Quantum Mechanics of Many-Electron Systems.

By P. A. M. Dirac, St. John's College, Cambridge.

(Communicated by R. H. Fowler, F.R.S.—Received March 12, 1929.)

§ 1. *Introduction.*

The general theory of quantum mechanics is now almost complete, the imperfections that still remain being in connection with the exact fitting in of the theory with relativity ideas. These give rise to difficulties only when high-speed particles are involved, and are therefore of no importance in the consideration of atomic and molecular structure and ordinary chemical reactions, in which it is, indeed, usually sufficiently accurate if one neglects relativity variation of mass with velocity and assumes only Coulomb forces between the various electrons and atomic nuclei. The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.

Already before the arrival of quantum mechanics there existed a theory of atomic structure, based on Bohr's ideas of quantised orbits, which was fairly successful in a wide field. To get agreement with experiment it was found necessary to introduce the spin of the electron, giving a doubling in the number of orbits of an electron in an atom. With the help of this spin and Pauli's exclusion principle, a satisfactory theory of multiplet terms was obtained when one made the additional assumption that the electrons in an atom all set themselves with their spins parallel or antiparallel. If s denoted the magnitude of the resultant spin angular momentum, this s was combined vectorially with the resultant orbital angular momentum l to give a multiplet of multiplicity $2s+1$. The fact that one had to make this additional assumption was, however, a serious disadvantage, as no theoretical reasons to support it could be given. It seemed to show that there were large forces coupling the spin vectors of the electrons in an atom, much larger forces than could be accounted for as due to the interaction of the magnetic moments of the electrons. The position was thus that there was empirical evidence in favour of these large forces, but that their theoretical nature was quite unknown.

The old orbit theory is now replaced by Hartree's method of the self-consistent field,* based on quantum mechanics. The simplifying feature of the old theory, according to which each electron has its own individual orbit, is retained, but the " orbit " is now a quantum-mechanical state of the single electron, represented by a wave function in three dimensions. The only action of one orbit on another is assumed to be that of a static distribution of electricity, causing a partial screening of the nucleus. A theoretical justification for Hartree's method, showing that its results must be in approximate agreement with those of the exact Schrödinger equation for the whole system, has been given by Gaunt.† The method, however, suffers from the same limitation as the old orbit theory. It cannot give an explanation of multiplet structure without an extraneous assumption of large forces coupling the spins.

The solution of this difficulty in the explanation of multiplet structure is provided by the *exchange* (austausch) interaction of the electrons, which arises owing to the electrons being indistinguishable one from another. Two electrons may change places without our knowing it, and the proper allowance for the possibility of quantum jumps of this nature, which can be made in a treatment of the problem by quantum mechanics, gives rise to the new kind of interaction. The energies involved, the so-called exchange energies, are quite large. In fact it is these exchange energies between electrons in different atoms that give rise to homopolar valency bonds, as shown by Heitler and London.‡

The application of the new exchange ideas to the problem of multiplet structure has been made by Wigner§ and Hund.‖ The new theory provides no justification for the assumption that the electrons all set themselves with their spins parallel or antiparallel. In fact it does not allow any meaning to be given to this assumption, since in quantum mechanics the component of the spin angular momentum of an electron in any direction is a q-number with the two eigen-values $\pm \frac{1}{2}h$, so that one cannot *in general* give a meaning to the direction of the spin of an electron in a given stationary state. What the

* D. R. Hartree, ' Proc. Camb. Phil. Soc.,' vol. 24, p. 89 (1928).

† J. A. Gaunt, ' Proc. Camb. Phil. Soc.,' vol. 24, p. 328 (1928). It is pointed out by Gaunt that there does not seem to be any theoretical justification for Hartree's method of calculating energies and that its extremely good agreement with observation is probably accidental. The somewhat different method proposed by Gaunt is the one that should be used in connection with the present paper.

‡ W. Heitler and F. London, ' Z. Physik,' vol. 44, p. 455 (1927).

§ E. Wigner, ' Z. Physik,' vol. 43, p. 624 (1927).

‖ F. Hund, ' Z. Physik,' vol. 43, p. 788 (1927).

new theory shows instead is that *for each stationary state of the atom there is one definite numerical value for s, the magnitude of the total spin vector.* If it were not for this theorem, a measurement of s for the atom in a given stationary state would lead to one or other of a number of possible results, according to a definite probability law. This theorem forms the basis of the theory of multiplets. It is quite sufficient to replace the previous idea of the electrons all setting themselves parallel or antiparallel, since it shows that we can take s to be a quantum number describing the states of the atom, while s combined vectorially with l gives a multiplet of multiplicity $2s + 1$.

Further developments of the theory of exchange have been made by Heitler, London and Heisenberg,* containing applications to molecules held together by homopolar valency bonds and to ferromagnetism. The treatment given by these authors makes an extensive use of group theory and requires the reader to be well acquainted with this branch of pure mathematics. Now group theory is just a theory of certain quantities that do not satisfy the commutative law of multiplication, and should thus form a part of quantum mechanics, which is the general theory of all quantities that do not satisfy the commutative law of multiplication. It should therefore be possible to translate the methods and results of group theory into the language of quantum mechanics and so obtain a treatment of the exchange phenomena which does not presuppose any knowledge of groups on the part of the reader. This is the object of the present paper. The treatment of groups on the lines of quantum mechanics has the advantage that it often gives a simple physical meaning to an abstract theorem in the theory of groups, enabling one to remember the theorem more easily and perhaps suggesting a simpler way of proving it. A further advantage of the treatment of the exchange phenomena on these lines is that one can avoid doing more work in the theory of groups than is strictly necessary for the physical applications, which results in a considerable shortening in the method.

In §§ 2 and 3 the general theory is given of systems containing a number of similar particles, showing the existence of exclusive sets of states (*i.e.*, sets such that a transition can never take place from a state in one set to a state in another), and giving their main properties. In § 4 an application is made to electrons, a proof being obtained of the fundamental theorem in italics above. The subsequent work is concerned with an approximate calculation of the energy levels of the states, the result of this being expressible by the single

* See various papers in the 'Z. Physik,' vols. 46–51. An excellent account of the whole theory is also contained in Weyl's book, 'Gruppentheorie und Quantummechanik.'

simple formula (26). This formula shows that in the first approximation the exchange interaction between the electrons may be replaced by a coupling between their spins, the energy of this coupling for each pair of electrons being equal to the scalar product of their spin vectors multiplied by a numerical coefficient given by the exchange energy. This form of coupling energy is, however, just what was required in the old orbit theory. We obtain in this way a justification for the assumptions of this old theory, in so far as they can be formulated without contradicting the quantum-mechanical description of the spin. The formula (26), combined with Hartree's method for determining approximate wave functions for the different electrons, should provide a powerful way of dealing with complicated atomic systems.

§ 2. *Permutations as Dynamical Variables.*

We consider a dynamical system composed of n similar particles, the rth particle being describable by certain generalised co-ordinates denoted by the single symbol q_r. Thus a wave function representing a state of the system will be a function of the variables $q_1, q_2, \ldots, q_n$, which may be written

$$\psi(q_1, q_2 \ldots q_n) = \psi(q)$$

for brevity. Suppose now that P is any permutation of $q_1, q_2, \ldots, q_n$. This P is an operator which can be applied to any wave function $\psi(q)$ to give as result another definite function of the q's, namely

$$\mathrm{P}\psi(q) = \psi(\mathrm{P}q),$$

where $\mathrm{P}q$ denotes the set of q's obtained by applying the permutation P to $q_1, q_2, \ldots, q_n$. Further P is a linear operator. Now in quantum mechanics any dynamical variable is a linear operator which can operate on any wave function, and conversely any linear operator that can operate on every wave function may be considered as a dynamical variable. Thus *any permutation* P *may be considered to be a dynamical variable.*

The present paper consists in a study of these permutations P as dynamical variables. There are no classical analogues to these variables and hence they give rise to phenomena, *e.g.*, the existence of exclusive sets of states and other exchange phenomena, which have no classical analogue. There are $n!$ of these variables, one of them, P_1 say, being the identity, which must thus be equal to unity. One can add and multiply these variables and form algebraic functions of them, in exactly the same way in which one can add and multiply and form algebraic functions of the ordinary co-ordinates and momenta. The product of any two permutations is a third permutation, and hence any function

of the permutations is reducible to a linear function of them. Any permutation P has a reciprocal P^{-1} satisfying $PP^{-1} = P^{-1}P = P_1 = 1$.

A permutation P, like any other dynamical variable, can be represented by a matrix. If we take the representation in which the q's are diagonal, P will be represented by a matrix, whose general element may be written

$$(q_1' q_2' \dots q_n' \mid P \mid q_1'' q_2'' \dots q_n'') = (q' \mid P \mid q'')$$

for brevity. This matrix must satisfy

$$\int (q' \mid P \mid q'')\, dq''\, \psi(q'') = P\psi(q') = \psi(Pq'),$$

and hence

$$(q' \mid P \mid q'') = \delta(Pq' - q''). \tag{1}$$

We are using the notation $\delta(x)$, where x is short for a set of variables $x_1, x_2, x_3, \dots$, to denote

$$\delta(x) = \delta(x_1)\,\delta(x_2)\,\delta(x_3)\dots$$

which vanishes except when each of the x's vanishes. With this notation we have

$$\delta(Pq' - q'') = \delta(q' - P^{-1}q''),$$

since the condition that the left-hand side shall not vanish, which is that the q'''s shall be given by applying the permutation P to the q''s, is the same as the condition that the right-hand side shall not vanish, which is that the q''s shall be given by applying the permutation P^{-1} to the q'''s. Thus we have an alternative expression for the matrix representing P.

$$(q' \mid P \mid q'') = \delta(q' - P^{-1}q''). \tag{2}$$

The conjugate complex of any dynamical variable is given when one writes $-i$ for i in the matrix representing that variable and also interchanges the rows with the columns. Thus we find for the conjugate complex of a permutation P, with the help of (2) and (1)

$$(q' \mid \bar{P} \mid q'') = \overline{(q'' \mid P \mid q')} = \delta(q'' - P^{-1}q')$$
$$= (q' \mid P^{-1} \mid q'')$$

or

$$\bar{P} = P^{-1}.$$

Thus a permutation is not in general a real variable, its conjugate complex being equal to its reciprocal.

Any permutation of the numbers 1, 2, 3, ..., n may be expressed in the cyclic notation, *e.g.*, for $n = 8$

$$P_a = (143)(27)(58)(6), \tag{3}$$

in which each number is to be replaced by the succeeding number in a bracket, unless it is the last in a bracket, when it is to be replaced by the first in that bracket. Thus P_a changes the numbers 12345678 into 47138625. The type of any permutation is specified by the partition of the number n which is provided by the number of numbers in each of the brackets. Thus the type of P_a is specified by the partition $8 = 3 + 2 + 2 + 1$. Permutations of the same type, *i.e.*, corresponding to the same partition, we shall call *similar*. (The usual language of group theory is to call them conjugate.) Thus, for example, P_a in (3) is similar to

$$P_b = (871)\,(35)\,(46)\,(2). \tag{4}$$

The whole of the $n!$ possible permutations may be divided into sets of similar permutations, each such set being called a *class*. The permutation $P_1 = 1$ forms a class by itself. Any permutation is similar to its reciprocal.

When two permutations P_a and P_b are similar, either of them P_b may be obtained by making a certain permutation P in the other P_a. Thus, in our example (3), (4) we can take P to be the permutation that changes 14327586 into 87135462, *i.e.*, the permutation

$$P = (18623)\,(475).$$

We then have the algebraic relation between P_b and P_a

$$P_b = PP_aP^{-1}. \tag{5}$$

To verify this, we observe that the product $P_a\psi$ of P_a with any wave function ψ is changed into $P_b\psi$ if one applies the permutation P to the P_a in the product but not to the ψ. If we multiply the product by P on the left we are applying this permutation to both the P_a and the ψ, so that we must insert another factor P^{-1} between the P_a and the ψ, giving us $PP_aP^{-1}\psi$ to equate to $P_b\psi$.

Equation (5) is the general formula showing when two permutations P_a and P_b are similar. Of course P is not uniquely determined when P_a and P_b are given, but the existence of any P satisfying (5) is sufficient to show that P_a and P_b are similar.

§ 3. *Permutations as Constants of the Motion.*

We now introduce a Hamiltonian H to describe the motion of the system, so that any stationary state of energy H′ is represented by a wave function ψ satisfying

$$H\psi = H'\psi,$$

in which H is regarded as an operator. This Hamiltonian can be an arbitrary function of the dynamical variables provided it is symmetrical between all the

particles. This symmetry condition requires that an element $(q' \mid \mathrm{H} \mid q'')$ of the matrix representing H shall be unaltered when one applies any permutation to the q''s and the same permutation to the q'''s, *i.e.*,

$$(q' \mid \mathrm{H} \mid q'') = (\mathrm{P}q' \mid \mathrm{H} \mid \mathrm{P}q'') \tag{6}$$

for arbitrary P.

The fact that H is symmetrical leads at once to the equation

$$\mathrm{PH} = \mathrm{HP}. \tag{7}$$

This equation may be verified by a similar argument to that used for equation (5), or alternatively by a direct application of the matrix representatives. Thus from (1)

$$(q' \mid \mathrm{PH} \mid q'') = \int \delta(\mathrm{P}q' - q''')\, dq'''\, (q''' \mid \mathrm{H} \mid q'') = (\mathrm{P}q' \mid \mathrm{H} \mid q'')$$

and from (2)

$$(q' \mid \mathrm{HP} \mid q'') = \int (q' \mid \mathrm{H} \mid q''')\, dq'''\, \delta(q''' - \mathrm{P}^{-1}q'') = (q' \mid \mathrm{H} \mid \mathrm{P}^{-1}q''),$$

and the two right-hand sides are now equal from (6). Equation (7) shows that *each permutation variable is a constant of the motion.* The P's are still constants when arbitrary perturbations are applied to the system, provided the perturbation energy to be added to the Hamiltonian is symmetrical. Thus the constancy of the P's is absolute.

In dealing with any system in quantum mechanics, when we have found a constant of the motion α, we know that if for any state α initially has the numerical value α' then it always has this value, so that we can assign different numbers α' to the different states and so obtain a classification of the states. This procedure is not so straightforward, however, when we have several constants of the motion α which do not commute (as is the case with our permutations P), since we cannot assign numerical values for all the α's simultaneously to any state. The existence of constants of the motion α which do not commute is a sign that the system is degenerate. We must now look for a function β of the α's which has one and the same numerical value β' for all those states belonging to one energy level H′, so that we can use β for classifying the energy levels of the system. We can express the condition for β by saying that it must be a function of H (a single-valued function is implied) according to the general definition of a function of a variable in quantum mechanics, or that β must commute with every variable that commutes with H, *i.e.* every constant of the motion. If the α's are the only constants of the motion, or if they are a set that commute with all other independent constants of the motion, our problem reduces to finding a function β of the α's which commutes

with all the α's. We can then assign a numerical value β' for β to each energy level of the system. If we can find several such functions β, they must all commute with each other, so that we can give them all numerical values simultaneously and obtain a complete classification of the energy levels.

An example of this procedure is provided by the study of the angular momentum of an isolated system. This angular momentum has three components m_x, m_y, m_z, each a constant of the motion, which do not commute. We look for a function of m_x, m_y, m_z which commutes with them all three. We can conveniently take for this function the variable k defined by

$$k(k+h) = m_x^2 + m_y^2 + m_z^2. \tag{8}$$

For each energy level of the system there will now be one definite numerical value k' for k. This constant of the motion k is the only significant one for purposes of classifying the states, as the others merely describe the degeneracy.

We follow this method in dealing with our permutations P. We must find a function χ of the P's such that $P\chi P^{-1} = \chi$ for every P. It is evident that a possible χ is ΣP_c, the sum of all the permutations P_c in a certain class c, *i.e.* the sum of a set of similar permutations, since ΣPP_cP^{-1} must consist of the same permutations summed in a different order. There will be one such χ for each class. Further, there can be no other independent χ, since an arbitrary function of the P's can be expressed as a linear function of them with numerical coefficients and it will not then commute with every P unless the coefficients of similar P's are always the same. We thus obtain all the χ's that can be used for classifying the states. It is convenient to define each χ as an average instead of a sum, thus

$$\chi_c = \Sigma P_c / n_c,$$

where n_c is the number of P's in the class c. An alternative expression for χ_c is

$$\chi_c = \Sigma_r P_r P_c P_r^{-1} / n!, \tag{9}$$

the summation being extended over all the $n!$ permutations P_r. For each permutation P there is one χ, $\chi(P)$ say, equal to the average of all permutations similar to P. One of the χ's is $\chi(P_1) = 1$.

The dynamical variables $\chi_1, \chi_2 \cdots \chi_m$ obtained in this way will each have a definite numerical value for every stationary state of the system. Thus for every permissible set of numerical values $\chi_1', \chi_2' \cdots \chi_m'$ for the χ's there will be a set of states of the system. Since the χ's are absolute constants of the motion these sets of states will be exclusive, *i.e.* transitions will never take place from a state in one set to a state in another.

The permissible sets of values χ' that one can give to the χ's are limited by the fact that there exist algebraic relations between the χ's. The product of any two χ's, $\chi_p\chi_q$, is of course expressible as a linear function of the P's, and since it commutes with every P it must be expressible as a linear function of the χ's, thus

$$\chi_p\chi_q = a_1\chi_1 + a_2\chi_2 + \ldots + a_m\chi_m \tag{10}$$

where the a's are numbers. Any numerical values χ' that one gives to the χ's must be eigen-values of the χ's and must satisfy these same algebraic equations. For every solution χ' of these equations there is one exclusive set of states. One solution is evidently $\chi_p' = 1$ for every χ_p, and this gives the set of states with symmetrical wave functions. A second obvious solution is $\chi_p' = \pm 1$, the $+$ or $-$ sign being taken according to whether the permutations in the class p are even or odd, and this gives the set of states with antisymmetrical wave functions. The other solutions may be worked out in any special case by ordinary algebraic methods, as the coefficients a in (10) may be obtained directly by a consideration of the types of permutation to which the χ's concerned refer. Any solution is, apart from a certain factor, what is called in group theory a character of the group of permutations. The χ's are all real variables, since each P and its conjugate complex P^{-1} are similar and will occur added together in the definition of any χ, so that the χ''s must be all real numbers.

The number of possible solutions of the equations (10) may easily be determined, since it must equal the number of different eigen-values of an arbitrary function B of the χ's. We can express B as a linear function of the χ's with the help of equations (10); thus

$$B = b_1\chi_1 + b_2\chi_2 + \ldots + b_m\chi_m. \tag{11}$$

Similarly we can express each of the quantities $B^2, B^3 \ldots B^m$ as a linear function of the χ's. From these m equations, together with the equation $\chi(P_1) = 1$, we can eliminate the m unknowns $\chi_1, \chi_2 \ldots \chi_m$, obtaining as result an algebraic equation of degree m for B,

$$B^m + c_1B^{m-1} + c_2B^{m-2} + \ldots + c_m = 0.$$

The m solutions of this equation give the m possible eigen-values for B, each of which will, according to (11), be a linear function of $b_1, b_2 \ldots b_m$ whose coefficients are a permissible set of values $\chi_1', \chi_2' \ldots \chi_m'$. These m sets of values χ' thus obtained must be all different, since if there were fewer than m different permissible sets of values χ' for the χ's there would exist a linear function of the χ's every one of whose eigen-values vanishes, which would mean that the

linear function itself vanishes and the χ's are not independent. Thus the number of permissible sets of numerical values for the χ's is just equal to m, which is the number of classes of permutations or the number of partitions of n. This number is therefore the number of exclusive sets of states.

The properties of the P's which are not properties of the χ's will only describe the degeneracy of the states. If we obtain a matrix representation of all the P's consistent with each of the χ's being a certain number χ', then the number of rows and columns of the matrices will be the degree of degeneracy of the states in the exclusive set χ', *i.e.*, the number of independent states belonging to each energy level. This degeneracy is an *essential* one and cannot be removed by any perturbation that is symmetrical between all the similar particles.

§ 4. *Multiplet Structure.*

The preceding theory of systems composed of similar particles will now be applied to the case when the particles are electrons. The new features which this requires us to take into consideration are the spin of the electrons and Pauli's exclusion principle.

The three Cartesian co-ordinates x, y, z of the rth electron we denote by the single symbol x_r. The spin angular momentum and magnetic moment of this electron will be of the form $\frac{1}{2}h\boldsymbol{\sigma}_r$ and $\frac{1}{2}eh/mc\,.\,\boldsymbol{\sigma}_r$, where $\boldsymbol{\sigma}_r$ is a vector whose components σ_{rx}, σ_{ry}, σ_{rz} satisfy

$$\sigma_{rx}^2 = 1, \qquad \sigma_{rx}\sigma_{ry} = i\sigma_{rz} = -\sigma_{ry}\sigma_{rx}, \qquad (12)$$

with similar relations obtained by cyclic permutation of the suffixes x, y and z. We take x_r and σ_{rz} to be the variables describing the rth electron that appear in the wave function. It is convenient to write the wave function

$$\psi(x_1\sigma_1 x_2\sigma_2 \ldots x_n\sigma_n) = \psi(x\sigma)$$

without the suffixes z attached to the σ's, these suffixes being understood whenever one is dealing with the variables in wave functions.

The exclusion principle now requires that ψ shall be antisymmetrical in the x's and σ's together, *i.e.*, that if any permutation is applied to the x's and also to the σ's, ψ must remain unchanged or change sign according to whether the permutation is an even or an odd one. Thus permutations applied to the x's and σ's together, produce only trivial effects and no useful results would be obtained by considering them as dynamical variables. We can, however, consider permutations P applied to the x's alone and apply our preceding theory to these. Any of these permutations is a constant of the motion when we neglect the forces due to the spins, so that the Hamiltonian does not involve

the spin variables $\boldsymbol{\sigma}$. We can now introduce our χ's as functions of these P's and assert that for any permissible set of numerical values χ' for the χ's there will be one exclusive set of states. Thus there exist these exclusive sets of states for systems containing many electrons even when we restrict ourselves to a consideration only of those states that satisfy the exclusion principle. The exclusiveness of the sets of states is now, of course, only approximate, since the χ's are constants only when we neglect the spin forces. There will actually be a small probability for a transition taking place from a state in one set to a state in another.

Since ψ is antisymmetrical, the result of any permutation P applied to the x's must equal $\pm$ times the result when the same permutation is applied to the σ's. Thus if we denote by P^σ a permutation applied to the σ's considered as a dynamical variable, we shall have

$$\mathrm{P}_r = \pm \mathrm{P}^\sigma_r, \tag{13}$$

for each of the $n!$ permutations P_r. Thus instead of studying the dynamical variables P we can get all the results we want, *e.g.*, the characters χ', by studying the variables P^σ. The P^σ's are much easier to study on account of the fact that the variables σ in the wave function have domains consisting each of only the two points 1 and -1, which are the two eigen-values of each σ_z. This fact results in there being fewer characters χ' for the group of permutations of the σ variables than for the group of general permutations, since it prevents a function of the variables $\sigma_1, \sigma_2, \ldots$, from being antisymmetrical in more than two of them.

The study of the dynamical variables P^σ is made specially easy by the fact that we can express them as algebraic functions of the dynamical variables $\boldsymbol{\sigma}$. Consider the quantity

$$\mathrm{O}_{12} = \tfrac{1}{2}\{1 + (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)\}.$$

With the help of (12) we find readily that

$$(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)^2 = (\sigma_{1x}\sigma_{2x} + \sigma_{1y}\sigma_{2y} + \sigma_{1z}\sigma_{2z})^2 = 3 - 2(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2), \tag{14}$$

and hence that

$$\mathrm{O}_{12}{}^2 = \tfrac{1}{4}\{1 + 2(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2) + (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)^2\} = 1. \tag{15}$$

Again, we find

$$\mathrm{O}_{12}\sigma_{1x} = \tfrac{1}{2}\{\sigma_{2x} - i\sigma_{1z}\sigma_{2y} + i\sigma_{1y}\sigma_{2z} + \sigma_{1x}\}$$

$$\sigma_{2x}\mathrm{O}_{12} = \tfrac{1}{2}\{\sigma_{1x} + i\sigma_{1y}\sigma_{2z} - i\sigma_{1z}\sigma_{2y} + \sigma_{2x}\}$$

and hence

$$\mathrm{O}_{12}\sigma_{1x} = \sigma_{2x}\mathrm{O}_{12}.$$

Similar relations hold for σ_{1y} and σ_{1z}, so that we have

$$O_{12}\boldsymbol{\sigma}_1 = \boldsymbol{\sigma}_2 O_{12}$$

or

$$O_{12}\boldsymbol{\sigma}_1 O_{12}^{-1} = \boldsymbol{\sigma}_2.$$

From this we can obtain with the help of (15)

$$O_{12}\boldsymbol{\sigma}_2 O_{12}^{-1} = \boldsymbol{\sigma}_1.$$

These " Vertauschungs " relations for O_{12} with $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ are precisely the same as those for P^{σ}_{12}, the permutation consisting of the interchange of the spin variables of electrons 1 and 2. Thus we can put

$$O_{12} = cP^{\sigma}_{12}$$

where c is a number. Equation (15) shows that $c = \pm 1$. To determine which of these values for c is the correct one, we observe that the eigen-values of P^{σ}_{12} are 1, 1, 1, -1, corresponding to the fact that there exist three independent symmetrical and one antisymmetrical function $F(\sigma_{1z}, \sigma_{2z})$ of the two variables σ_{1z}, σ_{2z}, whose domains contain only two points each. Thus the mean of the eigen-values of P^{σ}_{12} is $\frac{1}{2}$. Now the mean of the eigen-values of $(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)$ is evidently zero and hence the mean of the eigen-values of O_{12} is $\frac{1}{2}$. Thus we must have $c = +1$, and so we can put

$$P^{\sigma}_{12} = \tfrac{1}{2}\{1 + (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)\}.$$

In this way any permutation P^{σ} consisting simply of an interchange can be expressed as an algebraic function of the $\boldsymbol{\sigma}$'s. Any other permutation P^{σ} can be expressed as a product of interchanges and can therefore also be expressed as a function of the $\boldsymbol{\sigma}$'s. With the help of (13) we can now express the P's as algebraic functions of the $\boldsymbol{\sigma}$'s and eliminate the P^{σ}'s from the discussion. We have, since the $-$ sign must be taken in (13) when the permutations are interchanges,

$$P_{12} = -\tfrac{1}{2}\{1 + (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)\}. \tag{16}$$

The formula (16) may conveniently be used for the evaluation of the characters χ' which define the exclusive sets of states. We have, for example, for the permutations consisting of interchanges

$$\chi_{12} = \chi(P_{12}) = -\tfrac{1}{2}\left\{1 + \frac{2}{n(n-1)}\Sigma_{r<t}(\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_t)\right\}.$$

If we introduce the variable s to describe the magnitude of the total spin angular momentum, $\frac{1}{2}\Sigma_r \boldsymbol{\sigma}_r$ in units of h, through the formula

$$s(s+1) = (\tfrac{1}{2}\Sigma_r \boldsymbol{\sigma}_r, \tfrac{1}{2}\Sigma_t \boldsymbol{\sigma}_t),$$

analogous to (8), we have

$$2\Sigma_{r<t}(\sigma_r, \sigma_t) = (\Sigma_r \sigma_r, \Sigma_t \sigma_t) - \Sigma_r (\sigma_r, \sigma_r).$$
$$= 4s(s+1) - 3n.$$

Hence

$$\chi_{12} = -\tfrac{1}{2}\left\{1 + \frac{4s(s+1) - 3n}{n(n-1)}\right\} = -\frac{n(n-4) + 4s(s+1)}{2n(n-1)}. \quad (17)$$

Thus χ_{12} is expressible as a function of the variable s and of n the number of electrons. Any of the other χ's could be evaluated on similar lines and would be found to be a function of s and n only, since there are no other symmetrical functions of all the σ variables which could be involved. There is therefore one set of numerical values χ' for the χ's, and thus one exclusive set of states, for each eigen-value s' of s. The eigen-values of s are

$$\tfrac{1}{2}n, \quad \tfrac{1}{2}n - 1, \quad \tfrac{1}{2}n - 2, \ldots$$

the series terminating with $\frac{1}{2}$ or 0.

We obtain in this way a proof of the fundamental theorem of multiplet structure, that for each stationary state of the atom there is one definite numerical value s' for s. We obtain further that the probability of transitions occurring in which s changes is small, of the order of magnitude of the spin forces.

The degree of degeneracy of the states in any exclusive set $s = s'$ must equal the number of independent eigen-functions $F(\sigma_1, \sigma_2)$ that the variable s has, belonging to the eigen-value s'. This is just $2s' + 1$, as can be seen without detailed calculation from the fact that any of the Cartesian components of the total spin angular momentum, $\frac{1}{2}\Sigma\sigma_{rz}$ say, has the $2s' + 1$ eigen-values

$$s', \quad s' - 1, \quad s' - 2, \ldots \quad -s'.$$

When the spin forces are not neglected, each $(2s' + 1)$-fold degenerate energy level will in general be split up into $2s' + 1$ non-degenerate energy levels, differing slightly one from another. Thus the exclusive set of states $s = s'$ becomes a system of multiplets of multiplicity $2s' + 1$.

§ 5. *Determination of Energy Levels.*

We must now consider the application of perturbation theory for an approximate calculation of the energy levels. We shall take first the general case of a system with n similar particles, discussed in §§ 2 and 3. We shall follow the usual method in the theory of the perturbations of the stationary states of a degenerate system, according to which, if we label the states of the unperturbed

system α', α'', ..., we obtain the matrix $(\alpha' \mid V \mid \alpha'')$ representing the perturbation energy V and neglect all those matrix elements α', α'' for which the unperturbed states α' and α'' have two different energies. The remaining matrix elements will form a number of small matrices, one referring to each energy level of the unperturbed system, and having as the number of its rows and columns the number of independent states belonging to this energy level. The eigen-values of these matrices will then be, in the first approximation, the changes in the energy levels caused by the perturbation.

We suppose that for our unperturbed states each of the similar particles has its own "orbit," represented by a wave function $(q_r \mid \alpha)$ involving only the co-ordinates q_r of this one particle. We shall have altogether n orbits, one for each particle, which we assume for the present to be all different, and label $\alpha_1, \alpha_2, \ldots, \alpha_n$. The wave function representing an unperturbed state of the whole system will then be the product.

$$(q_1 \mid \alpha_1)\,(q_2 \mid \alpha_2) \ldots (q_n \mid \alpha_n) = (q \mid \alpha) \tag{18}$$

say, for brevity. If we apply an arbitrary permutation P_a to the α's, we shall obtain another wave function

$$(q_1 \mid \alpha_r)\,(q_2 \mid \alpha_s) \ldots (q_n \mid \alpha_t) = (q \mid P_a\alpha) \tag{19}$$

representing another unperturbed state with the same energy. There are thus altogether $n!$ unperturbed states with this energy, if we assume there are no other causes of degeneracy. The matrix elements of V that we must take into consideration are therefore of the type $(P_a\alpha \mid V \mid P_b\alpha)$, where P_a and P_b are two permutations of the α's, and form a matrix with $n!$ rows and columns. The eigen-values of this matrix are what we must calculate.

It is necessary in the present discussion to distinguish between the two kinds of permutations, those of the q's and those of the α's. The essential difference between them can perhaps be seen most clearly in the following way. Let us consider a permutation in the general case, say that consisting of the interchange of 2 and 3. This may be interpreted either as the interchange of the objects 2 and 3 or as the interchange of the objects in the places 2 and 3, these two operations producing in general quite different results. The first of these interpretations is the one we have been using throughout §§ 2 and 3, the objects concerned being the q's. A permutation with this interpretation can be applied to an arbitrary function of the q's. A permutation with the second interpretation has a meaning, however, when applied to a function of the q's only if each of the q's has a definite specifiable place in the function. This is not the case for a general function of the q's, but it is the case for any of

the $n!$ functions of the type (19), the place of each q being specified by the α with which it is bracketed. Any permutation applied to the q's in given places now produces the same result as the reciprocal permutation applied to the α's. A permutation *of* the q's (*i.e.*, one with the first interpretation) since it can be applied to any function of the q's, may be regarded as an ordinary dynamical variable. On the other hand a permutation of places or of the α's can be considered as a dynamical variable only in a very restricted sense, since it has a meaning only when multiplied into one of the $n!$ wave functions (19) or into some linear combination of them. We denote such a permutation of the α's, considered as a dynamical variable in this restricted sense, by a symbol P^a.

We can form algebraic functions of the variables P^a, which will be other variables in the same restricted sense. In particular we can form $\chi(P^a{}_a)$, the average of all P^a's similar to $P^a{}_a$. This must equal $\chi(P_a)$, the average of the similar permutations of the q's, since the total set of all permutations of a given type must evidently be the same whether the permutations are applied to the objects q or to the places α.

If we set up arbitrarily a one-one correspondence between the q's and the α's, as is done automatically when we label both the q's and the α's by the numbers 1, 2, 3, ..., n, as in (18), then, if we have any permutation of the q's, we can give a meaning to this same permutation of the α's. This meaning is such that

$$(q \mid \alpha) = (Pq \mid P\alpha).$$

In this equation we can apply a permutation P_a to the α's on both sides, which will give us

$$(q \mid P_a\alpha) = (Pq \mid P_aP\alpha) \tag{20}$$

an equation which shows us the connection between permutations of the q's and those of the α's when applied to the wave function (19).

The matrix $(P_a\alpha \mid V \mid P_b\alpha)$, which we must now study, may be obtained from the matrix $(q' \mid V \mid q'')$ representing the same dynamical variable V by a canonical transformation, in which the transformation functions are just $(q \mid P_a\alpha)$, the wave function (19), and its conjugate complex $(P_a\alpha \mid q)$, provided these functions are properly normalised. Thus

$$(P_a\alpha \mid V \mid P_b\alpha) = \iint (P_a\alpha \mid q')\, dq'\, (q' \mid V \mid q'')\, dq''\, (q'' \mid P_b\alpha). \tag{21}$$

Again, for arbitrary P,

$$(P_aP\alpha \mid V \mid P_bP\alpha) = \iint (P_aP\alpha \mid q')\, dq'\, (q' \mid V \mid q'')\, dq''\, (q'' \mid P_bP\alpha)$$

$$= \iint (P_aP\alpha \mid Pq')\, dq'\, (Pq' \mid V \mid Pq'')\, dq''\, (Pq'' \mid P_bP\alpha)$$

when we apply the permutation P to the variables of integration q' and q''. With the help of (20), this reduces to

$$(P_a P\alpha \mid V \mid P_b P\alpha) = \iint (P_a\alpha \mid q')\, dq'\, (Pq' \mid V \mid Pq'')\, dq''\, (q'' \mid P_b\alpha). \quad (22)$$

Now since V is symmetrical between all the particles, we must have

$$(q' \mid V \mid q'') = (Pq' \mid V \mid Pq''),$$

like (6), and hence, comparing (21) and (22), we obtain

$$(P_a\alpha \mid V \mid P_b\alpha) = (P_a P\alpha \mid V \mid P_b P\alpha) \quad (23)$$

Let $(P\alpha \mid V \mid \alpha) = V_P$ for brevity. Then, taking $P = P_b^{-1}$ in (23), we obtain

$$(P_a\alpha \mid V \mid P_b\alpha) = (P_a P_b^{-1}\alpha \mid V \mid \alpha) = V_{P_a P_b^{-1}}.$$

Thus the general matrix element $(P_a\alpha \mid V \mid P_b\alpha)$ depends only on the ratio $P_a P_b^{-1}$, and of the total of $(n!)^2$ matrix elements there are only $n!$ different ones. The coefficient of any V_P in this matrix will be a matrix, each of whose elements is 0 or 1, the 1 occurring when

$$(P_a\alpha \mid V \mid P_b\alpha) = V_P,$$

i.e., when $P_a P_b^{-1} = P$. But this matrix, multiplied into any wave function $(q \mid P_b\alpha)$, gives the result $(q \mid P_a\alpha)$ with $P_a P_b^{-1} = P$, *i.e.*, it gives the result $(q \mid PP_b\alpha)$, so that it is precisely the matrix representing the variable P^a, or the permutation P applied to the α's. Thus the whole matrix $(P_a\alpha \mid V \mid P_b\alpha)$ is equal to the matrix representing $\Sigma V_P P^a$, where the summation is over all the $n!$ permutations P, and we can put

$$V = \Sigma V_P P^a. \quad (24)$$

This formula shows that the perturbation energy V is equal to a linear function of the permutation variables P^a, with numerical coefficients V_P, which are the exchange energies. It is, of course, only an approximate formula, as it holds only with neglect of those matrix elements of V that refer to two different energy levels of the unperturbed system. It can, however, be used for the calculation of the energy levels in the first approximation, and is very convenient for this purpose as the expression $\Sigma V_P P^a$ is easily handled. This expression, it should be remembered, is a dynamical variable only in the restricted sense mentioned above, but this sense is just sufficiently general for equation (24) to be valid with neglect of those matrix elements of V referring to two different energy levels of the unperturbed system.

As an example of an application of (24) we shall determine the average energy of all those states arising from a given state of the unperturbed system

that belong to one exclusive set. This requires us to calculate the average eigen-value of V when the χ's have specified numerical values χ'. Now the average eigen-value of P^α equals that of $P^\alpha_a P^\alpha (P^\alpha_a)^{-1}$ for arbitrary P^α_a and thus equals that of $(n!)^{-1} \Sigma_a P^\alpha_a P^\alpha (P^\alpha_a)^{-1}$, which is $\chi'(P^\alpha)$ or $\chi'(P)$. Hence the average eigen-value of V is $\Sigma V_P \chi'(P)$. A similar method could be used for calculating the average eigen-value of any function of V, it being only necessary to replace each P^α by $\chi(P)$ to perform the averaging.

The modifications required in the theory when the orbits $\alpha_1, \alpha_2, \alpha_3, \ldots$, or the undisturbed system are not all different may easily be made. Suppose, for example, that α_1 and α_2 are the same. Then the permutation P^α_{12} that causes an interchange of α_1 and α_2 must equal unity. Only functions of the P^α's that commute with P^α_{12} now have a meaning. This, however, is sufficient for us to be able to follow out the same sort of argument as before, and obtain a result of the same form (24). The term in the summation in (24) that involves the permutation P^α_{12} now does not occur, since it could be added on to the term involving the identical permutation P^α_1. For the remaining terms, any two terms P^α_a and P^α_b must have the same coefficient if the permutations P^α_a and P^α_b can be obtained from one another by the interchange of α_1 and α_2. This results in $\Sigma V_P P^\alpha$ commuting with P^α_{12} and thus having a meaning. The condition $P^\alpha_{12} = 1$ will impose restrictions on the possible numerical values χ' that the χ's can have and will reduce the number of characters.

§ 6. *The Energy Levels in the Case of Electrons.*

We shall now consider the application of the formula (24) to the case of electrons. If we assume only Coulomb forces between the electrons, then the perturbation will consist of a number of terms, each involving the co-ordinates of one or at most two electrons, so that all the exchange energies V_P will vanish except those referring to the identical permutation P_1 and to simple interchanges of two orbits, P^α_{rs}. Thus (24) reduces to

$$V = V_1 + \Sigma_{r<s} V_{rs} P^\alpha_{rs},$$

V_{rs} being the exchange energy of orbits r and s. Since the P^α's have exactly the same properties as the P's, we can replace the P^α's in this expression for V by P's without changing its eigen-values. This gives us

$$V = V_1 + \Sigma_{r<s} V_{rs} P_{rs}, \tag{25}$$

where the $=$ sign is now to be interpreted as denoting the equality of the eigen-values of the two sides and not the complete equality of the two sides as dynamical variables or operators.

With the help of (16), the result (25) may be put in the more expressive form

$$V = V_1 - \tfrac{1}{2}\Sigma_{r<s} V_{rs}\{1 + (\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_s)\}. \quad (26)$$

This shows that, for the purpose of calculating energies, the exchange interaction due to the equivalence of the electrons may be replaced by a constant perturbation energy $-\frac{1}{2}\Sigma_{r<s} V_{rs}$, together with a coupling between the spin vectors with energy $-\frac{1}{2}V_{rs}(\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_s)$ for each pair of electrons r, s. It is this coupling which may be considered as giving rise, for instance, to the large differences in energy between the singlet and triplet terms of helium. The total number of eigen-values of the right-hand side of (26) is 2^n, a factor 2 occurring for the representation of the spin vector of each of the n electrons. These 2^n eigen-values will not, in general, all be different, as each one will occur repeated a number of times to give the correct multiplicity of the corresponding term.

When two of the orbits of the unperturbed system are the same, say the orbits 1 and 2 are the same, the only eigen-values of the right-hand side of (25) or (26) that will be eigen-values of V are those consistent with the equation $P_{12} = 1$ or $P^{\sigma}_{12} = -1$. In this case we have $V_{12} = 0$ and

$$V_{1r} = V_{2r} \quad (r = 3, 4, 5, \ldots),$$

which results in the right-hand side of (26) being symmetrical between $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$. It follows from this that any eigen-function $F(\sigma_{1z}, \sigma_{2z}, \sigma_{3z}, \ldots)$ of this right-hand side, considered as an operator, must be either symmetrical or antisymmetrical between σ_{1z} and σ_{2z}. The condition $P^{\sigma}_{12} = -1$ now shows that only the antisymmetrical ones, representing states for which the spins $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ are antiparallel, must be taken into account. The number of eigen values of (26) that must be used is thus reduced by a factor 4, on account of there being only one antisymmetrical eigen-function for every three symmetrical ones. The case of more than two orbits the same cannot occur with electrons.

In our theory of the energies we have nowhere had to assume that the wave functions $(q \mid \alpha_1)$, $(q \mid \alpha_2)$, ..., representing the various orbits in the unperturbed system, are orthogonal, or that they are eigen-functions of any unperturbed Hamiltonian H_0. This enables an important generalisation to be made in the application of our results. It is not necessary that we should be able to split up our Hamiltonian for the whole system H into a Hamiltonian for the unperturbed system H_0 and a perturbation energy V, and then use the eigen-functions of H_0 to give our $(q \mid \alpha_1)$, $(q \mid \alpha_2)$, ... We can take our $(q \mid \alpha_1)$,

$(q \mid \alpha_2)$ to be *any* functions giving a good approximation to the actual distribution of electrons in the system, and must then throughout the analysis replace V, which now no longer exists, by the whole Hamiltonian H. The wave functions supplied by Hartree's theory can thus very conveniently be used. The only mathematical conditions which the $(q \mid \alpha_1)$, $(q \mid \alpha_2)$, ..., need satisfy is that in the matrix $(\alpha' \mid \mathrm{H} \mid \alpha'')$ representing H, those matrix elements for which the α'''s are not simply a permutation of the α''s must be small.

As an example of the application of (25), Heitler's formula* for the interaction of two atoms A and B, each with n valency electrons, will be deduced. The fundamental theorem of multiplet structure shows that there will be a quantum number s describing the magnitude of the resultant spin of the electrons in both atoms. This same theorem shows that, provided the interaction between the atoms is small compared with the exchange energies within either of them, the whole energy of the system will depend very largely on the magnitudes s_A and s_B of the resultant spins for the two atoms separately, so that for each energy level of the whole system there must be definite numerical values for s_A and s_B, which will thus be two more quantum numbers describing states of the whole system. For valency electrons the resultant spin vector has its maximum possible value (we can if we like in this case speak of the electron spins all being parallel) and hence

$$s_A = s_B = \tfrac{1}{2}n. \tag{27}$$

Again, if ζ is the valency of the homopolar bond uniting the two atoms (*i.e.*, $\zeta = 1, 2, \ldots$, for a single, double, ... bond)

$$s = n - \zeta. \tag{28}$$

We now apply our formula (25), taking the summation only over pairs of orbits r, s of which one is in each atom, since we want the interaction energy between the two atoms. This gives

$$\mathrm{V} = \mathrm{V}_1 + \Sigma_{\mathrm{AB}}\, \mathrm{V}_{rs}\, \mathrm{P}_{rs}.$$

As a rough approximation we may take all the exchange energies V_{rs} between two orbits one in each atom to be equal. Calling these exchange energies V_Q, we get

$$\mathrm{V} = \mathrm{V}_1 + \mathrm{V}_Q\, \Sigma_{\mathrm{AB}}\, \mathrm{P}_{rs}.$$

We must now evaluate $\Sigma_{\mathrm{AB}}\, \mathrm{P}_{rs}$, summed over all pairs one in each atom, which we can best do by first summing over all possible pairs and then subtracting

* W. Heitler, 'Z. Physik,' vol. 47, p. 835 (1928), equation (33).

the two sums for pairs both in atom A and both in atom B respectively. Thus

$$\begin{aligned}\Sigma_{AB} P_{rs} &= \Sigma P_{rs} - \Sigma_{AA} P_{rs} - \Sigma_{BB} P_{rs} \\ &= -\tfrac{1}{4}\{2n(2n-4) + 4s(s+1)\} + \tfrac{1}{4}\{n(n-4) + 4s_A(s_A+1)\} \\ &\qquad + \tfrac{1}{4}\{n(n-4) + 4s_B(s_B+1)\}\end{aligned}$$

by a three-fold application of (17), in the first case to a system of $2n$ electrons. This reduces to

$$\begin{aligned}\Sigma_{AB} P'_{rs} &= -\tfrac{1}{4}\{2n^2 + 4s(s+1) - 4s_A(s_A+1) - 4s_B(s_B+1)\} \\ &= \zeta - (n-\zeta)^2.\end{aligned}$$

Thus

$$V = V_1 + V_Q\{\zeta - (n-\zeta)^2\},$$

which is Heitler's result.
