

*Note on Exchange Phenomena in the Thomas Atom.* By P. A. M. DIRAC, Ph.D., St John's College.

[Read 19 May 1930.]

### *Introduction.*

For dealing with atoms involving many electrons the accurate quantum theory, involving a solution of the wave equation in many-dimensional space, is far too complicated to be practicable. One must therefore resort to approximate methods. The best of these is Hartree's method of the self-consistent field\*. Even this, however, is hardly practicable when one has to deal with very many electrons, so that one then requires a still simpler and rougher method. Such a method is provided by Thomas' atomic model†, in which the electrons are regarded as forming a perfect gas satisfying the Fermi statistics and occupying the region of phase space of lowest energy. This region of phase space is assumed to be saturated, with two electrons with opposite spins in each volume  $(2\pi\hbar)^3‡$ , and the remainder is assumed to be empty. Although this model hitherto has not been justified theoretically, it seems to be a plausible approximation for the interior of a heavy atom and one may expect it to give with some accuracy the distribution of electric charge there.

The method of the self-consistent field has recently been established on a very much better theoretical basis in a paper by Fock§, which shows how one can take into account the exchange phenomena between the equivalent electrons. Fock shows that if one takes the best approximation to the many-dimensional wave function that is of the form of a product of a number of three-dimensional wave functions, one for each electron, then the three-dimensional wave functions will satisfy just Hartree's equations. In this way a theoretical justification for Hartree's method is obtained. The exclusion principle of Pauli, however, requires that the wave function representing a number of electrons shall always be antisymmetrical. One would therefore expect to get a better approximation if one first made the product of a number of three-dimensional wave functions antisymmetrical, by applying permutations and taking a linear combination, and then made it approach as closely as possible to

\* Hartree, *Proc. Camb. Phil. Soc.*, Vol. 24, p. 111 (1927).

† Thomas, *Proc. Camb. Phil. Soc.*, Vol. 23, p. 542 (1926). See also Fermi, *Zeit. für Phys.*, Vol. 48, p. 73 (1928).

‡  $\hbar$  denotes Planck's constant divided by  $2\pi$ .

§ Fock, *Zeit. für Phys.*, Vol. 61, p. 126 (1930).

the accurate many-dimensional wave function. The three-dimensional wave functions will then, as found by Fock, satisfy equations somewhat different from Hartree's, containing extra terms which may be considered as representing the exchange phenomena.

Each three-dimensional wave function will give rise to a certain electric density. This electric density is really a matrix, like all dynamical variables in the quantum theory (although one usually considers only its diagonal elements, as one can insert these directly into one's picture of the atom). By adding the electric densities arising from all the wave functions we can obtain the total electric density for the atom. If we adopt the equations of the self-consistent field as amended for exchange, then this total electric density (the matrix) has one important property, namely, *if the value of the total electric density at any time is given, then its value at any later time is determined by the equations of motion.* This means that the whole state of the atom is completely described simply by this electric density; it is not necessary to specify the individual three-dimensional wave functions that make up the total electric density. Thus one can deal with any number of electrons by working with just one matrix density function.

In problems involving a large number of electrons it may be that the electric density, considered as a function of the coordinates and momenta, varies appreciably only through regions of phase space large compared with  $(2\pi\hbar)^3$ . One can then make the further approximation of neglecting the fact that the coordinates and momenta do not commute, so that the problem reduces to a classical problem. One still, however, has the condition that the number of electrons per volume  $(2\pi\hbar)^3$  of phase space cannot exceed two.

We shall examine with this further approximation the electron distribution in the state of lowest energy of an atom, for which a certain region of phase space is occupied with the maximum density of electrons and the remainder is empty. We shall find that the equation governing the boundary between the occupied and unoccupied phase space is just that of Thomas' model, with a small extra term representing the exchange effect. In this way we obtain a theoretical justification for Thomas' model and also see what correction must be made in it to allow for exchange.

### *Equations of the self-consistent field.*

When an atom with  $n$  electrons is treated by the method of the self-consistent field, each electron is given its own three-dimensional wave function. We thus have  $n$  three-dimensional wave functions,  $(q|1), (q|2), \dots, (q|n)$  say, the single variable  $q$  being written to denote the three Cartesian coordinates  $x, y, z$  and also

a spin variable. These wave functions, according to Hartree, satisfy equations of the form of Schrödinger's equation, thus†

$$i\hbar \frac{\partial}{\partial t} (q' | r) = \int (q' | H(r) | q'') dq'' (q'' | r) \dots\dots\dots(1),$$

where  $(q' | H(r) | q'')$  is the matrix representing the Hamiltonian for the  $r$ th electron. This Hamiltonian is of the form

$$H(r) = H_0 + \sum_{s \neq r} B_s \dots\dots\dots(2),$$

where  $H_0$  is the Hamiltonian for an electron moving in the field of the nucleus alone and  $B_s$  is the change in this Hamiltonian caused by the distribution of electric charge corresponding to the wave function  $(q | s)$ . We can express  $B_s$  in terms of the interaction energy  $V$  between two electrons. This  $V$  will be represented by a matrix of the form  $(q_1' q_2' | V | q_1'' q_2'')$ , which satisfies

$$(q_1' q_2' | V | q_1'' q_2'') = (q_2' q_1' | V | q_2'' q_1'') \dots\dots\dots(3),$$

on account of the symmetry between the two electrons. We shall then have

$$(q' | B_s | q'') = \iint (s | q''') dq''' (q' q''' | V | q'' q^{iv}) dq^{iv} (q^{iv} | s) \dots\dots\dots(4),$$

provided that the wave function  $(q | s)$  and its conjugate complex  $(s | q)$  are normalised. If  $V$  includes only the Coulomb force between the electrons, we have

$$(q_1' q_2' | V | q_1'' q_2'') = e^2 / r(q_1' q_2') \cdot \delta(q_1' - q_1'') \delta(q_2' - q_2'') \dots\dots\dots(5),$$

where  $r(q_1' q_2')$  is the distance between the two points specified by  $q_1'$  and  $q_2'$ , and equation (4) reduces to

$$(q' | B_s | q'') = e^2 \delta(q' - q'') \int | (q''' | s) |^2 / r(q' q''') \cdot dq'''.$$

Our wave functions  $(q | 1), (q | 2), \dots, (q | n)$  should be all orthogonal to one another. Further, if  $(q | 1)^*, (q | 2)^*, \dots, (q | n)^*$  are any set of  $n$  orthogonal independent linear functions of the  $(q | 1), (q | 2), \dots, (q | n)$ , then the  $(q | )^*$ 's should describe the distribution of electrons just as well as the  $(q | )$ 's, and should therefore also satisfy the equations (1), (2), (4) for the self-consistent field. This, however, is not the case. We must therefore modify our equations to give them this invariant property, which may be done in the following way.

† The integral sign is understood to include a summation over both states of spin.

We denote the matrix defined in (4) by  $B_{ss}$  instead of  $B_s$ , and generalise this definition to

$$(q' | B_{sr} | q'') = \iint (s | q''') dq''' (q' q''' | V | q'' q^{iv}) dq^{iv} (q^{iv} | r) \dots\dots(6).$$

We then assume the following wave equations:

$$i\hbar \frac{\partial}{\partial t} (q' | r) = \int (q' | H_0 | q'') dq'' (q'' | r) + \Sigma_s \int (q' | B_{ss} | q'') dq'' (q'' | r) - \Sigma_s \int (q' | B_{sr} | q'') dq'' (q'' | s) \dots\dots\dots(7).$$

It is evident that equations (6) and (7) will remain invariant under a transformation to the  $(q|)^*$ 's, since these equations are of tensor form with respect to the suffixes  $r$  and  $s$  (the left and right positions of a suffix in a symbol corresponding to the covariant and contravariant positions in relativity theory). Equations (7) are equivalent to Fock's equations, with the difference that Fock uses energy parameters instead of the operator  $i\hbar \partial/\partial t$  and also his equations do not take the spin variables into account but allow two electrons to be in the same state instead.

It is easily verified that, with equations (7), if the wave functions are initially all orthogonal they always remain orthogonal. We have, denoting  $\Sigma_s B_{ss}$  by  $B$  for brevity,

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \int (u | q') dq' (q' | r) &= \int (u | q') dq' \left[ i\hbar \frac{\partial}{\partial t} (q' | r) \right] - \int \left[ -i\hbar \frac{\partial}{\partial t} (u | q') \right] dq' (q' | r) \\ &= \iint (u | q') dq' (q' | H_0 + B | q'') dq'' (q'' | r) \\ &\quad - \Sigma_s \iint (u | q') dq' (q' | B_{sr} | q'') dq'' (q'' | s) \\ &\quad - \iint (u | q'') dq'' (q'' | H_0 + B | q') dq' (q' | r) \\ &\quad + \Sigma_s \iint (s | q'') dq'' (q'' | B_{rs} | q') dq' (q' | r). \end{aligned}$$

The first and third terms here cancel, while the second term becomes, with the help of (6),

$$- \Sigma_s \iiint (u | q') dq' (s | q''') dq''' (q' q''' | V | q'' q^{iv}) dq^{iv} (q^{iv} | r) dq'' (q'' | s),$$

and the fourth becomes similarly

$$\Sigma_s \iiint (s | q'') dq'' (u | q''') dq''' (q'' q''' | V | q' q^{iv}) dq^{iv} (q^{iv} | s) dq' (q' | r).$$

These two terms also cancel, on account of (3), and hence

$$\int (u | q') dq' (q' | r)$$

is constant and always vanishes if it vanishes initially. This result still holds when the atom is perturbed in any way, provided that the perturbing energy is symmetrical between all the electrons.

The total electron density is, by definition, represented by the matrix

$$(q' | \rho | q'') = \Sigma_r (q' | r) (r | q'') \dots \dots \dots (8).$$

It has the property that its square is equal to itself, since

$$\begin{aligned} (q' | \rho^2 | q'') &= \int (q' | \rho | q''') dq''' (q''' | \rho | q'') \\ &= \Sigma_{rs} \int (q' | r) (r | q''') dq''' (q''' | s) (s | q'') \\ &= \Sigma_{rs} (q' | r) \delta_{rs} (s | q'') = (q' | \rho | q''), \end{aligned}$$

on account of the orthogonality and normalization conditions for the wave functions. The condition  $\rho^2 = \rho$  is equivalent to the condition that the eigenvalues of  $\rho$  are all 0 or 1.

We shall now obtain the equation of motion for  $\rho$ . We have

$$\begin{aligned} i\hbar \frac{d}{dt} (q' | \rho | q'') &= \Sigma_r \left\{ \left[ i\hbar \frac{\partial}{\partial t} (q' | r) \right] (r | q'') - (q' | r) \left[ -i\hbar \frac{\partial}{\partial t} (r | q'') \right] \right\} \\ &= \Sigma_r \left\{ \int (q' | H_0 + B | q''') dq''' (q''' | r) (r | q'') \right. \\ &\quad - \Sigma_s \int (q' | B_{rs} | q''') dq''' (q''' | s) (s | q'') \\ &\quad - \int (q' | r) (r | q''') dq''' (q''' | H_0 + B | q'') \\ &\quad \left. + \Sigma_s \int (q' | r) (s | q''') dq''' (q''' | B_{rs} | q'') \right\} \\ &= \int (q' | H_0 + B | q''') dq''' (q''' | \rho | q'') \\ &\quad - \iiint (q' q^{1v} | V | q''' q^v) dq''' dq^{1v} dq^v (q''' | \rho | q^{1v}) (q^v | \rho | q'') \\ &\quad - \int (q' | \rho | q''') dq''' (q''' | H_0 + B | q'') \\ &\quad + \iiint (q''' q^v | V | q'' q^{1v}) dq''' dq^{1v} dq^v (q' | \rho | q^v) (q^{1v} | \rho | q''), \end{aligned}$$

if one substitutes for  $B_{sr}$  its value given by (6) and eliminates the wave functions with the help of (8). This result may be written symbolically

$$i\hbar \dot{\rho} = (H_0 + B - A)\rho - \rho(H_0 + B - A) \dots\dots\dots(9),$$

where  $A$  is represented by the matrix

$$\begin{aligned} (q' | A | q'') &= \iint (q' q^{1v} | V | q''' q'') dq''' dq^{1v} (q''' | \rho | q^{1v}) \\ &= \iint (q''' q' | V | q'' q^{1v}) dq''' dq^{1v} (q^{1v} | \rho | q''') \\ &\dots\dots\dots(10), \end{aligned}$$

these two expressions being equal on account of (3). A corresponding definition for  $B$  is

$$(q' | B | q'') = \iint (q' q''' | V | q'' q^{1v}) dq''' dq^{1v} (q^{1v} | \rho | q''') \dots\dots\dots(11).$$

We now have equations (9), (10) and (11) describing the motion of the distribution of electrons and involving only the one unknown  $\rho$ . We can therefore deal with the distribution of electrons by working with just the total electron density  $\rho$  and need not bother about the individual wave functions  $(q|1)$ ,  $(q|2)$ , ...,  $(q|n)$ . The equation of motion for  $\rho$ , namely equation (9), is of the standard form in quantum mechanics with the Hamiltonian

$$H = H_0 + B - A \dots\dots\dots(12).$$

This Hamiltonian is unusual in that it is not independent of  $\rho$ , but contains a part  $B - A$  which depends linearly on  $\rho$ .

In the case of a Coulomb law of interaction, for which  $V$  is given by (5), equations (11) and (10) for  $B$  and  $A$  reduce to

$$\left. \begin{aligned} (q' | B | q'') &= e^2 \delta(q' - q'') \int \frac{(q''' | \rho | q''')}{r(q' q''')} dq''' \\ (q' | A | q'') &= e^2 \frac{(q' | \rho | q'')}{r(q' q'')} \end{aligned} \right\} \dots\dots(13).$$

$B$  is just the potential produced by the total distribution of charge and is to be expected in the Hamiltonian  $H$ .  $A$  is an extra term representing the exchange effects.

When  $H_0$  does not involve the time explicitly, the equation of motion (9) admits of a simple integral, namely the quantity

$$D \{ \rho (H_0 + \frac{1}{2} B - \frac{1}{2} A) \},$$

where  $D$  denotes the diagonal sum (or integral) of the matrix following it. To verify the constancy of this quantity, we observe that, from (11),

$$\begin{aligned} D(\rho \dot{B}) &= \iint (q'' | \rho | q') (q' | \dot{B} | q'') dq' dq'' \\ &= \iiint (q'' | \rho | q') (q' q''' | V | q'' q^{iv}) (q^{iv} | \dot{\rho} | q''') dq' dq'' dq''' dq^{iv} \\ &= \iiint (q^{iv} | \dot{\rho} | q''') (q''' q' | V | q^{iv} q'') (q'' | \rho | q') dq' dq'' dq''' dq^{iv} \\ &= D(\dot{\rho} B), \end{aligned}$$

and similarly  $D(\rho \dot{A}) = D(\dot{\rho} A)$ .

$$\begin{aligned} \text{Hence } \frac{d}{dt} D\{\rho(H_0 + \tfrac{1}{2}B - \tfrac{1}{2}A)\} \\ &= D\{\dot{\rho}(H_0 + \tfrac{1}{2}B - \tfrac{1}{2}A)\} + \tfrac{1}{2}D(\rho \dot{B}) - \tfrac{1}{2}D(\rho \dot{A}) \\ &= D\{\dot{\rho}(H_0 + \tfrac{1}{2}B - \tfrac{1}{2}A)\} + \tfrac{1}{2}D(\dot{\rho} B) - \tfrac{1}{2}D(\dot{\rho} A) \\ &= D(\dot{\rho} H). \end{aligned}$$

$$\begin{aligned} \text{Thus } i\hbar \frac{d}{dt} D\{\rho(H_0 + \tfrac{1}{2}B - \tfrac{1}{2}A)\} &= D\{(H\rho - \rho H)H\} \\ &= D(H\rho H) - D(\rho H^2) = 0, \end{aligned}$$

since the diagonal sum of a product is not changed by a cyclic permutation of the factors. Hence  $D\{\rho(H_0 + \tfrac{1}{2}B - \tfrac{1}{2}A)\}$  is a constant of the motion. It may be interpreted as the energy integral,  $D(\rho H_0)$  being the proper energy of the electrons (their kinetic energy plus their potential energy in the field of the nucleus),  $\tfrac{1}{2}D(\rho B)$  being their interaction energy and  $-\tfrac{1}{2}D(\rho A)$  being a correction for exchange effects.

### *Reduction to a classical density function.*

We shall now examine what the equation of motion (9) becomes when the electron density  $\rho$  is spread over such a large volume of phase space that we can neglect the fact that the momenta  $p$  do not commute with the coordinates  $q$  and reduce our description of the atom to a classical one. We shall also now neglect the spin variables. Each element  $(q' | \alpha | q'')$  of the matrix representing any dynamical variable  $\alpha$  will now be connected with a certain Fourier component in the  $p$ -variables of  $\alpha(qp)$  considered as a function of commuting  $q$ 's and  $p$ 's. We shall have, in fact,

$$(q' | \alpha | q'') = (2\pi\hbar)^{-3} \int \alpha(qp) e^{i(q'-q'')p/\hbar} dp \dots\dots\dots (14),$$

the connection being most accurate when the  $q$  on the right-hand side is taken to be the mean of  $q'$  and  $q''$ . The converse equation will be

$$a(qp) = \int (q' | a | q'') e^{-i(q'-q'')p/\hbar} d(q' - q'') \dots (15).$$

If we apply equation (15) to the matrix  $(q' | B | q'')$  defined by the first of equations (13), we obtain for the corresponding classical function  $B(qp)$  of commuting  $q$ 's and  $p$ 's,

$$\begin{aligned} B(qp) &= e^2 \int \delta(q' - q'') e^{-i(q'-q'')p/\hbar} d(q' - q'') \cdot \int \frac{(q''' | \rho | q''')}{r(qq''')} dq''' \\ &= e^2 \int \frac{(q''' | \rho | q''')}{r(qq''')} dq''' \\ &= \frac{e^2}{(2\pi\hbar)^3} \int \frac{dq'}{r(qq')} \int \rho(q'p') dp' \dots (16), \end{aligned}$$

with the help of (14) applied to  $\rho$ . Thus  $B(qp)$  is independent of  $p$ , as it must be since the matrix  $(q' | B | q'')$  is diagonal, and it is just that function of  $q$  which corresponds to the potential arising from a distribution of electrons of density  $\rho(qp)$  per volume  $(2\pi\hbar)^3$  of phase space.

If we now apply (15) to the matrix  $(q' | A | q'')$  defined by the second of equations (13), we obtain

$$\begin{aligned} A(qp) &= e^2 \int \frac{(q' | \rho | q'')}{r(q'q'')} e^{-i(q'-q'')p/\hbar} d(q' - q'') \\ &= \frac{e^2}{(2\pi\hbar)^3} \int \rho(qp') dp' \int \frac{e^{-i(q'-q'')(p-p')/\hbar}}{r(q'q'')} d(q' - q''), \end{aligned}$$

with the help of (14) applied to  $\rho$ . The second integral here can be evaluated. The  $q' - q''$  and  $p - p'$  appearing in the exponential are really vectors and their product should be understood to mean their scalar product. If we denote by  $\theta$  the angle between these two vectors and by  $\sigma$  the magnitude of  $q' - q''$ , which is the same as  $r(q'q'')$ , we have for this second integral

$$\begin{aligned} \int \frac{e^{-i(q'-q'')(p-p')/\hbar}}{r(q'q'')} d(q' - q'') &= \int_0^\infty \sigma d\sigma \int_{-1}^1 e^{-i\sigma|p-p'|\cos\theta/\hbar} 2\pi d(\cos\theta) \\ &= \int_0^\infty 4\pi\hbar d\sigma \sin\{\sigma|p-p'|/\hbar\} / |p-p'| \\ &= 4\pi\hbar^2 / |p-p'|^2. \end{aligned}$$

Hence

$$A(qp) = \frac{e^2}{2\pi^2\hbar} \int \frac{\rho(qp')}{|p-p'|^2} dp' \dots (17).$$



These expressions for  $B(q)$  and  $A(qp)$  substituted in (12) will give us a classical Hamiltonian governing the motion of the distribution of electrons. To proceed with the solution we shall now make an assumption concerning the form of  $\rho$ , which seems to be plausible when we are dealing with the state of lowest energy of the atom, namely the assumption that for each value of  $q$  the phase space is saturated, with two electrons per volume  $(2\pi\hbar)^3$ , in a region for which the magnitude of the momentum  $p$  is less than a certain value  $P$ , and is empty outside this region. In symbols

$$\rho(qp) = 2 \quad |p| < P \\ = 0 \quad |p| > P,$$

where  $P$  is a certain function of  $q$ . This assumption gives us at once from (16)

$$B(q) = \frac{e^2}{3\pi^2\hbar^3} \int \frac{P^3(q') dq'}{r(qq')^2} \dots\dots\dots(18),$$

and from (17), after a straightforward integration,

$$A(qp) = \frac{e^2}{\pi\hbar} \left[ \frac{P^2 - |p|^2}{|p|} \log \left| \frac{\dot{P} + |p|}{P - |p|} \right| + 2P \right] \dots(19).$$

For a stationary state of the atom  $\rho$  must be constant, so that the Poisson bracket of  $\rho$  with  $H$  must vanish. With  $\rho$  of the form which we have assumed, this condition becomes that  $H$  must be constant along the boundary between the saturated and unoccupied phase space, i.e.  $H(qP)$  must be constant. Now

$$H(qP) = H_0(qP) + B(q) - 2e^2/\pi\hbar \cdot P.$$

The constancy of this gives us a condition for the unknown function  $P$ .

For an atom with atomic number  $Z$ , we shall have

$$H_0 = -\frac{Ze^2}{r} + \frac{p^2}{2m},$$

where  $r$  denotes distance from the nucleus. We may assume spherical symmetry, so that  $P$  is a function of  $r$  only. We now get for the value of  $H$  on the boundary

$$H(qP) = -\frac{Ze^2}{r} + \frac{P^2}{2m} - \frac{2e^2P}{\pi\hbar} \\ + \frac{4e^2}{3\pi\hbar^3} \left[ \frac{1}{r} \int_0^r r'^2 P^3(r') dr' + \int_r^\infty r' P^3(r') dr' \right].$$

By equating to zero the differential coefficient of this with respect to  $r$ , we get

$$\frac{Ze^2}{r^2} + \frac{d}{dr} \left( \frac{P^2}{2m} - \frac{2e^2P}{\pi\hbar} \right) - \frac{4e^2}{3\pi\hbar^2} \frac{1}{r^2} \int_0^r r'^2 P^3(r') dr' = 0.$$

Multiplication by  $r^2$  and a further differentiation then give

$$\frac{d}{dr} \left\{ r^2 \frac{d}{dr} \left( \frac{P^2}{2m} - \frac{2e^2}{\pi h} P \right) \right\} = \frac{4e^2}{3\pi h^2} r^2 P^3 \dots\dots\dots(20).$$

This is a differential equation which determines  $P$ , the maximum momentum for an electron, as a function of  $r$ . It differs from the corresponding equation in Thomas' theory only on account of its having a term linear in  $P$  on the left-hand side, which term may be considered to represent the exchange effects. This term will not be very important in the interior of a heavy atom, since the ratio of its coefficient to that of the  $P^2$  term is  $4e^2m/\pi h$ , which is  $4/\pi$  times the momentum of an electron in the first Bohr orbit in the hydrogen atom. For a sufficiently large value of  $r$ , however, the extra term causes  $P$  to become negative, and then to oscillate with decreasing amplitude and increasing period as  $r \rightarrow \infty$ . A negative value for  $P$  has, of course, no physical meaning. The fact that our theory gives this meaningless result for the outside of the atom is hardly surprising, since the approximation we made of regarding  $\rho$  as a function of commuting  $q$ 's and  $p$ 's is certainly not valid for this region. We may expect equation (20) to be more accurate than Thomas' equation in the interior of the atom, in spite of the fact that it is inapplicable outside.