case of n even.

Equation (62) also remains valid in this case. However, (63) is changed to

$$\boldsymbol{\rho}^2 = 4\sum_{j=1}^{\frac{1}{2}n}\sum_{h=1}^{\frac{1}{2}n} \mathbf{R}(j)\,\mathbf{R}(h) = 4\sum_{j=1}^{\frac{1}{2}n} \mathbf{R}^2(j) = 4\sum_{j=1}^{\frac{1}{2}n} \mathbf{R}(j) = 2\boldsymbol{\rho}. \tag{66}$$

The matrix $\boldsymbol{\rho}$ has so lost its idempotency. At the same time equation (64) remains valid. Summarizing, in the independent one-electron approximation with the Pauli principle, the quantum-mechanical system is defined by equations (64), (66) and (62).

Here we have used the Pauli principle only in the form that no more than two electrons can be placed in an electronic state. The symmetry property of the eigenfunction by an interchange of two electrons was not used here. Later in § 5 we shall consider it in more detail. It is worth mentioning that in the case when the system contains electrons with unpaired spins, we have to introduce the density of the electrons with unpaired spins separately in the common three-dimensional space. (This is the case e.g. in the presence of free radicals or in excited states of an even electron system too.) Let us suppose that the total number of electrons is n and of this n_1 are with unpaired spins. The states of these last electrons being occupied only by single electrons, we may write the definition of the total density matrix of the unpaired (free) electrons, $\boldsymbol{\rho}_u$,

$$\boldsymbol{\rho}_u = \sum_{j=n-n_1+1}^{n} \mathbf{R}(j), \tag{67}$$

where we have chosen the first $n-n_1$ electrons to be those which are paired and the last n_1 to be unpaired. The total electron density matrix is now, instead of (60) or (65),

$$\boldsymbol{\rho} = 2\sum_{j=1}^{\frac{1}{2}(n-n_1)} \mathbf{R}(j) + \sum_{j=n-n_1+1}^{n} \mathbf{R}(j), \tag{68}$$

where it is evident that $n-n_1$ is an even number by definition. Since the one-electron density matrices retain their normalization and orthogonality properties even in this case, the equations corresponding to equations (63) and (66) are now

$$\boldsymbol{\rho}^2 = 4\sum_{j=1}^{\frac{1}{2}(n-n_1)} \mathbf{R}^2(j) + \sum_{j=n-n_1+1}^{n} \mathbf{R}^2(j) = 4\sum_{j=1}^{\frac{1}{2}(n-n_1)} \mathbf{R}(j) + \sum_{j=n-n_1+1}^{n} \mathbf{R}(j) = 2\boldsymbol{\rho} - \boldsymbol{\rho}_u \tag{69}$$

and

$$\boldsymbol{\rho}_u^2 = \sum_{j=n-n_1+1}^{n} \mathbf{R}^2(j) = \sum_{j=n-n_1+1}^{n} \mathbf{R}(j) = \boldsymbol{\rho}_u. \tag{70}$$

It is easily seen that equation (64) still holds but in addition we may write yet another equation for $\boldsymbol{\rho}_u$, namely

$$\mathbf{H}\boldsymbol{\rho}_u - \boldsymbol{\rho}_u\mathbf{H} = \mathbf{0}. \tag{71}$$

Finally, besides the equation (62), which is still valid, we have the equation

$$\operatorname{tr}\boldsymbol{\rho}_u = n_1. \tag{72}$$

Therefore in the case of the presence of n_1 unpaired electrons among the n electrons in total, we have to solve the system of equations composed of (64), (71), (69), (70), (62) and (72).

For practical purposes it is necessary to solve first the problem for the unpaired electrons only, given by the system of equations (71), (70) and (72), not containing in the independent electron case the total electron density, only that of the unpaired electrons; then we can solve the system of equations (64), (69) and (62), where in (69) the formerly obtained $\boldsymbol{\rho}_u$ is to be substituted. In the case of interaction, however, this procedure is already not very simple.

4. One-electron eigenfunctions with interaction without spin

We start with the Hartree assumption (1928) for the eigenfunction that the form (3) of ψ may be used here too, but the Hamiltonian cannot be represented as a sum of one-electron Hamiltonians as in (1).

We may now express the Hamiltonian as a sum of one-electron Hamiltonians like (1) *plus* a part of the Hamiltonian which contains co-ordinates of more than one electron (in practice generally the sum of terms, each containing co-ordinates of two electrons). This can be written in the form

$$\mathbf{H} = \sum_{j=1}^{n} \mathbf{H}(j) + \mathbf{H}', \tag{73}$$

where $\mathbf{H}'$ is the interaction term of the electrons.

Since $\boldsymbol{\psi}$ is again given in the form (25), equation (I 29) may be written now by using equations (30) to (33), which are still valid

$$\mathbf{H}[\mathbf{R}(n) \circ (\mathbf{R}(n-1) \circ \ldots \circ (\mathbf{R}(1))\ldots)] - [\mathbf{R}(n) \circ (\mathbf{R}(n-1) \circ \ldots \circ (\mathbf{R}(1))\ldots)]\,\mathbf{H} = \mathbf{0}. \tag{74}$$

Using the form (73) of the Hamiltonian and making transformations similar to that in (35), we have

$$\begin{aligned}&[\mathbf{H}(n)\,\mathbf{R}(n) - \mathbf{R}(n)\,\mathbf{H}(n)] \circ (\mathbf{R}(n-1) \circ \ldots \circ (\mathbf{R}(1))\ldots)\\ &\quad + \mathbf{R}(n) \circ \left[\sum_{j=1}^{n-1} \mathbf{R}(n-1) \circ (R(n-2) \circ \ldots \circ ([\mathbf{H}(j)\,\mathbf{R}(j) - \mathbf{R}(j)\,\mathbf{H}(j)] \circ (\mathbf{R}(j-1) \circ \ldots))\ldots)\right]\\ &\quad + \mathbf{H}'\mathbf{R} - \mathbf{R}\mathbf{H}' = \mathbf{0}.\end{aligned} \tag{75}$$

Let us now form the m^{n-1}th partial trace of equation (75) making use of equations (56), (16) and—being interested in one state only—of the permanently valid equation (59). Using the commutative character of the normal trace-forming operation we obtain from (75)

$$\mathbf{H}(n)\,\mathbf{R}(n) - \mathbf{R}(n)\,\mathbf{H}(n) + \mathbf{tr}_{m^{n-1}}(\mathbf{H}'\mathbf{R} - \mathbf{R}\mathbf{H}') = \mathbf{0}. \tag{76}$$

We restrict ourselves to the case where there is a *two-particle interaction* only between the electrons. In this case the elements of $\mathbf{H}'$ may affect at most two electrons. Let us investigate first the terms, in which one of these electrons is the nth electron (with the one-electron density matrix $\mathbf{R}(n)$) and the other one any of the other electrons. Take first the case that the other electron is the $(n-1)$th. Then any term of this kind occurring in (76) may be written by using equation (53), (56) and (59)

$$\begin{aligned}&\mathbf{tr}_{m^{n-1}}[(\mathbf{H}'_{n,\,n-1}[\mathbf{R}(n) \circ (\mathbf{R}(n-1))] - [\mathbf{R}(n) \circ (\mathbf{R}(n-1))]\,\mathbf{H}'_{n,\,n-1}) \circ (\mathbf{R}(n-2) \circ \ldots)]\\ &\quad = \mathbf{tr}_m\{\mathbf{tr}_{m^{n-2}}[(\mathbf{H}'_{n,\,n-1}[\mathbf{R}(n) \circ (\mathbf{R}(n-1))] - [\mathbf{R}(n) \circ (\mathbf{R}(n-1))]\,\mathbf{H}'_{n,\,n-1}) \circ (\mathbf{R}(n-2) \circ \ldots)]\}\\ &\quad = \mathbf{tr}_m\{\mathbf{H}'_{n,\,n-1}[\mathbf{R}(n) \circ (\mathbf{R}(n-1))] - [\mathbf{R}(n) \circ (\mathbf{R}(n-1))]\,\mathbf{H}'_{n,\,n-1}\},\end{aligned} \tag{77}$$

where $\mathbf{H}'_{n,\,n-1}$ is that part of $\mathbf{H}'$ which contains the co-ordinates of the nth and of the $(n-1)$th electrons and no others. As the integration over the space of the electron $(n-1)$ involved in the forming of $\mathbf{tr}_m$ leaves $\mathbf{R}(n)$ constant, we may continue the transformation of (77)

$$\begin{aligned}&\mathbf{tr}_m(\mathbf{H}'_{n,\,n-1}[\mathbf{R}(n)\circ(\mathbf{R}(n-1))]-[(\mathbf{R}(n)\circ(\mathbf{R}(n-1))]\,\mathbf{H}'_{n,\,n-1})\\&\qquad=\mathbf{tr}_m(\mathbf{H}'_{n,\,n-1}[\mathbf{I}_{(m)}\circ(\mathbf{R}(n-1))])\,\mathbf{R}(n)-\mathbf{R}(n)\,\mathbf{tr}_m([\mathbf{I}_m\circ(\mathbf{R}(n-1))]\,\mathbf{H}'_{n,\,n-1}).\end{aligned}\tag{78}$$

We shall introduce the following symbols:

$${}_{n-1}\mathbf{H}'(n)=\mathbf{tr}_m(\mathbf{H}'_{n,\,n-1}[\mathbf{I}_m\circ(\mathbf{R}(n-1))])\tag{79}$$

and

$$\mathbf{H}'_{n-1}(n)=\mathbf{tr}_m([\mathbf{I}_m\circ(R(n-1))]\,\mathbf{H}'_{n,\,n-1}),\tag{80}$$

where we have emphasized in this notation, that the resulting matrices depend only on the co-ordinates of electron n, since the partial trace-forming resulted in an integration over the whole space of electron $(n-1)$. The matrices given by (79) and (80) are some kinds of average values averaging over the various positions of the $(n-1)$th electron.

Performing a suitable arranging and rearranging of the sequence of the electrons we obtain similar results for the other contributing electron pairs. In general these are

$${}_j\mathbf{H}'(n)=\mathbf{tr}_m(H'_{n,\,j}[\mathbf{I}_m\circ(\mathbf{R}(j))])\tag{81}$$

and

$$\mathbf{H}'_j(n)=\mathbf{tr}_m([\mathbf{I}_m\circ(\mathbf{R}(j))]\,.\,\mathbf{H}'_{n\ j}).\tag{82}$$

So the term (78) can be written

$${}_{n-1}\mathbf{H}'(n)\,\mathbf{R}(n)-\mathbf{R}(n)\,\mathbf{H}'_{n-1}(n)\tag{83}$$

and similarly the term containing the contribution of the jth electron

$${}_j\mathbf{H}'(n)\,\mathbf{R}(n)-\mathbf{R}(n)\,\mathbf{H}'_j(n).\tag{84}$$

If we consider a term in (76), which contains a pair of electrons j, k, none of which are identical with electron n, using rules 1, 2 and 6, and equation (56), we obtain

$$\begin{aligned}\mathbf{tr}_{m^{n-1}}(\mathbf{H}'_{jk}\mathbf{R}-\mathbf{R}\mathbf{H}'_{jk})&=\mathbf{tr}_{m^{n-1}}[(\mathbf{I}_m\circ(\mathbf{H}'_{jk}))\,(\mathbf{R}(n)\circ(\mathbf{R}(n-1)\circ\ldots))\\&\qquad-(\mathbf{R}(n)\circ(\mathbf{R}(n-1)\circ\ldots))\,(\mathbf{I}_m\circ(\mathbf{H}'_{jk}))]\\&=\mathbf{tr}_{m^{n-1}}[\mathbf{R}(n)\circ(\mathbf{H}'_{jk}\,.\,(\mathbf{R}(n-1)\circ\ldots))-\mathbf{R}(n)\circ((\mathbf{R}(n-1)\circ\ldots)\,.\,\mathbf{H}'_{jk})]\\&=\mathbf{tr}_{m^{n-1}}[\mathbf{R}(n)\circ\{\mathbf{H}'_{jk}(\mathbf{R}(n-1)\circ\ldots)-(\mathbf{R}(n-1)\circ\ldots)\,\mathbf{H}'_{jk}\}]\\&=\mathbf{R}(n)\,\mathrm{tr}\,[\mathbf{H}'_{jk}(\mathbf{R}(n-1)\circ\ldots)-(\mathbf{R}(n-1)\circ\ldots)\,\mathbf{H}'_{jk}]=\mathbf{R}(n)\,.\,0=\mathbf{0},\end{aligned}\tag{85}$$

the last conclusion drawn from the fact of commutability of the *normal* trace forming.

So in (76) only the matrices (81) and (82) make contribution (for all $j=1,2,\ldots,n-1$), and so (76) may be written by (84) and (85)

$$\mathbf{H}(n)\,\mathbf{R}(n)-\mathbf{R}(n)\,\mathbf{H}(n)+\sum_{j=1}^{n-1}[{}_j\mathbf{H}'(n)\,\mathbf{R}(n)-\mathbf{R}(n)\,\mathbf{H}'_j(n)]=\mathbf{0}.\tag{86}$$

Introducing now the modified Hamiltonians

$${}_{\mathrm{mod}}\mathbf{H}(n)=\mathbf{H}(n)+\sum_{j=1}^{n-1}{}_j\mathbf{H}'(n)\tag{87}$$

and

$$\mathbf{H}_{\mathrm{mod}}(n)=\mathbf{H}(n)+\sum_{j=1}^{n-1}\mathbf{H}'_j(n),\tag{88}$$

we can finally write equation (86) in the form

$$_{\mathrm{mod}}\mathbf{H}(n)\,\mathbf{R}(n)-\mathbf{R}(n)\,\mathbf{H}_{\mathrm{mod}}(n)=\mathbf{0}. \tag{89}$$

Observe that the left- and the right-hand side operators (87) and (88) are *in general not the same* as a consequence of the inequality (55) and that the modified Hamiltonians depend now on **R**.

By rearranging the electrons suitably it may be seen that equations similar to (89) are also valid for the other electrons.

So we obtain a system of equations instead of (47)

$$_{\mathrm{mod}}\mathbf{H}(j)\,\mathbf{R}(j)-\mathbf{R}(j)\,\mathbf{H}_{\mathrm{mod}}(j)=\mathbf{0}\quad(j=1,2,\ldots,n), \tag{90}$$

with the definitions

$$_{\mathrm{mod}}\mathbf{H}(j)=\mathbf{H}(j)+\sum_{\substack{l=1\\l\neq j}}^{n}{}'\,{}_{l}\mathbf{H}'(j) \tag{91}$$

and

$$\mathbf{H}_{\mathrm{mod}}(j)=\mathbf{H}(j)+\sum_{\substack{l=1\\l\neq j}}^{n}{}'\,\mathbf{H}'_{l}(j), \tag{92}$$

${}_{l}\mathbf{H}'(j)$ and $\mathbf{H}'_{l}(j)$ being given by (81) and (82).

These are the Hartree equations for the system of n interacting electrons. Unfortunately it is not here possible to obtain an equation like (64) because of the difference in the operators for each electron (see operators (91) and (92)). The system of equations (90) may be solved by the self-consistent-field iteration method using the auxiliary conditions (39) and (59).

5. ONE-ELECTRON EIGENFUNCTIONS WITH INTERACTION TAKING INTO ACCOUNT SPIN AND THE PAULI PRINCIPLE

The Pauli principle was used in § 3 only partly, in that we have not used the symmetry property of the system for an interchange of two electrons.

The Pauli principle states that such an interchange of two electrons has to *change the sign* of the eigenfunction of the total system.

Before continuing the development of the equations we shall introduce a further notation. Extending the sense of $\mathbf{R}(j)$ defined by equation (32) we introduce

$$\mathbf{R}(j,l)=\boldsymbol{\phi}(j)\,\boldsymbol{\phi}^{+*}(l), \tag{93}$$

where j and l denote two different electrons. Evidently if they denote the same electron, (93) regains the form (32). The matrix $\mathbf{R}(j,l)$ has the same size as $\mathbf{R}(j)$ or $\mathbf{R}(l)$ has, i.e. $m\times m$. The arguments of (93) are not interchangable and so the matrix $\mathbf{R}(j,l)$ is not a symmetric matrix in general.

The most general form of the wave function fulfilling the whole Pauli principle is a determinantal wave function of the following form (see Slater 1929)

$$\psi_i(y)=\frac{1}{\surd(n!)}\begin{vmatrix}\phi_{k_1}(y_1) & \phi_{k_2}(y_1) & \cdots & \phi_{k_n}(y_1)\\ \phi_{k_1}(y_2) & \phi_{k_2}(y_2) & \cdots & \phi_{k_n}(y_2)\\ \vdots & \vdots & & \vdots\\ \phi_{k_1}(y_n) & \phi_{k_2}(y_n) & \cdots & \phi_{k_n}(y_n)\end{vmatrix}, \tag{94}$$

where the normalization was already taken into consideration. Expanding the determinant according to the usual rules we obtain

$$\psi_i(y) = \frac{1}{\sqrt{(n!)}} \sum_{\mathbf{P}} \pm \phi_{k_1}(\mathbf{P}y_1)\, \phi_{k_2}(\mathbf{P}y_2) \ldots \phi_{k_n}(\mathbf{P}y_n), \tag{95}$$

where $\mathbf{P}$ denotes a permutation for the variables $y_1, y_2, \ldots, y_n$. If $\mathbf{P}$ is an even permutation, the $+$ sign, if it is an odd permutation, the $-$ sign is valid in the sum (95). $\mathbf{P}$ runs through all possible permutations, $n!$ cases in total.

Each term of the sum in (95) is a wave function of the form (3). So the matrix $\boldsymbol{\psi}$ may be defined

$$\boldsymbol{\psi} = \frac{1}{\sqrt{(n!)}} \sum_{\mathbf{P}} \pm \boldsymbol{\phi}(\mathbf{P}n) \circ (\boldsymbol{\phi}(\mathbf{P}(n-1)) \circ (\boldsymbol{\phi}(\mathbf{P}(n-2)) \circ \ldots \circ (\boldsymbol{\phi}(\mathbf{P}1)) \ldots)), \tag{96}$$

analogous to (25).

Equation (31) is replaced now by a more complicated equation

$$\begin{aligned}\mathbf{R} &= \boldsymbol{\psi}\boldsymbol{\psi}^{+*}\\ &= \frac{1}{n!}\sum_{\mathbf{P}}\sum_{\mathbf{Q}} \pm \boldsymbol{\phi}(\mathbf{P}n)\, \boldsymbol{\phi}^{+*}(\mathbf{Q}n) \circ (\boldsymbol{\phi}(\mathbf{P}(n-1))\, \boldsymbol{\phi}^{+*}(\mathbf{Q}(n-1)) \circ \ldots \circ (\boldsymbol{\phi}(\mathbf{P}1)\, \boldsymbol{\phi}^{+*}(\mathbf{Q}1)) \ldots)\end{aligned} \tag{97}$$

where one must use the $+$ sign if both of the permutations $\mathbf{P}$ and $\mathbf{Q}$ are even or both are odd and the $-$ sign if one of them is even and the other odd. Both $\mathbf{P}$ and $\mathbf{Q}$ run through all possible permutations of $1, 2, \ldots, n$. Using the notation (93), we can write equation (97) in the form

$$\mathbf{R} = \frac{1}{n!}\sum_{\mathbf{P}}\sum_{\mathbf{Q}} \pm \mathbf{R}(\mathbf{P}n, \mathbf{Q}n) \circ (\mathbf{R}(\mathbf{P}(n-1), \mathbf{Q}(n-1)) \circ \ldots \circ (\mathbf{R}(\mathbf{P}1, \mathbf{Q}1)) \ldots). \tag{98}$$

The form of the Hamiltonian is given by equation (73) even in this case. Equation (74) is to be replaced by

$$\begin{aligned}\frac{1}{n!}\sum_{\mathbf{P}}\sum_{\mathbf{Q}} \pm [\mathbf{H}\,.\{\mathbf{R}(\mathbf{P}n, \mathbf{Q}n) \circ (\mathbf{R}(\mathbf{P}(n-1), \mathbf{Q}(n-1)) \circ \ldots \circ (\mathbf{R}(\mathbf{P}1, \mathbf{Q}1)) \ldots)\}\\ -\{\mathbf{R}(\mathbf{P}n, \mathbf{Q}n) \circ (\mathbf{R}(\mathbf{P}(n-1), \mathbf{Q}(n-1)) \circ \ldots \circ (\mathbf{R}(\mathbf{P}1, \mathbf{Q}1)) \ldots)\}\mathbf{H}] = \mathbf{0}.\end{aligned} \tag{99}$$

Making use of the form (73) we now have an equation corresponding to (75),

$$\begin{aligned}\frac{1}{n!}\sum_{\mathbf{P}}\sum_{\mathbf{Q}} \pm \Big[[\mathbf{H}(n)\, \mathbf{R}(\mathbf{P}n, \mathbf{Q}n) - \mathbf{R}(\mathbf{P}n, \mathbf{Q}n)\, \mathbf{H}(n)] \circ (\mathbf{R}(\mathbf{P}(n-1), \mathbf{Q}(n-1)) \circ \ldots \circ (\mathbf{R}(\mathbf{P}1, \mathbf{Q}1)) \ldots)\\ + \mathbf{R}(\mathbf{P}n, \mathbf{Q}n) \circ \Big(\sum_{j=1}^{n-1} \mathbf{R}(\mathbf{P}(n-1), \mathbf{Q}(n-1)) \circ (\mathbf{R}(\mathbf{P}(n-2), \mathbf{Q}(n-2)) \circ \ldots \circ ([\mathbf{H}(j)\, \mathbf{R}(\mathbf{P}j, \mathbf{Q}j)\\ - \mathbf{R}(\mathbf{P}j, \mathbf{Q}j)\, \mathbf{H}(j)] \circ (\mathbf{R}(\mathbf{P}(j-1), \mathbf{Q}(j-1)) \circ \ldots)) \ldots) \Big) \Big] + \mathbf{H}'\mathbf{R} - \mathbf{R}\mathbf{H}' = \mathbf{0}.\end{aligned} \tag{100}$$

We shall now form the m^{n-1}th partial trace of (100). To obtain the simplified result immediately we note that if $\mathbf{P}j = \mathbf{Q}l$ and j and l are not separated by any operator ($\mathbf{H}$ or any parts of $\mathbf{H}$), then the partial trace formation (integration) gives for this term a non-vanishing contribution, whereas if $\mathbf{P}j \neq \mathbf{Q}l$, the orthogonality makes this term $\mathbf{0}$. Using this fact, forming the partial trace of order m^{n-1} of (100), we obtain formally an equation similar to (76) again. However, in the evaluation of terms, where j and l are separated by a part of $\mathbf{H}$ the partial trace formation is changed. Since parts of $\mathbf{H}'$ are involved, which contain a pair of electron co-ordinates, there is a possibility for a non-vanishing contribution from terms with $\mathbf{P}j \neq \mathbf{Q}l$ also. Performing the transformations of (100) leads to the self-consistent-field equations of the Hartree–Fock treatment. We shall not here follow the details of this but refer to the literature; see, for example, McWeeny (1960).

6. Linear combination of atomic orbitals

In molecular problems, especially in the case of large molecules, the amount of work involved in a detailed calculation is extremely large. Therefore even the cruder approximations have, in this case, greater importance. We shall therefore fit the linear combination of atomic orbitals (l.c.a.o.) approximation in the molecular-orbital theory into our present scheme.

In the l.c.a.o. approximation we characterize each electron around an atom by a prescribed normalized atomic wave function and a coefficient which determines the weight of the atomic wave function in question. So in our representations the subdivision of the space of an electron is made roughly into the same number of parts as there are atoms present in the system. Our previous m now denotes therefore the number of atoms in the system.

First, we may imagine it as being built up of one-electron functions each having a specific well-defined state. In this case there is only one state to be considered for every electron and so we may write

$$\phi_j = \left\| \begin{matrix} c_{j1} \\ c_{j2} \\ c_{j3} \\ \vdots \\ c_{jm} \end{matrix} \right\| \quad (j = 1, 2, \ldots, n), \tag{101}$$

where m denotes the number of nuclei and n the number of electrons in the system. The one-electron wave function is approximated in this case by

$$\phi_j = \sum_{l=1}^{m} c_{jl}\varphi_l, \tag{102}$$

where φ_l is a normalized atomic wave function of the atom placed at position l.

Using equation (32) we now have from (101)

$$\mathbf{R}(j) = \left\| \begin{matrix} |c_{j1}|^2 & c_{j1}c_{j2}^* & c_{j1}c_{j3}^* & \ldots & c_{j1}c_{jm}^* \\ c_{j1}^*c_{j2} & |c_{j2}|^2 & c_{j2}c_{j3}^* & \ldots & c_{j2}c_{jm}^* \\ c_{j1}^*c_{j3} & c_{j2}^*c_{j3} & |c_{j3}|^2 & \ldots & c_{j3}c_{jm}^* \\ \vdots & \vdots & \vdots & & \vdots \\ c_{j1}^*c_{jm} & c_{j2}^*c_{jm} & c_{j3}^*c_{jm} & \ldots & |c_{jm}|^2 \end{matrix} \right\|, \tag{103}$$

and from this, using (61),

$$\boldsymbol{\rho} = \left\Vert \begin{matrix} \sum_{j=1}^{n} |c_{j1}|^2 & \sum_{j=1}^{n} c_{j1}c_{j2}^* & \sum_{j=1}^{n} c_{j1}c_{j3}^* & \cdots & \sum_{j=1}^{n} c_{j1}c_{jm}^* \\ \sum_{j=1}^{n} c_{j1}^*c_{j2} & \sum_{j=1}^{n} |c_{j2}|^2 & \sum_{j=1}^{n} c_{j2}c_{j3}^* & \cdots & \sum_{j=1}^{n} c_{j2}c_{jm}^* \\ \sum_{j=1}^{n} c_{j1}^*c_{j3} & \sum_{j=1}^{n} c_{j2}^*c_{j3} & \sum_{j=1}^{n} |c_{j3}|^2 & \cdots & \sum_{j=1}^{n} c_{j3}c_{jm}^* \\ \vdots & \vdots & \vdots & & \vdots \\ \sum_{j=1}^{n} c_{j1}^*c_{jm} & \sum_{j=1}^{n} c_{j2}^*c_{jm} & \sum_{j=1}^{n} c_{j3}^*c_{jm} & \cdots & \sum_{j=1}^{n} |c_{jm}|^2 \end{matrix} \right\Vert. \tag{104}$$

Here we have not in the notation taken advantage of the single or double occupation of a state, so if a state were occupied by two electrons, it would be represented by two different j values in the summations. However, if there are some one-electron states occupied by two electrons, these states occur twice in the sums of (104). If we made further operations with the matrix $\boldsymbol{\rho}$, there would arise a difficulty due to *not all states being orthogonal* (just the states occurring twice for a doubly occupied level are the same and therefore not orthogonal to each other). Therefore at this point it is more convenient to introduce the following separation of the states. Let us suppose that the total number of electrons, n, consists of n_1 unpaired electrons and $\frac{1}{2}(n-n_1)$ pairs of electrons in the same states, but with opposite spins. (In the ground state generally $n_1 = 0$, or 1, but in the excited states we have to take into account even larger n_1 values.)

The density matrix $\boldsymbol{\rho}$ can now be written (an identical transformation of (104)) as

$$\boldsymbol{\rho} = (\rho_{ik}) \quad (i, k = 1, 2, \ldots, m), \tag{105}$$

where the elements of the matrix are given by

$$\rho_{ik} = 2\sum_{j=1}^{\frac{1}{2}(n-n_1)} c_{ji}c_{jk}^* + \sum_{j=\frac{1}{2}(n-n_1)+1}^{\frac{1}{2}(n+n_1)} c_{ji}c_{jk}^*. \tag{106}$$

At the same time we define the density matrix of the unpaired electrons by

$$\boldsymbol{\rho}_u = (\rho_{u,ik}) \quad (i, k = 1, 2, \ldots, m), \tag{107}$$

where the matrix elements are

$$\rho_{u,ik} = \sum_{j=\frac{1}{2}(n-n_1)+1}^{\frac{1}{2}(n+n_1)} c_{ji}c_{jk}^*. \tag{108}$$

It is easy to see that the matrix $\boldsymbol{\rho}_u$ has to satisfy (70) and (72), whereas $\boldsymbol{\rho}$ has to satisfy (69) and (62).

In the special case of $n_1 = 0$ (e.g. a closed shell), $\boldsymbol{\rho}_u = \mathbf{0}$ and so $\boldsymbol{\rho}$ has to satisfy equations (66) and (62).

The Hamiltonian is to be chosen in the form (1) in the simple Hückel approximation (see, for example, Hückel 1931; Coulson 1952; Eyring, Walter & Kimball 1944).

In this approximation the one-electron Hamiltonians $\mathbf{H}(j)$ do not depend on j, and all electrons are governed by the same Hamiltonian. So both the matrices $\boldsymbol{\rho}$ and $\boldsymbol{\rho}_u$ obey the same type of equations, namely equations (64) and (71), where the matrices $\mathbf{H}$ are given either by evaluating some integrals over the original *atomic* wave functions, or semi-empirically.

In the more sophisticated l.c.a.o. approximations we have to choose the Hamiltonian in another way. The most general choice is that of (73), the approximations being different in the actual approximations of the *elements* of $\mathbf{H}(i)$ and $\mathbf{H}'$. In molecular problems however $\mathbf{H}'$ may be chosen always as the sum of two-electron interaction terms, which makes a considerable restriction of the number of possibilities.

We shall not follow here the individual approximations but only refer to some of them in the literature (Lennard-Jones 1937; Mulliken, Rieke & Brown 1941; Coulson & Longuet-Higgins 1947, 1948*a*, *b*; Wheland & Mann 1949; Roothaan 1951; Pariser & Parr 1953*a*, *b*; Pople 1953; Pritchard & Sumner 1956; Streitwieser 1960, 1961).

7. Representation of the Hamiltonian

So far we have not specified the way in which the Hamiltonian is determined. In all the formulas the exact or the modified Hamiltonian was occurring, i.e. an operator for the execution of the prescribed operations in a co-ordinate system. For the sake of more simple handling, however, one may express all the matrices and operators occurring in a *representation*, which enables us to perform the calculations using only simple matrix operations on them.

Let us suppose that there exists a complete basis for all the functions coming up as eigen-functions, i.e. we can express *all* functions occurring in the form

$$f(1,2,\ldots,n)=\sum_{j=1}^{\infty}c_j g_j\,(1,2,\ldots,n), \tag{109}$$

where $1,2,\ldots,n$ denote that the functions contain as variables all the co-ordinates of the electrons $1,2,\ldots,n$. We shall presume the basis $g_1, g_2, \ldots$ to be an orthonormal basis. (This is important in order that the representation of unity, i.e. the basis itself, shall remain a unit matrix.) The representation of the function $f(1,2,\ldots,n)$ is in this case defined as

$$\mathbf{f}=\left\Vert\begin{matrix} c_1 \\ c_2 \\ \vdots \\ c_j \\ \vdots \end{matrix}\right\Vert. \tag{110}$$

In the same way any operator $\mathbf{A}$ operating on a function of all the n electrons gives as a result a function. So we may write

$$(\mathbf{Af})=\left\Vert\begin{matrix} d_1 \\ d_2 \\ \vdots \\ d_j \\ \vdots \end{matrix}\right\Vert, \tag{111}$$

expressing the relation

$$\mathbf{A}f(1,2,\ldots,n)=\sum_{j=1}^{\infty}d_j g_j(1,2,\ldots,n). \tag{112}$$

Substituting (109) into (112), we have

$$\sum_{k=1}^{\infty} c_k \mathbf{A} g_k(1,2,\ldots,n) = \sum_{j=1}^{\infty} d_j g_j(1,2,\ldots,n). \tag{113}$$

Multiplying both sides of (113) by $g_j^*(1,2,\ldots,n)$ from the left and integrating over the whole space (taking into account that $\mathbf{A}$ operates on g_k, but not on g_j^*) we obtain by the orthonormality of the basis $g_1, g_2, \ldots,$

$$\sum_{k=1}^{\infty} c_k \int g_j^*(1,2,\ldots,n)\, \mathbf{A} g_k(1,2,\ldots,n)\, \mathrm{d}\tau = d_j. \tag{114}$$

Introducing the notation

$$A_{jk} = \int g_j^*(1,2,\ldots,n)\, \mathbf{A} g_k(1,2,\ldots,n)\, \mathrm{d}\tau, \tag{115}$$

we can write (114) as

$$\sum_{k=1}^{\infty} A_{jk} c_k = d_j. \tag{116}$$

We see that the quantities A_{jk} do not depend on the special choice of f and this latter comes in only in the quantities c_k. So introducing the matrix

$$\mathbf{A} = \left\| \begin{matrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \ddots \end{matrix} \right\|, \tag{117}$$

and using (110) and (111) we can write equation (116) in the form

$$\mathbf{Af} = (\mathbf{Af}), \tag{118}$$

i.e. the function of the operator can be taken over in this representation by defining the matrix $\mathbf{A}$ by (117) and performing simple matrix multiplication with the representation of the function f, $\mathbf{f}$. The result is given in the same representation.

In this way we define the representation of the Hamiltonian as the matrix $\mathbf{H}$ with the elements

$$H_{jk} = \int g_j^* \mathbf{H} g_k \,\mathrm{d}\tau, \tag{119}$$

which gives according to the rules of quantum-mechanics a Hermitian matrix (see, for example, von Neumann 1932).

The representation now replaces the subdivision of the space, so that instead of the m^n 'space regions' we have to deal with the various basis functions only. In general the complete basis is an infinite set of the functions $g_1, g_2, \ldots$. For practical purposes, however, it is usually enough to take into account a finite number of them. Let us assume that it is sufficient to take into account the first r of the functions: $g_1, g_2, \ldots, g_r$. In some of the cases the basis functions in all of the co-ordinates of the n electrons can also be built up of one electron basis functions. Let us assume that such a one electron basis for the jth electron is

$$g_1(j), \quad g_2(j), \quad \ldots, \quad g_r(j),$$

where we have found r functions as sufficient to describe the phenomena in due exactness. Naturally for every $j = 1, 2, \ldots, n$ we obtain the corresponding basis. One can now construct from these one-electron basis functions the n-electron basis, e.g. by choosing it as

$$g_i(1,2,\ldots,n) = \prod_{j=1}^{n} g_{k_{j,i}}(j), \tag{120}$$

where $k_{j,i}$ are integers between 1 and r. Clearly there are r^n such basis functions in all. As the one-electron basis functions form an orthonormal set, we easily see that if $\mathbf{H}(j)$ is a part of the Hamiltonian containing only the co-ordinates of the jth electron,

$$\int g_i^*(1,2,\ldots,n)\,\mathbf{H}(j)\,g_l(1,2,\ldots,n)\,\mathrm{d}\tau = \int g^*_{k_{j,i}}(j)\,\mathbf{H}(j)\,g_{k_{j,l}}(j)\,\mathrm{d}\tau_j \prod_{\substack{p=1\\p\neq j}}^{n}{}' \delta_{k_{p,i}k_{p,l}}, \tag{121}$$

where δ_{ij} is the Kronecker symbol. Similarly, since the part $\mathbf{H}'$ of the Hamiltonian contains terms only where co-ordinates of *two* of the electrons occur, e.g. the term $\mathbf{H}_{jk}$, we have for this

$$\int g_i^*(1,2,\ldots,n)\,\mathbf{H}_{jk}\,g_l(1,2,\ldots,n)\,\mathrm{d}\tau = \int g^*_{k_{j,i}}(j)\,g^*_{k_{k,l}}(k)\,\mathbf{H}_{jk}\,g_{k_{j,l}}(j)\,g_{k_{k,l}}(k)\,\mathrm{d}\tau_j\,\mathrm{d}\tau_k \prod_{\substack{p=1\\p\neq j,k}}^{n}{}'' \delta_{k_{p,i}k_{p,l}}. \tag{122}$$

So all the matrix elements used in the previous section—even those used in the modified Hamiltonians (see equations (81), (82), (91), (92))—can be calculated in a straightforward manner.

In practical calculations it is sometimes not possible on grounds of economy to include all parts of the Hamiltonian; or sometimes a somewhat altered Hamiltonian is used instead of the exact one in order to make the calculations somewhat simpler. In the next section we shall see how the error can be estimated in this case. Furthermore sometimes the Hamiltonian is changed as a physical perturbation indicates it. Even in this case the error in the calculation is to be estimated.

In any case we shall assume in the following that the exact Hamiltonian and even the approximating Hamiltonian are given in their matrix form in some representation.

8. ESTIMATING THE ERROR OF AN APPROXIMATION

In equations (I 53), (I 54) and (I 55) we have constructed a measure for the *mathematical* error of an approximation. These equations were related to the construction of a density matrix **R**. We shall now extend this somewhat for use when we try to construct the total density matrix, $\boldsymbol{\rho}$. Here we sketch only the case for the one-electron approximation taking into account the Pauli principle. In this case the exact equations to be fulfilled are (69), (70), (71), (64), (72) and (62). Let us suppose, that we have an approximate $\boldsymbol{\rho}_0$ for $\boldsymbol{\rho}$ and an approximate $\boldsymbol{\rho}_{u0}$ for $\boldsymbol{\rho}_u$. We have to calculate the error and how to correct it in the most effective way for the six equations. The definitions of the errors may easily be seen as being analogous to (I 53), (I 54) and (I 55)

$$D_1 = \mathrm{tr}\,[(\boldsymbol{\rho}_{u0}^2-\boldsymbol{\rho}_{u0})^+(\boldsymbol{\rho}_{u0}^2-\boldsymbol{\rho}_{u0})], \tag{123}$$

$$D_2 = \mathrm{tr}\,[(\boldsymbol{\rho}_0^2-2\boldsymbol{\rho}_0+\boldsymbol{\rho}_{u0})^+\,(\boldsymbol{\rho}_0^2-2\boldsymbol{\rho}_0+\boldsymbol{\rho}_{u0})], \tag{124}$$

$$D_3 = \mathrm{tr}\,[(\mathbf{H}\boldsymbol{\rho}_{u0}-\boldsymbol{\rho}_{u0}\mathbf{H})^+\,(\mathbf{H}\boldsymbol{\rho}_{u0}-\boldsymbol{\rho}_{u0}\mathbf{H})], \tag{125}$$

$$D_4 = \mathrm{tr}\,[(\mathbf{H}\boldsymbol{\rho}_0-\boldsymbol{\rho}_0\mathbf{H})^+\,(\mathbf{H}\boldsymbol{\rho}_0-\boldsymbol{\rho}_0\mathbf{H})], \tag{126}$$

$$D_5 = (\mathrm{tr}\,\boldsymbol{\rho}_{u0}-n_1)^2 \tag{127}$$

and

$$D_6 = (\mathrm{tr}\,\boldsymbol{\rho}_0-n)^2. \tag{128}$$

If $\boldsymbol{\rho}_0$ and $\boldsymbol{\rho}_{u0}$ are already good approximations, a better approximation consists in a change of these matrices by $\delta\boldsymbol{\rho}$ and by $\delta\boldsymbol{\rho}_u$ respectively.

We first neglect the higher powers of $\delta\boldsymbol{\rho}$ and $\delta\boldsymbol{\rho}_u$ and so obtain the first order changes in the D's of (123) to (128) after similar transformations as were made in (I 71), (I 72) and (I 73):

$$\delta_1 D_1 = 2\,\mathrm{tr}\,[\mathbf{d}_1^+\,\delta\boldsymbol{\rho}_u], \tag{129}$$

$$\delta_1 D_2 = 2\,\mathrm{tr}\,[\mathbf{d}_2^+\,\delta\boldsymbol{\rho}_u] + 2\,\mathrm{tr}\,[\mathbf{d}_3^+\,\delta\boldsymbol{\rho}], \tag{130}$$

$$\delta_1 D_3 = 2\,\mathrm{tr}\,[\mathbf{d}_4^+\,\delta\boldsymbol{\rho}_u], \tag{131}$$

$$\delta_1 D_4 = 2\,\mathrm{tr}\,[\mathbf{d}_5^+\,\delta\boldsymbol{\rho}], \tag{132}$$

$$\delta_1 D_5 = 2(\mathrm{tr}\,\boldsymbol{\rho}_{u0} - n_1)\,\mathrm{tr}\,\delta\boldsymbol{\rho}_u, \tag{133}$$

$$\delta_1 D_6 = 2(\mathrm{tr}\,\boldsymbol{\rho}_0 - n)\,\mathrm{tr}\,\delta\boldsymbol{\rho}, \tag{134}$$

where δ_1 denotes everywhere that only first-order terms in the corrections were retained and

$$\mathbf{d}_1 = (\boldsymbol{\rho}_{u0}^2 - \boldsymbol{\rho}_{u0})\,\boldsymbol{\rho}_{u0}^+ + \boldsymbol{\rho}_{u0}^+(\boldsymbol{\rho}_{u0}^2 - \boldsymbol{\rho}_{u0}) - (\boldsymbol{\rho}_{u0}^2 - \boldsymbol{\rho}_{u0}), \tag{135}$$

$$\mathbf{d}_2 = \boldsymbol{\rho}_0^2 - 2\boldsymbol{\rho}_0 + \boldsymbol{\rho}_{u0}, \tag{136}$$

$$\mathbf{d}_3 = \mathbf{d}_2\boldsymbol{\rho}_0^+ + \boldsymbol{\rho}_0^+\,\mathbf{d}_2 - 2\mathbf{d}_2, \tag{137}$$

$$\mathbf{d}_4 = \mathbf{H}^2\boldsymbol{\rho}_{u0} + \boldsymbol{\rho}_{u0}\,\mathbf{H}^2 - 2\mathbf{H}\boldsymbol{\rho}_{u0}\,\mathbf{H}, \tag{138}$$

$$\mathbf{d}_5 = \mathbf{H}^2\boldsymbol{\rho}_0 + \boldsymbol{\rho}_0\,\mathbf{H}^2 - 2\mathbf{H}\boldsymbol{\rho}_0\,\mathbf{H}. \tag{139}$$

Here we have used the fact that $\mathbf{H}$ is always symmetrical (considering real matrices), but have not taken advantage of the symmetry of $\boldsymbol{\rho}_0$ and $\boldsymbol{\rho}_{u0}$, since there may be approximations in which these conditions are not yet fulfilled.

Now according to our method in § 4 of I the best correction for D_1 would be

$$\delta\boldsymbol{\rho}_u = -\lambda_1\,\mathbf{d}_1; \tag{140}$$

for D_2 $$\delta\boldsymbol{\rho}_u = -\lambda_2\,\mathbf{d}_2 \tag{141}$$

and $$\delta\boldsymbol{\rho} = -\mu_1\,\mathbf{d}_3; \tag{142}$$

for D_3 $$\delta\boldsymbol{\rho}_u = -\lambda_3\,\mathbf{d}_4; \tag{143}$$

for D_4 $$\delta\boldsymbol{\rho} = -\mu_2\,\mathbf{d}_5; \tag{144}$$

for D_5 $$\delta\boldsymbol{\rho}_u = -\lambda_4\mathbf{I}_{(m)}; \tag{145}$$

and for D_6 $$\delta\boldsymbol{\rho} = -\mu_3\mathbf{I}_{(m)}; \tag{146}$$

where λ_1, λ_2, λ_3, λ_4, μ_1, μ_2 and μ_3 are numbers to be determined from the higher-order terms of the error.

Evidently (140), (141), (143) and (145) cannot generally be fulfilled simultaneously. So we try to make an approximation in the form

$$\delta\boldsymbol{\rho}_u = -\lambda_1\,\mathbf{d}_1 - \lambda_2\,\mathbf{d}_2 - \lambda_3\,\mathbf{d}_4 - \lambda_4\mathbf{I}_{(m)} \tag{147}$$

and similarly since (142), (144) and (146) cannot generally be satisfied simultaneously, we consider instead of them the form

$$\delta\boldsymbol{\rho} = -\mu_1\,\mathbf{d}_3 - \mu_2\,\mathbf{d}_5 - \mu_3\mathbf{I}_{(m)}. \tag{148}$$

Introducing (147) and (148) into the correction formulas for D_1 to D_6 up to the second-order terms, we obtain the equations

$$\delta_2 D_i = \sum_{j=1}^{4}\sum_{k=1}^{4}\lambda_j\lambda_k P_{jk}^{(i)} + 2\sum_{j=1}^{4}\sum_{k=1}^{3}\lambda_j\mu_k Q_{jk}^{(i)} + \sum_{j=1}^{3}\sum_{k=1}^{3}\mu_j\mu_k S_{jk}^{(i)} + 2\sum_{j=1}^{4}\lambda_j P_j^{(i)} + 2\sum_{j=1}^{3}\mu_j S_j^{(i)}\ (i=1,\ldots,6), \tag{149}$$

where δ_2 means the second-order approximation and the coefficients $P^{(i)}_{jk}$, $Q^{(i)}_{jk}$, $S^{(i)}_{jk}$, $P^{(i)}_{j}$ and $S^{(i)}_{j}$ are given explicitly in tables 1 to 6. In these tables the following notation is used:

$$\mathbf{d}_{im} = \boldsymbol{\rho}_{u0}\,\mathbf{d}_i + \mathbf{d}_i\,\boldsymbol{\rho}_{u0} - \mathbf{d}_i \quad (i=1,2,4), \tag{150}$$

$$\mathbf{d}_{im} = \boldsymbol{\rho}_{0}\,\mathbf{d}_i + \mathbf{d}_i\,\boldsymbol{\rho}_{0} - 2\mathbf{d}_i \quad (i=3,5), \tag{151}$$

and

$$\mathbf{H}_i = \mathbf{H}\mathbf{d}_i - \mathbf{d}_i\,\mathbf{H} \quad (i=1,\dots,5) \tag{152}$$

TABLE 1. COEFFICIENTS OF $\delta_2 D_1$

$$\begin{aligned}
Q^{(1)}_{jk} &= 0 \quad (j=1,\dots,4;\ k=1,2,3)\\
S^{(1)}_{jk} &= 0 \quad (j,k=1,2,3)\\
S^{(1)}_{j} &= 0 \quad (j=1,2,3)\\
P^{(1)}_{11} &= \operatorname{tr}[\mathbf{d}^{+}_{1m}\mathbf{d}_{1m}] + 2\operatorname{tr}[(\boldsymbol{\rho}^2_{u0}-\boldsymbol{\rho}_{u0})^{+}\mathbf{d}^2_1]\\
P^{(1)}_{12} &= \operatorname{tr}[\mathbf{d}^{+}_{1m}\mathbf{d}_{2m}] + \operatorname{tr}[(\boldsymbol{\rho}^2_{u0}-\boldsymbol{\rho}_{u0})^{+}(\mathbf{d}_1\mathbf{d}_2+\mathbf{d}_2\mathbf{d}_1)]\\
P^{(1)}_{13} &= \operatorname{tr}[\mathbf{d}^{+}_{1m}\mathbf{d}_{4m}] + \operatorname{tr}[(\boldsymbol{\rho}^2_{u0}-\boldsymbol{\rho}_{u0})^{+}(\mathbf{d}_1\mathbf{d}_4+\mathbf{d}_4\mathbf{d}_1)]\\
P^{(1)}_{14} &= \operatorname{tr}[\mathbf{d}^{+}_{1m}(2\boldsymbol{\rho}_{u0}-\mathbf{I}_{(m)})] + 2\operatorname{tr}[(\boldsymbol{\rho}^2_{u0}-\boldsymbol{\rho}_{u0})^{+}\mathbf{d}_1]\\
P^{(1)}_{22} &= \operatorname{tr}[\mathbf{d}^{+}_{2m}\mathbf{d}_{2m}] + 2\operatorname{tr}[(\boldsymbol{\rho}^2_{u0}-\boldsymbol{\rho}_{u0})^{+}\mathbf{d}^2_2]\\
P^{(1)}_{23} &= \operatorname{tr}[\mathbf{d}^{+}_{2m}\mathbf{d}_{4m}] + \operatorname{tr}[(\boldsymbol{\rho}^2_{u0}-\boldsymbol{\rho}_{u0})^{+}(\mathbf{d}_2\mathbf{d}_4+\mathbf{d}_4\mathbf{d}_2)]
\end{aligned}$$

$$\begin{aligned}
P^{(1)}_{24} &= \operatorname{tr}[\mathbf{d}^{+}_{2m}(2\boldsymbol{\rho}_{u0}-\mathbf{I}_{(m)})] + 2\operatorname{tr}[(\boldsymbol{\rho}^2_{u0}-\boldsymbol{\rho}_{u0})^{+}\mathbf{d}_2]\\
P^{(1)}_{33} &= \operatorname{tr}[\mathbf{d}^{+}_{4m}\mathbf{d}_{4m}] + 2\operatorname{tr}[(\boldsymbol{\rho}^2_{u0}-\boldsymbol{\rho}_{u0})^{+}\mathbf{d}^2_4]\\
P^{(1)}_{34} &= \operatorname{tr}[\mathbf{d}^{+}_{4m}(2\boldsymbol{\rho}_{u0}-\mathbf{I}_{(m)})] + 2\operatorname{tr}[(\boldsymbol{\rho}^2_{u0}-\boldsymbol{\rho}_{u0})^{+}\mathbf{d}_4]\\
P^{(1)}_{44} &= \operatorname{tr}[(2\boldsymbol{\rho}_{u0}-\mathbf{I}_{(m)})^{+}(2\boldsymbol{\rho}_{u0}-\mathbf{I}_{(m)})] + 2\operatorname{tr}[\boldsymbol{\rho}^2_{u0}-\boldsymbol{\rho}_{u0}]\\
P^{(1)}_{1} &= -\operatorname{tr}[\mathbf{d}^{+}_{1}\mathbf{d}_{1}]\\
P^{(1)}_{2} &= -\operatorname{tr}[\mathbf{d}^{+}_{1}\mathbf{d}_{2}]\\
P^{(1)}_{3} &= -\operatorname{tr}[\mathbf{d}^{+}_{1}\mathbf{d}_{4}]\\
P^{(1)}_{4} &= -\operatorname{tr}\mathbf{d}_{1}
\end{aligned}$$

TABLE 2. COEFFICIENTS OF $\delta_2 D_2$

$$\begin{aligned}
P^{(2)}_{11} &= \operatorname{tr}[\mathbf{d}^{+}_{1}\mathbf{d}_{1}]\\
P^{(2)}_{12} &= \operatorname{tr}[\mathbf{d}^{+}_{1}\mathbf{d}_{2}]\\
P^{(2)}_{13} &= \operatorname{tr}[\mathbf{d}^{+}_{1}\mathbf{d}_{4}]\\
P^{(2)}_{14} &= \operatorname{tr}\mathbf{d}_{1}\\
P^{(2)}_{22} &= \operatorname{tr}[\mathbf{d}^{+}_{2}\mathbf{d}_{2}]\\
P^{(2)}_{23} &= \operatorname{tr}[\mathbf{d}^{+}_{2}\mathbf{d}_{4}]\\
P^{(2)}_{24} &= \operatorname{tr}\mathbf{d}_{2}\\
P^{(2)}_{33} &= \operatorname{tr}[\mathbf{d}^{+}_{4}\mathbf{d}_{4}]\\
P^{(2)}_{34} &= 0\\
P^{(2)}_{44} &= m
\end{aligned}$$

$$\begin{aligned}
Q^{(2)}_{11} &= \operatorname{tr}[\mathbf{d}^{+}_{1}\mathbf{d}_{3m}]\\
Q^{(2)}_{12} &= \operatorname{tr}[\mathbf{d}^{+}_{1}\mathbf{d}_{5m}]\\
Q^{(2)}_{13} &= 2\operatorname{tr}[\mathbf{d}^{+}_{1}(\boldsymbol{\rho}_{0}-\mathbf{I}_{(m)})]\\
Q^{(2)}_{21} &= \operatorname{tr}[\mathbf{d}^{+}_{2}\mathbf{d}_{3m}]\\
Q^{(2)}_{22} &= \operatorname{tr}[\mathbf{d}^{+}_{2}\mathbf{d}_{5m}]\\
Q^{(2)}_{23} &= 2\operatorname{tr}[\mathbf{d}^{+}_{2}(\boldsymbol{\rho}_{0}-\mathbf{I}_{(m)})]\\
Q^{(2)}_{31} &= \operatorname{tr}[\mathbf{d}^{+}_{4}\mathbf{d}_{3m}]\\
Q^{(2)}_{32} &= \operatorname{tr}[\mathbf{d}^{+}_{4}\mathbf{d}_{5m}]\\
Q^{(2)}_{33} &= 2\operatorname{tr}[\mathbf{d}^{+}_{4}(\boldsymbol{\rho}_{0}-\mathbf{I}_{(m)})]\\
Q^{(2)}_{41} &= \operatorname{tr}\mathbf{d}_{3m}\\
Q^{(2)}_{42} &= \operatorname{tr}\mathbf{d}_{5m}\\
Q^{(2)}_{43} &= 2(\operatorname{tr}\boldsymbol{\rho}_{0}-m)
\end{aligned}$$

$$\begin{aligned}
S^{(2)}_{11} &= \operatorname{tr}[\mathbf{d}^{+}_{3m}\mathbf{d}_{3m}] + 2\operatorname{tr}[\mathbf{d}^{+}_{2}\mathbf{d}^2_{3}]\\
S^{(2)}_{12} &= \operatorname{tr}[\mathbf{d}^{+}_{3m}\mathbf{d}_{5m}] + \operatorname{tr}[\mathbf{d}^{+}_{2}(\mathbf{d}_3\mathbf{d}_5+\mathbf{d}_5\mathbf{d}_3)]\\
S^{(2)}_{13} &= 2\operatorname{tr}[\mathbf{d}^{+}_{3m}(\boldsymbol{\rho}_{0}-\mathbf{I}_{(m)})] + 2\operatorname{tr}[\mathbf{d}^{+}_{2}\mathbf{d}_{3}]\\
S^{(2)}_{22} &= \operatorname{tr}[\mathbf{d}^{+}_{5m}\mathbf{d}_{5m}] + 2\operatorname{tr}[\mathbf{d}^{+}_{2}\mathbf{d}^2_{5}]\\
S^{(2)}_{23} &= 2\operatorname{tr}[\mathbf{d}^{+}_{5m}(\boldsymbol{\rho}_{0}-\mathbf{I}_{(m)})] + 2\operatorname{tr}[\mathbf{d}^{+}_{2}\mathbf{d}_{5}]\\
S^{(2)}_{33} &= 4\operatorname{tr}[(\boldsymbol{\rho}_{0}-\mathbf{I}_{(m)})^{+}(\boldsymbol{\rho}_{0}-\mathbf{I}_{(m)})] + 2\operatorname{tr}\mathbf{d}_{2}\\
P^{(2)}_{1} &= -\operatorname{tr}[\mathbf{d}^{+}_{1}\mathbf{d}_{2}] \qquad S^{(2)}_{1} = -\operatorname{tr}[\mathbf{d}^{+}_{3}\mathbf{d}_{3}]\\
P^{(2)}_{2} &= -\operatorname{tr}[\mathbf{d}^{+}_{2}\mathbf{d}_{2}] \qquad S^{(2)}_{2} = -\operatorname{tr}[\mathbf{d}^{+}_{3}\mathbf{d}_{5}]\\
P^{(2)}_{3} &= -\operatorname{tr}[\mathbf{d}^{+}_{2}\mathbf{d}_{4}] \qquad S^{(2)}_{3} = -\operatorname{tr}\mathbf{d}_{3}\\
P^{(2)}_{4} &= -\operatorname{tr}\mathbf{d}_{2}
\end{aligned}$$

TABLE 3. COEFFICIENTS OF $\delta_2 D_3$

$$\begin{aligned}
Q^{(3)}_{jk} &= 0 \quad (j=1,\dots,4;\ k=1,2,3)\\
S^{(3)}_{jk} &= 0 \quad (j,k=1,2,3)\\
S^{(3)}_{j} &= 0 \quad (j=1,2,3)\\
P^{(3)}_{11} &= \operatorname{tr}[\mathbf{H}^{+}_{1}\mathbf{H}_{1}]\\
P^{(3)}_{12} &= \operatorname{tr}[\mathbf{H}^{+}_{1}\mathbf{H}_{2}]\\
P^{(3)}_{13} &= \operatorname{tr}[\mathbf{H}^{+}_{1}\mathbf{H}_{4}]\\
P^{(3)}_{14} &= 0\\
P^{(3)}_{22} &= \operatorname{tr}[\mathbf{H}^{+}_{2}\mathbf{H}_{2}]
\end{aligned}$$

$$\begin{aligned}
P^{(3)}_{23} &= \operatorname{tr}[\mathbf{H}^{+}_{2}\mathbf{H}_{4}]\\
P^{(3)}_{24} &= 0\\
P^{(3)}_{33} &= \operatorname{tr}[\mathbf{H}^{+}_{4}\mathbf{H}_{4}]\\
P^{(3)}_{34} &= 0\\
P^{(3)}_{44} &= 0\\
P^{(3)}_{1} &= -\operatorname{tr}[\mathbf{d}^{+}_{1}\mathbf{d}_{4}]\\
P^{(3)}_{2} &= -\operatorname{tr}[\mathbf{d}^{+}_{2}\mathbf{d}_{4}]\\
P^{(3)}_{3} &= -\operatorname{tr}[\mathbf{d}^{+}_{4}\mathbf{d}_{4}]\\
P^{(3)}_{4} &= 0
\end{aligned}$$

For all the coefficients we have generally

$$P^{(i)}_{jk} = P^{(i)}_{kj} \tag{153}$$

and

$$S^{(i)}_{jk} = S^{(i)}_{kj} \tag{154}$$

for all i, j and k. Naturally the same is not true for $Q^{(i)}_{jk}$, where j and k are not symmetrical indices. In tables 1 to 6 the property of $\operatorname{tr}\mathbf{d}_4$ and $\operatorname{tr}\mathbf{d}_5$ of vanishing generally has been taken into account.

$\delta_2 D_1$ is to be minimized for λ_1 and at the same time the other $\lambda_2, \lambda_3, \lambda_4$ occurring in it are to be held constant. $\delta_2 D_2$ is to be minimized according to λ_2 *and* μ_1 while $\lambda_1, \lambda_3, \lambda_4, \mu_2$ and μ_3 occurring in it should be regarded as constant parameters. In $\delta_2 D_3$ λ_3 is the minimizing variable, in $\delta_2 D_4$ μ_2, in $\delta_2 D_5$ λ_4 and finally in $\delta_2 D_6$ μ_3. So we obtain seven linear equations for the seven unknowns $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \mu_1, \mu_2, \mu_3$, which may be determined in this way. The system of linear equations leading to the solution is easily obtained by differentiating the proper equations (149) according to the variables mentioned above.

TABLE 4. COEFFICIENTS OF $\delta_2 D_4$

$P^{(4)}_{jk} = 0 \quad (j, k = 1, \ldots, 4)$
$Q^{(4)}_{jk} = 0 \quad (j = 1, \ldots, 4;\ k = 1, 2, 3)$
$P^{(4)}_{j} = 0 \quad (j = 1, \ldots, 4)$
$S^{(4)}_{11} = \operatorname{tr}[\mathbf{H}_3^+ \mathbf{H}_3]$
$S^{(4)}_{12} = \operatorname{tr}[\mathbf{H}_3^+ \mathbf{H}_5]$
$S = 0^{(4)}_{13}$
$S^{(4)}_{22} = \operatorname{tr}[\mathbf{H}_5^+ \mathbf{H}_5]$
$S^{(4)}_{23} = 0$
$S^{(4)}_{33} = 0$
$S^{(4)}_{1} = -\operatorname{tr}[\mathbf{d}_3^+ \mathbf{d}_5]$
$S^{(4)}_{2} = -\operatorname{tr}[\mathbf{d}_5^+ \mathbf{d}_5]$
$S^{(4)}_{3} = 0$

TABLE 5. COEFFICIENTS OF $\delta_2 D_5$

$Q^{(5)}_{jk} = 0 \quad (j = 1, \ldots, 4;\ k = 1, 2, 3)$
$S^{(5)}_{jk} = 0 \quad (j, k = 1, 2, 3)$
$S^{(5)}_{j} = 0 \quad (j = 1, 2, 3)$
$P^{(5)}_{11} = (\operatorname{tr} \mathbf{d}_1)^2$
$P^{(5)}_{12} = \operatorname{tr} \mathbf{d}_1 \operatorname{tr} \mathbf{d}_2$
$P^{(5)}_{13} = 0$
$P^{(5)}_{14} = m \operatorname{tr} \mathbf{d}_1$
$P^{(5)}_{22} = (\operatorname{tr} \mathbf{d}_2)^2$
$P^{(5)}_{23} = 0$
$P^{(5)}_{24} = m \operatorname{tr} \mathbf{d}_2$
$P^{(5)}_{33} = 0$
$P^{(5)}_{34} = 0$
$P^{(5)}_{44} = m^2$
$P^{(5)}_{1} = -\operatorname{tr} \mathbf{d}_1 (\operatorname{tr} \boldsymbol{\rho}_{u0} - n_1)$
$P^{(5)}_{2} = -\operatorname{tr} \mathbf{d}_2 (\operatorname{tr} \boldsymbol{\rho}_{u0} - n_1)$
$P^{(5)}_{3} = 0$
$P^{(5)}_{4} = -m(\operatorname{tr} \boldsymbol{\rho}_{u0} - n_1)$

TABLE 6. COEFFICIENTS OF $\delta_2 D_6$

$P^{(6)}_{jk} = 0 \quad (j, k = 1, \ldots, 4)$
$Q^{(6)}_{jk} = 0 \quad (j = 1, \ldots, 4;\ k = 1, 2, 3)$
$P^{(6)}_{j} = 0 \quad (j = 1, \ldots, 4)$
$S^{(6)}_{11} = (\operatorname{tr} \mathbf{d}_3)^2$
$S^{(6)}_{12} = 0$
$S^{(6)}_{13} = m \operatorname{tr} \mathbf{d}_3$
$S^{(6)}_{22} = 0$
$S^{(6)}_{23} = 0$
$S^{(6)}_{33} = m^2$
$S^{(6)}_{1} = -\operatorname{tr} \mathbf{d}_3 (\operatorname{tr} \boldsymbol{\rho}_0 - n)$
$S^{(6)}_{2} = 0$
$S^{(6)}_{3} = -m(\operatorname{tr} \boldsymbol{\rho}_0 - n)$

The non-negativity of the roots can be disregarded on the same grounds as given at the end of § 4 of I.

With the λ's and μ's determined in this way we obtain the next approximation for $\boldsymbol{\rho}$ and $\boldsymbol{\rho}_u$,

$$\boldsymbol{\rho}_{u1} = \boldsymbol{\rho}_{u0} - \lambda_1 \mathbf{d}_1 - \lambda_2 \mathbf{d}_2 - \lambda_3 \mathbf{d}_4 - \lambda_4 \mathbf{I}_{(m)} \tag{155}$$

and

$$\boldsymbol{\rho}_1 = \boldsymbol{\rho}_0 - \mu_1 \mathbf{d}_3 - \mu_2 \mathbf{d}_5 - \mu_3 \mathbf{I}_{(m)}, \tag{156}$$

and with these the whole process can be repeated with a subsequent estimation of the errors.

If some of the equations are exactly satisfied, the set of equations for the next approximation simplifies considerably. For instance, in the closed shell case $\boldsymbol{\rho}_u = \mathbf{0}$ exactly, and so $D_1 = D_3 = D_5 = 0$ in every case. Only the deviations D_2, D_4 and D_6 remain and consequently only the correction (148). The system of equations to be solved reduces in this case to three equations for the determination of μ_1, μ_2 and μ_3, taking $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$. However, sometimes the temporary deviation from $\boldsymbol{\rho}_u = \mathbf{0}$ supplies a faster convergence by not taking

$\lambda_2 = \lambda_4 = 0$, but instead determining these also from the excess equations. At first sight it seems to be inconvenient that the errors D_1, D_3 and D_5, which were in the initial approximation all zero, temporarily will have some positive values. This loss will, however, be offset by the faster convergence on the top of the 'hills' in a straight direction instead of winding in the 'valleys'.

9. APPLICATION OF THE 'SUBJECTIVE' MINIMIZATION

In I the possibility to define a single error quantity for a calculation was investigated and called 'subjective' error procedure, since there was not a unique way to attribute the proper weights for the deviations. In our present application we may proceed in the same way as was done in equations (I 86) to (I 89) in the general theory.

For the one-electron density matrices we have to introduce besides the normalization factor α given by (I 84)

$$\gamma = 1/n \tag{157}$$

to normalize the equations (149).

From the equation (149) we have to construct in this way

$$N = \gamma[D_1\,\delta_2 D_1 + D_2\,\delta_2 D_2 + \alpha D_3\,\delta_2 D_3 + \alpha D_4\,\delta_2 D_4 + D_5\,\delta_2 D_5 + D_6\,\delta_2 D_6], \tag{158}$$

where the D_i's are given by (123) to (128), the $\delta_2 D_i$'s by (149) and γ is given by (157).

N contains λ_1, λ_2, λ_3, λ_4, μ_1, μ_2 and μ_3 quadratically and we have to minimize it according to these variables. This gives seven simultaneous linear equations for the seven unknowns and the solutions supply through (147) and (148) the next approximation.

We give here the seven equations to be solved in terms of the quantities of tables 1 to 6. Although they seem to be a bit complicated, many of the coefficients vanish due to symmetry in most cases, but even if they do not, the evaluation of the coefficients does not cause any difficulties. Four of the equations are

$$\sum_{k=1}^{4} \lambda_k \bar{P}_{jk} + \sum_{k=1}^{3} \mu_k \bar{Q}_{jk} = -\bar{P}_j \quad (j = 1, \ldots, 4), \tag{159}$$

and the other three

$$\sum_{k=1}^{4} \lambda_k \bar{Q}_{kj} + \sum_{k=1}^{3} \mu_k \bar{S}_{jk} = -\bar{S}_j \quad (j = 1, 2, 3), \tag{160}$$

where we have used the notation

$$\alpha_1 = \alpha_2 = \alpha_5 = \alpha_6 = 1, \tag{161}$$

$$\alpha_3 = \alpha_4 = \alpha; \tag{162}$$

further,

$$\bar{P}_{jk} = \sum_{i=1}^{6} \alpha_i D_i P^{(i)}_{jk} \quad (j, k = 1, \ldots, 4), \tag{163}$$

$$\bar{Q}_{jk} = \sum_{i=1}^{6} \alpha_i D_i Q^{(i)}_{jk} \quad (j = 1, \ldots, 4;\ k = 1, 2, 3), \tag{164}$$

$$\bar{S}_{jk} = \sum_{i=1}^{6} \alpha_i D_i S^{(i)}_{jk} \quad (j, k = 1, 2, 3), \tag{165}$$

$$\bar{P}_j = \sum_{i=1}^{6} \alpha_i D_i P^{(i)}_j \quad (j = 1, \ldots, 4), \tag{166}$$

$$\bar{S}_j = \sum_{i=1}^{6} \alpha_i D_i S^{(i)}_j \quad (j = 1, 2, 3). \tag{167}$$

Although there is a certain arbitrariness in the definition of the overall error N, in practice this procedure seems to give a good estimate for the error and the corrections performed with its aid are much faster than the methods of the previous chapters.

I am indebted to Professor C. A. Coulson, F.R.S., for calling my attention to some of the problems presented in this article, further to Professor P.-O. Löwdin, Uppsala and Dr K. Ohno, Uppsala for their helpful criticism on the subject. I would like to express my thanks to Unesco for giving a grant in the field of the 'Peaceful uses of atomic energy' which made it possible to prepare this work and to the AB Atomenergi, Studsvik, where this work was performed.

References

Coulson, C. A. 1952 *Valence*, Oxford University Press.
Coulson, C. A. & Longuet-Higgins, H. C. 1947 *Proc. Roy. Soc.* A, **191**, 39.
Coulson, C. A. & Longuet-Higgins, H. C. 1948*a* *Proc. Roy. Soc.* A, **193**, 447.
Coulson, C. A. & Longuet-Higgins, H. C. 1948*b* *Proc. Roy. Soc.* A, **195**, 188.
Eyring, H., Walter, J. & Kimball, G. E. 1944 *Quantum chemistry*. New York: John Wiley and Sons.
Hartree, D. R. 1928 *Proc. Camb. Phil. Soc.* **24**, 89.
Hoffmann, T. A. 1965 *Phil. Trans.* **257**, 309.
Hückel, E. 1931 *Z. Phys.* **70**, 204.
Lennard-Jones, J. E. 1937 *Proc. Roy. Soc.* A, **158**, 280.
McWeeny, R. 1960 *Rev. Mod. Phys.* **32**, 335.
Mulliken, R. S., Rieke, C. A. & Brown, W. G. 1941 *J. Amer. Chem. Soc.* **63**, 41.
von Neumann, J. 1932 *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer.
Pariser, R. & Parr, R. G. 1953*a* *J. Chem. Phys.* **21**, 466.
Pariser, R. & Parr, R. G. 1953*b* *J. Chem. Phys.* **21**, 767.
Pople, J. A. 1953 *Trans. Faraday Soc.* **49**, 1375.
Pritchard, H. O. & Sumner, F. H. 1956 *Proc. Roy. Soc.* A, **235**, 136.
Roothaan, C. C. J. 1951 *Rev. Mod. Phys.* **23**, 69.
Slater, J. C. 1929 *Phys. Rev.* **34**, 1293.
Streitwieser, A. Jr. 1960 *J. Amer. Chem. Soc.* **82**, 4123.
Streitwieser, A. Jr. 1961 *Molecular orbital theory*. New York: John Wiley and Sons.
Wheland, G. W. & Mann, D. E. 1949 *J. Chem. Phys.* **17**, 264.