THE APPLICATION OF GROUP THEORY TO THE QUANTUM DYNAMICS OF MONATOMIC SYSTEMS

By Carl Eckart
University of Chicago

Table of Contents

I. General Theory

II. Specification of the Groups to be Considered

III. Construction of Irreducible Representations of the Rotation Group, etc.

IV. Application to the Wave Equation

THIS report is intended to furnish an account of so much of the theory of groups and their representations as has been found useful in the solution of quantum dynamical problems. The general theory (Part I) is general only in comparison to the applications which follow. This is evident in the treatment of Schur's lemma, which is proven for a very special case, mathematically speaking, but with quite sufficient generality for the purposes of quantum dynamics.

The greater portion of the report is taken up by applications to the theory of complex atomic spectra. It has seemed better to limit the field of application and to use the available space for a fairly detailed account of some problems, rather than to draw a few general illustrations from each of several fields. This may lead the reader to form a false estimate of the applicability of the methods; it may merely be mentioned that they have also been applied to the theories of molecular spectra, valence forces, electron conduction, and ferromagnetism.

I. General Theory

1. **The wave equation and the manifold of its solutions.** The most important computational task set by quantum dynamics is the solution of the characteristic value problem

$$H\psi = W\psi. \tag{1}$$

Here ψ is a function of the translatory and spin coordinates of the electrons: these will be denoted by $x = (x_1 x_2 \cdots s_1 s_2 \cdots)$; H is a linear operator, and W a number independent of x. Certain other conditions (the "natural" boundary conditions) are imposed on ψ; those solutions of Eq. (1) which satisfy these are called the characteristic solutions of the equation, and those values of W for which at least one such solution exists are called the characteristic values or energy levels. This problem has been discussed in its relationship to the physical principles of the quantum theory elsewhere in this journal;[1] in this review it is proposed to discuss the problem

[1] E. C. Kemble, Phys. Rev. Supplement **1**, 157 (1929).

on its own merits, with a view to summarizing certain mathematical methods which are capable of yielding exact information about its solution (though unfortunately they do not suffice for its complete solution). These same methods prove to be of even greater value when combined with the perturbation theory.

To present these methods in as simple a manner as possible, it has seemed desirable to treat them quite independently of the general quantum theory. It is presumed that the reader is familiar with this theory, however. In distinction to the point of view of quantum dynamics, attention will be focused on those characteristic solutions of Eq. (1) for which W has a single definite value. Solutions for different values of W will occasionally be considered, but only after suitable announcement of the fact. The linearity of the operator H has as its consequence that, if $\psi=u$ and $\psi=v$ are solutions of Eq. (1) for the given value of W, so is $\psi=au+bv$ whenever a and b are constants independent of the coordinates.[2] If the boundary conditions are homogeneous, $au+bv$ will also be a characteristic solution if u and v are such; it is assumed that this is the case. These statements may conveniently be summarized by saying that the solutions of Eq. (1) form a linear manifold or space; the last term is justified by the fact that the functions of the linear manifold satisfy the same abstract axioms as the vectors of a space.[3]

In particular, there is an integer ν, called the number of dimensions of the manifold, determined by the statement that any $\nu+1$ members of the manifold are linearly dependent, while it is always possible to find ν members which are linearly independent. Any ν independent members, $u_1u_2\cdots u_\nu$, are called a basis for the manifold, and any other member ψ can be expressed as[4]

$$\psi = a_n u_n \,, \tag{2}$$

where the constants a_n are determinate. The basis corresponds to the set of "unit" vectors along an oblique set of axes in a space, the a_n to the coordinates measured along these axes with the appropriate units. The number ν is also called the degree of degeneracy and the statistical weight of the energy level; the case $\nu=1$ is of course not excluded, and it is assumed that ν is finite.

It is possible to define a unit of length in this space: it is known that in case the wave equation does not involve the spin coordinates, integrals of the type[5]

[2] An operator R is said to be linear if

$$R(f+g) = Rf + Rg,$$
$$R(af) = a(Rf),$$

where f and g are arbitrary functions on which it is able to act and a is a constant.

[3] H. Weyl; Gruppen theorie und Quantenmechanik, Chap. I (Leipzig, 1928). This book will be referred to as Weyl, *l.c.*

[4] The usual summation convention is adopted.

[5] The asterisk indicates that g^* is the conjugate complex function to g.

$$\int f(x)g^*(x)d\tau, \quad (d\tau = \rho(x)dx_1dx_2 \cdots) \tag{3}$$

play an important part in the theory of its solution; the integral is to be extended over all values of the variables x which have a physical significance. If the spin variables s_i (each of which is capable of assuming only the values $+\frac{1}{2}$ or $-\frac{1}{2}$) do enter into Eq. (1), the role of these integrals is played by

$$\sum_{s_1 s_2} \cdots \int f(x_i s_i)g^*(x_i s_i)d\tau. \tag{3a}$$

To include these two cases it is convenient to represent both by the same notation

$$(f, g). \tag{4}$$

It will be called the scalar product of the functions f and g. The following properties of the scalar product will be verified:

$$\begin{aligned} (f, g) &= (g^*, f^*) = \text{constant}; \\ (af, g) &= a(f, g), \text{ if } a \text{ is a constant}; \\ (f + f', g) &= (f, g) + (f'g); \\ (f, f) &> 0, \text{ unless } f \equiv 0, \text{ in which case it vanishes.} \end{aligned} \tag{5}$$

Because of the last of the properties the scalar product is suitable for a definition of length in the manifold of W: if $(u,u) = c^2$, then $c \geqq 0$ may be called the length of u. Furthermore, if

$$(f, g) = 0$$

then f and g are said to be orthogonal.

One other property of the inner product deserves special mention: if $f = f_1(x_1 \cdots x_n s_1 \cdots s_n) f_2(x_{n+1} \cdots x_n s_{n+1} \cdots s_n)$ and g also has the same form, then

$$(f, g) = (f_1, g_1)_1 (f_2, g_2)_2$$

where( )$_1$ and ( )$_2$ are operations analogous to ( ) but involving only the coordinates $x_1 \cdots x_n s_1 \cdots s_n$ and $x_{n+1} \cdots x_n s_{n+1} \cdots s_n$ respectively. This formula is analogous to breaking up an N-fold integral into the product of an n-fold and $(N\text{-}n)$-fold integral when the integrand can be factored as above.

For the purposes of this review, H may be restricted to the class of hermitian operators, which satisfy the relation

$$(Hf, g) = (f, Hg).$$

This formula indicates that the operator H is such that an operation analogous to an integration by parts can be carried out on the scalar product of Hf and g.

It is evident that the ν functions $u_1 u_2 \cdots u_\nu$ may always be chosen orthogonal to one another, and of unit length. As it will be necessary to use this

fact from time to time, a proof may be given here. Suppose that $u_1 u_2 \cdots u_\mu$, $\mu < \nu$, are orthogonal to one another, and each have the length 1, while v is any other member of the manifold which is linearly independent of $u_1 \cdots u_\mu$. If $u_{\mu+1}$ can then be determined so that it is orthogonal to $u_1 \cdots u_\mu$ and is of unit length, the theorem has, to all intents and purposes, been proven. But this is true, for

$$u' = v - u_1(v, u_1) - u_2(v, u_2) - \cdots - u_\mu(v, u_\mu)$$

is orthogonal to $u_1 \cdots u_\mu$ and cannot vanish, since v is linearly independent of the other functions. If its length is c:

$$(u', u') = c^2,$$

then $u_{\mu+1} = u'/c$ satisfies the required conditions. While this procedure can always be used to construct an orthogonal basis for any given manifold, it is a lengthy computational process, and to be avoided if possible. The basis u_n will therefore not be assumed to be orthogonal unless that fact is explicitly stated.

The manifold of W, hereafter called $\mathfrak{M}_w$, contains other linear manifolds, just as a three dimensional space contains straight lines and planes (one and two dimensional linear manifolds). These may be defined by means of their bases: let $v_1\, v_2 \cdots v_\mu$ be $\mu < \nu$ linearly independent functions of $\mathfrak{M}_w$. These generate the submanifold $\mathfrak{M}_1$, whose functions are

$$\psi = \sum_{\alpha=1}^{\mu} a_\alpha v_\alpha$$

and which is μ dimensional.

If one linear submanifold $\mathfrak{M}_1$, of μ dimensions be given, it is always possible to find another, $\mathfrak{M}_2$, of $\mu' \leqq \nu - \mu$ dimensions such that

$$(\psi, \psi') = 0$$

if ψ is in the one and ψ' in the other. The two manifolds are said to be orthogonal. To prove the linearity of $\mathfrak{M}_2$, note that if (ψ, ψ') and (ψ, ψ'') are both zero, $(\psi, a\psi' + b\psi'')$ is also zero. The bases of two orthogonal manifolds must be linearly independent, so that $\mu + \mu' \leqq \nu$. The actual construction may be carried out by a modification of the process of orthogonalization.

If $\mu + \mu' = \nu$, $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are said to be complementary. In this case, any member ψ of $\mathfrak{M}_w$, may be expressed as the sum of two functions ψ_1 and ψ_2 which are contained in $\mathfrak{M}_1$ and $\mathfrak{M}_2$ respectively. They are called the components of ψ in the two complementary submanifolds.

2. **The integrals of the wave equation and the finite matrices representing them.** Quantum dynamics also contemplates linear operators, other than H, which are capable of acting on functions of the coordinates. They are called dynamical variables, but need not necessarily possess analogues among

the variables of classical dynamics. Let R be one of these; if it commutes[6] with H, i.e., if

$$HR = RH \ , \tag{7}$$

it is said to be an integral of the Eq. (1). Integrals possess the following property: if u is a solution of Eq. (1) so is $\psi = Ru$; for

$$H\psi = RHu = HRu = R(Wu) = W\psi.$$

This does not prove that if u is a *characteristic* solution of Eq. (1), Ru will also be characteristic. This is usually the case, however, and may be included in the definition of the term "integral" as a separate requirement. The product of two integrals is obviously again an integral.

Let R, S, $T \cdots$ be integrals of Eq. (1) and $u_1 \cdots u_\nu$ a basis of the manifold of W. Then, because of this property of integrals and the properties of the basis,

$$Ru_n = R_{nm}u_m, \ Su_n = S_{nm}u_m, \text{ etc.}, \tag{8}$$

where the coefficients R_{nm}, S_{nm}, are numbers independent of the coordinates. They may be arranged in matrix array, the left hand index numbering the rows, and the right hand index the columns. Furthermore,

$$RSu_n = S(R_{nm}u_m) = R_{nm}S_{mk}u_k, \tag{9}$$

so that the matrix associated to the product of two operators is itself the product of the matrices associated with the factors of the operator product. *The order of factors in the two products is the same.* Whenever a set of matrices is associated to a set of operators in this manner, they are said to be multiplicatively associated. If, in addition, one and only one matrix is associated to each operator then the set of matrices is called a representation of the set of operators.[7] Both conditions are obviously fulfilled by the matrices (8); the functions u_n are said to be the basis of the representation.

It is not possible to construct a representation of more general dynamical variables using u_n as basis. It is true that it is possible to associate matrices of ν rows and columns to any dynamical variable (by taking the "time average" of the infinite matrix associated to it by quantum theory) but these do not satisfy the requirement that the matrix product be associated to the operator product. The following discussion will therefore apply directly only to integrals.

The definition of the matrices by Eq. (8) is obviously not independent of

[6] If R and S are two operators, their product $T=RS$ is defined by $Tf=g$, where g is determined from

$$Rf=g_1$$
$$Sg_1=g.$$

The factors of the product act successively in the order of reading from *left to right.* $RS \neq SR$ in general, but it is assumed that $Q(RS)=(QR)S$.

[7] The converse condition, that only one operator be associated to each matrix, need not necessarily be fulfilled. This definition is implied whenever the term representation is used in the following discussion.

the choice of basis. Let $v_1 \cdots v_\nu$ be another basis for the manifold; then, because of the properties of a basis

$$v_\alpha = a_{\alpha n} u_n, \tag{12}$$

and the determinant of the matrix $a_{\alpha n}$ does not vanish. If

$$Rv_\alpha = R'_{\alpha\beta} v_\beta, \tag{13}$$

then

$$R'_{\alpha\beta}\, a_{\beta n} = a_{\alpha m} R_{mn}, \tag{14}$$

as follows at once on substituting Eq. (12) into (13) and using Eq. (8). Since the determinant of $a_{\alpha n}$ does not vanish, its reciprocal exists and Eq. (14) may also be written

$$R'_{\alpha\beta} = a_{\alpha n} R_{nm} a^{-1}{}_{m\beta}. \tag{14a}$$

Representations which are related in this manner are said to be equivalent, and to arise from one another by a canonic transformation with the matrix $a_{\alpha n}$. The accented matrices also constitute a representation of the operators. It is to be emphasized that the matrix $a_{\alpha n}$ is the same for every matrix $R_{mn}, S_{mn}, \cdots$ of the representation. The concept of equivalent representations is fundamental in the following discussion, where equivalent representations are not considered to differ essentially from one another.

3. **Further restriction of the operators to be considered. The group postulates.** In the foregoing paragraph it has been remarked that only integrals of Eq. (1) will be considered in this review; further restrictions are now to be placed on the class of integrals which form the primary objects of study. This will somewhat restrict the generality of the results, but in return for this, the theorems will be much more numerous and specific in their content.

The first restriction is the assumption that the integrals under consideration[8] form a group. Any set of elements is said to be a group[9] if the following three postulates are obeyed:

I. If R and S are members of the group, their products, RS and SR, have a well-defined significance and are also members of the set. The associative law $R(ST) = (RS)T$ is obeyed by the multiplication.

II. An element I, called "the identity" or "unity" and having the properties $RI = IR = R$ for any R, is a member of the group.

III. Each element R uniquely determines another element of the group, called its reciprocal R^{-1} and defined by $R^{-1}R = RR^{-1} = I$.

The first of these postulates is satisfied by the set of all integrals of Eq. (1). The element I may be defined by the relation

[8] It is not assumed that *all* integrals of Eq. (1) satisfy the group postulates, but only that some do. The restriction imposed on the form of Eq. (1) is thus very slight.

[9] H. Weyl, *l.c.* p. 97.

$$If = f; \tag{15}$$

it is a (trivial) integral of Eq. (1).

The third group postulate is more stringent: it excludes all integrals which have no reciprocal and all which have more than one. (The operator $\partial/\partial x$ cannot be a member of any group, since its reciprocal is not unique.)

The theory of abstract groups does not specify the analytic character of the elements of the group: they may be operators, matrices, or ordinary numbers. For the moment, operator groups are under consideration, but it may be remarked that the matrices representing a group of operators often satisfy the group postulates also.

The second restriction to be placed on the integrals is the assumption that they are unitary: i.e., satisfy the relation

$$(R^{-1}f, g) = (f, Rg), \tag{10a}$$

which may also be written

$$(f, g) = (Rf, Rg). \tag{10b}$$

The justification for this name is obtained by noting that f and Rf have the same length: R leaves the unit of length unchanged. The unitary operators are thus analogous to rotations of the vectors of a space. This may be brought out more clearly by considering the matrices representing unitary operators: let u_n be a normallized orthogonal basis for $\mathfrak{M}_w$, so that

$$(u_n, u_m) = \delta_{nm}$$

and

$$Ru_n = R_{nm}u_m,$$
$$R^{-1}u_n = R^{-1}{}_{nm}u_m;$$

then

$$\begin{aligned} R_{nm}{}^{-1} &= (R^{-1}u_n, u_m) \\ &= (u_n, Ru_m), \text{ by (10b)}; \\ &= R^*{}_{mn} \text{ by (5)}. \end{aligned}$$

Hence

$$\delta_{kn} = R_{kn}R^{-1}{}_{nm} = R_{kn}R_{mn}{}^*; \tag{10c}$$

except for the asterisk, these relations are the same as those satisfied by the coefficients of an orthogonal substitution (rotation) of coordinates. A matrix whose elements satisfy Eq. (10c) will be called a unitary matrix. It is to be noted that the matrices representing a set of unitary operators are themselves unitary only when the basis is normallized and orthogonal. According to the definition of the Eq. (14), the representation will always be equivalent to one of unitary matrices; this theorem will be used frequently in the following.

Those integrals of the wave equation which are unitary and form a group will hereafter be called the group $\mathfrak{G}$ of the wave equation, to distinguish them from other integrals and other groups of operators.

4. **Invariant manifolds.** Let ψ be a member of any linear manifold $\mathfrak{M}$ and $I, R, S, T \cdots$ the members of the group $\mathfrak{G}$; if $R\psi$, $S\psi$, $T\psi \cdots$ are also members of $\mathfrak{M}$, it is said to be invariant under the group $\mathfrak{G}$. The manifold $\mathfrak{M}_w$ is invariant under $\mathfrak{G}$; it often happens that a submanifold of $\mathfrak{M}_w$, say $\mathfrak{M}_1$ is also invariant under $\mathfrak{G}$.[10] If this is the case, $\mathfrak{M}_w$ is said to be reducible.

The complementary manifold to $\mathfrak{M}_1$ is also invariant under $\mathfrak{G}$. Let ψ be any member of $\mathfrak{M}_1$, and ψ' of $\mathfrak{M}_2$, the complementary manifold; then by definition

$$(\psi, \psi') = 0.$$

By Eq. (10a), if R is any member of $\mathfrak{G}$,

$$(R^{-1}\psi, \psi') = (\psi, R\psi').$$

But, because of the group postulate III, and the invariance of $\mathfrak{M}_1$, $R^{-1}\psi$ is also in $\mathfrak{M}_1$. Hence the left side of this equation vanishes, and therefore $R\psi'$ is in $\mathfrak{M}_2$, which was to be proven. When the manifolds $\mathfrak{M}_1$ and $\mathfrak{M}_2$ have both been determined, $\mathfrak{M}_w$ is said to have been reduced.

$\mathfrak{M}_1$ and $\mathfrak{M}_2$ may themselves be reducible. If this is the case, the process of reduction may be applied to each, and then to their components in turn, until finally $\mathfrak{M}_w$ has been completely reduced into irreducible submanifolds, $\mathfrak{M}_1$, $\mathfrak{M}_2$, $\mathfrak{M}_3, \cdots \mathfrak{M}_q$.

The manifolds $\mathfrak{M}_i$ will each be of a certain number of dimensions, say ν_i, where $\Sigma\nu_i = \nu$. (The possibility $\nu_i = 1$ is, of course, not excluded.) A basis of functions say u_{in}, $(n = 1, 2, \cdots, \nu_i)$ can be found for each; since these functions are independent by definition, and are ν in number when i varies from 1 to q, it follows that they constitute a basis for the whole of $\mathfrak{M}_w$.

The importance of the invariant submanifolds lies in the fact that the whole of the reasoning of Section 2 can be repeated for each of these (in place of $\mathfrak{M}_w$) *in-so-far as only integrals of the group* $\mathfrak{G}$ *are considered.* Each $\mathfrak{M}_i$ is the basis of a representation[11] of the group $\mathfrak{G}$:

$$Ru_{in} = {}^{i}R_{nm}u_{im}; \tag{16}$$

representations which are based on irreducible manifolds are themselves called irreducible.

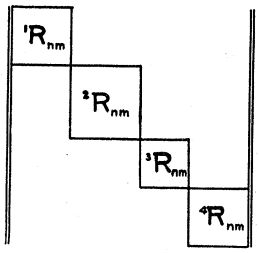
The q matrices (16) may also be considered as a single matrix of ν rows and columns, each row or column being numbered with two indices i n, in the order $11, 12, \cdots, 1\nu_1; 21, 22, \cdots, 2\nu_2; 31 \cdots, q\nu_q$. Analytically, the matrices defined by Eq. (16) have the form

$$R_{in,jm} = {}^{i}R_{nm}\delta_{ij}; \tag{17}$$

[10] Example: Three dimensional Euclidean space is invariant under the group of rotations about the z-axis, the xy plane and the z-axis are invariant submanifolds. They are complementary and irreducible.

[11] Matrices belonging to different representations of the group $\mathfrak{G}$ will be distinguished by an index at the upper left of the letter; *such indices are never summed unless it is explicitly indicated.*

when arranged in matrix array, they have the form:



(17a)

in which all elements not in the squares along the principal diagonal are zero. Every matrix representing an element of the group $\mathfrak{G}$ has this form. As a particular case, all the $\mathfrak{M}_i$ may be one dimensional, in which case every matrix ${}^iR_{nm}$ is an ordinary number, and (17a) is a diagonal matrix.

Let $w_1 w_2 \cdots w_\nu$ be any other basis of $\mathfrak{M}_w$ and

$$Rw_\alpha = R_{\alpha\beta} w_\beta. \tag{18}$$

The representation $R_{\alpha\beta}$ must be equivalent to the representation (17); this is the significance of the term "reducible" as applied to representations. The set $\mathfrak{M}_w$ is usually given in terms of a basis like w_α and the first problem which presents itself is the calculation of the canonic transformation which transforms it into u_{in}:

$$u_{in} = a_{in,\alpha} w_\alpha. \tag{19}$$

This problem will be solved with considerable generality in the following pages.

It is convenient to introduce another convention into the definition of complete reduction: two or more of the smaller matrices ${}^iR_{nm}$ may be equivalent. If this is the case, they may always be made identical by a canonic transformation, and this will be assumed to have been accomplished in the process of complete reduction. It is convenient to add a third index to the bases u_{in}, which indicates to which one of a number of identical representations the function belongs. The basis is then designated by u_{ipn}, and

$$Ru_{ipn} = {}^iR_{nm} u_{ipm},$$

where p has any value from 1 to q_i (q_i being the number of times the representation ${}^iR_{nm}$ appears in (17a)). The Eq. (19) then becomes

$$u_{ipn} = a_{ipn,\alpha} w_\alpha. \tag{19a}$$

5. **Schur's lemmas.** As a preliminary step toward the solution of this problem, certain criteria for the reducibility of a representation will be deduced. The definition of reducibility immediately results in one: let R_{nm} be a reducible representation of $\mathfrak{G}$, whose basis is $u_1 \cdots u_\nu$:

$$Ru_n = R_{nm} u_m. \tag{20}$$

Since the manifold $\mathfrak{M}$ subtended by $u_1 \cdots u_\nu$ is reducible, there exist $\mu<\nu$ functions of the form

$$v_\alpha = a_{\alpha n}u_n, \quad (\alpha = 1, 2 \cdots \mu; n = 1, 2, \cdots \nu;) \tag{21}$$

such that

$$Rv_\alpha = R'_{\alpha\beta}v_\beta. \tag{22}$$

On substituting Eqs. (21) and (20) into Eq. (22) and using the independence of the u_n, it follows that

$$R'_{\alpha\beta}a_{\beta m} = a_{\alpha n}R_{nm}. \tag{23}$$

This is the necessary condition that R_{nm} be reducible; it is formally identical with Eq. (14), but it is to be noted that $a_{\alpha n}$ in Eq. (23) is a rectangular matrix of $\mu<\nu$ rows and ν columns, while in Eq. (14), it is a square matrix of ν rows and columns. It is again to be emphasized that this equation is valid for every element R of the group and that $a_{\alpha n}$ is the same matrix for each R.

The Eq. (23) is also the sufficient condition for the reducibility of R_{nm} (if the trivial matrix $a_{\alpha n}$ all of whose elements are zero be excluded) since if a non-trivial matrix exists satisfying Eq. (23), it may always be used to define v_α in Eq. (21), which then satisfy Eq. (22) because of Eq. (20).

Lemma I.

If R_{nm} be an irreducible representation of a group $\mathfrak{G}$ of unitary elements and $R'_{\alpha\beta}$ any other representation of $\mathfrak{G}$ having fewer rows and columns than R_{nm}, then no matrix $a_{\alpha n}$ exists and satisfies Eq. (23) for all R, other than $a_{\alpha n}\equiv 0$.

A second extremely useful theorem is the following:

Lemma II.

If R_{nm} be an irreducible representation of $\mathfrak{G}$ and $R'_{\alpha\beta}$ another representation of $\mathfrak{G}$ having the same number of rows and columns as R_{nm}, while $a_{\alpha n}$ is a square matrix satisfying Eq. (14), then either (1). The determinant of $a_{\alpha n}$ does not vanish and the two representations are equivalent; or (2). Every element of $a_{\alpha n}$ is zero.

The theorem is a matter of definition if the determinant of $a_{\alpha n}$ does not vanish, therefore only the second case need be considered. A matrix $b_{\alpha\beta}$ may be found in any case such that it transforms $a_{\alpha n}$ to diagonal form:

$$b_{\alpha\beta}a_{\beta n}b_{nm}^{-1} = a_\alpha\delta_{\alpha m}. \tag{24}$$

Since the determinant of a matrix is invariant under a canonic transformation, at least one of the a_α must vanish. It is no loss of generality to assume $a_1=a_2=\cdots a_\mu=0$, $a_{\mu+1}\neq 0$, $a_{\mu+2}\neq 0$, $\cdots$, $a_\nu\neq 0$. Substituting in Eq. (14) and rearranging:

$$b_{\gamma\alpha}R'_{\alpha\beta}b_{\beta(k)}^{-1}a_{(k)} = a_{(\gamma)}b_{(\gamma)n}R_{nm}b_{mk}^{-1} \tag{25a}$$

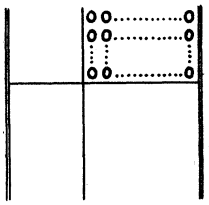
which is not to be summed over γ and k. Calling

$$b_{\gamma n}R_{nm}b_{mk}^{-1} = {}^1R_{\gamma k}, \quad b_{\gamma\alpha}R'_{\alpha\beta}b_{\beta k} = {}^2R_{\gamma k},$$

(these matrices are also representations of the group $\mathfrak{G}$) Eq. (25a) becomes

$$ {}^1R_{\gamma(k)}a_{(k)} = a_{(\gamma)}{}^2R_{(\gamma)k}. \tag{26} $$

This being valid for every R, it follows that ${}^1R_{\gamma k}$ has the form



(27)

since ${}^1R_{\gamma k}=0$ whenever $\gamma \leqq \mu$, $k>\mu$, because of Eq. (26). Now let R_{mn} be based on the functions u_n which, by hypothesis, subtend an irreducible manifold $\mathfrak{M}$. Then the functions

$$ v_\alpha = b_{\alpha n}u_n $$

satisfy the relation

$$ Rv_\gamma = {}^1R_{\gamma\beta}v_\beta $$

and hence, because of (27), $v_1 \cdots v_\mu$ subtend an invariant manifold, which is contrary to hypothesis if $\mu<\nu$. Hence every a_α must be zero, and therefore every $a_{\alpha n}$, which was to be proven.

The properties of irreducible representations were first studied by Frobenius and Burnside.[12] The above proof of the second lemma is based on a more general proof given by I. Schur, and the two theorems may be called Schur's Lemmas, though the theorem proven by Schur is somewhat different and more general than either.[13]

As a corollary to the second lemma, it may be shown that

If R_{mn} is an irreducible representation of $\mathfrak{G}$ and a_{nm} commutes with all R_{mn}:

$$ R_{mn}a_{nk} = a_{mn}R_{nk}, $$

then $a_{nm}=a\delta_{nm}$, where a is independent of n.

A matrix $b_{\alpha\beta}$ exists satisfying Eq. (25), but it is now no longer necessary that one of the a_α vanish. In Eq. (26), ${}^1R_{\gamma k}={}^2R_{\gamma k}$, so that it becomes

$$ (a_k - a_\gamma){}^1R_{\gamma k} = 0 $$

[12] G. Frobenius, Sitzungsberichte der Berliner Akademie, 1896–99. Burnside, Theory of Groups of Finite Order, Camb. (1911).

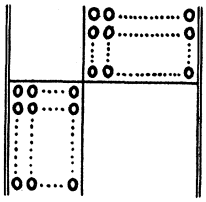
[13] I. Schur, Sitzungsberichte der Berliner Akademie, XVIII (1905). The theorem proven in this paper reads:

Let R_{mn} and $R'_{\alpha\beta}$ be any two irreducible representations of the same group (not necessarily of unitary operators), of ν and ν' rows and columns respectively. If $a_{\alpha n}$ is a matrix satisfying

$$ R'_{\alpha\beta}a_{\beta m}=a_{\alpha n}R_{nm} $$

for all R, then either $a_{\alpha n}=0$, or $\nu=\nu'$ and the determinant of $a_{\alpha n}$ does not vanish.

which is true for all R. Hence if all the a_k are not equal, all ${}^1R_{\gamma k}$ must have the form



which is contrary to the assumption of the theorem; but if all the a_k are equal, then the original matrix must also have the form $a_{nm} = a\delta_{nm}$.

6. **An application to perturbation theory.** The solution of the problem stated at the end of Section 4 may profitably be interrupted here to show the relation of the concept of irreducible manifolds to the problem of solving the wave equation by perturbation methods. Whenever the operator H can be written in the form

$$H = H^0 + \epsilon H',$$

where ϵ is a numerical parameter of given value, both the solutions and characteristic values of the equation

$$(H^0 + \epsilon H')\Psi = W\Psi \tag{28}$$

may be expanded in a power series in ϵ:

$$\begin{aligned} \Psi &= \psi + \epsilon\psi' + \epsilon^2\psi'' + \cdots, \\ W &= W + \epsilon W' + \epsilon^2 W'' + \cdots. \end{aligned} \tag{29}$$

Even when ϵ is not small, the first one or two terms of these series are often fair approximations to their sums.

The successive members of these series satisfy the equations

$$H^0\psi - W\psi = 0, \tag{30}$$

$$H^0\psi' - W\psi' = (W' - H')\psi, \tag{31}$$

$$H^0\psi'' - W\psi'' = (W' - H'\,\psi' + W''\psi, \tag{32}$$

. . . . . .

These may be solved successively for ψ, ψ', ψ'', etc.; each of these functions must itself satisfy the boundary conditions of the problem. The characteristic values and functions of Eq. (30) are assumed known: let them be W_k and $v_{k\alpha}(\alpha = 1 \cdots \nu)$ respectively. The functions $v_{k\alpha}$ subtend the manifold of the energy level W_k. For the following it will be necessary to adopt the general view-point of quantum theory and consider all possible characteristic values W_k and their functions. The $v_{k\alpha}$ may be supposed to constitute a complete and normalized orthogonal set, so that

$$H'v_{k\alpha} = \sum_l \sum_\rho H'_{k\alpha,l\beta}v_{l\beta}, \tag{33}$$

where

$$H'_{k\alpha,l\beta} = (H'v_{k\alpha}, v_{l\beta}).$$

This matrix contains an infinite number of rows and columns, and possibly the indices k and l may take on all values in a certain continuous range also; this possibility will not be considered.

Now, if

$$\psi = \sum_{\alpha} A_\alpha v_{k\alpha} \tag{34}$$

it will satisfy Eq. (30) for $W = W_k$, and if ψ' be assumed to have the form

$$\psi' = \sum_{l} \sum_{\beta} B_{l\beta} v_{l\beta}, \tag{35}$$

Eq. (31) becomes

$$\sum_{l} \sum_{\beta} B_{l\beta}(W_l - W_k) v_{l\beta} = \sum_{\alpha} A_\alpha W' v_{k\alpha} - \sum_{\alpha} \sum_{l} \sum_{\beta} A_\alpha H'_{k\alpha,l\beta} v_{l\beta}.$$

On equating the coefficients of $v_{l\beta}$ on both sides, it follows that

$$B_{l\beta} = \frac{\sum_{\alpha} A_\alpha H'_{k\alpha,l\beta}}{W_k - W_l} \qquad (k \neq l), \tag{36}$$

and

$$0 = A_\beta W' - \sum_{\alpha} A_\alpha H'_{k\alpha,k\beta}, \qquad (k = l); \tag{37}$$

the last term of Eq. (37) is not to be summed over k.

The Eqs. (36) determine the $B_{l\beta}$ in terms of the A_α, and the Eqs. (37) are ν homogeneous linear equations in the ν unknowns A_α. Unless the determinant of these vanishes, the A_α must all be zero; if

$$\begin{vmatrix} H'_{k1,k1} - W' & H'_{k1,k2} & \cdots & H'_{k1,k\nu} \\ H'_{k2,k1} & H'_{k2,k2} - W' & \cdots & H'_{k2,k\nu} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ H'_{k\nu,k1} & H'_{k\nu,k2} & \cdots & H'_{k\nu,k\nu} - W' \end{vmatrix} = 0 \tag{38}$$

then the A_α need not vanish and their ratios are determinate. Eq. (38) is an algebraic equation of the ν-th degree in W', and is called the secular equation. Its roots are the only values of W' for which Eq. (31) possesses a solution satisfying the boundary conditions of the problem.

The properties of the finite matrices $H'_{k\alpha,l\beta}$ for fixed k and l are now to be determined on the assumption that the elements of the group $\mathfrak{G}$ are integrals of both Eqs. (28) and (30), so that

$$\begin{aligned} RH &= HR, \\ RH' &= H'R, \end{aligned} \tag{39}$$

and

$$Rv_{k\alpha} = {}^{k}R_{\alpha\beta}v_{k\beta},$$

whenever R is a member of $\mathfrak{G}$. From Eq. (33) it follows that

$$H'Rv_{k\alpha} = R\{H'_{k\alpha,l\beta}v_{l\beta}\} = \sum_{l}\sum_{\beta\gamma}H'_{k\alpha,l\beta}{}^{l}R_{\beta\gamma}v_{l\gamma},$$

and

$$RH'v_{k\alpha} = H'\{{}^{k}R_{\alpha\beta}v_{k\beta}\} = \sum_{l}\sum_{\beta\gamma}{}^{k}R_{\alpha\beta}H'_{k\beta,l\gamma}v_{l\gamma}.$$

Comparing with the second of Eqs. (39), it is seen that

$${}^{k}R_{\alpha\beta}H'_{k\beta,l\gamma} = H'_{k\alpha,l\beta}{}^{l}R_{\beta\gamma}, \tag{40}$$

which is not to be summed over k and l. The infinite matrix H' can thus be broken up into finite matrices, each of which satisfies an equation similar to that entering the hypothesis of Schur's lemmas.

If all the representations ${}^{k}R_{\alpha\beta}$ are irreducible, it follows at once that $H'_{k\alpha,l\beta}$ (and therefore $B_{l\beta}$) vanishes unless ${}^{l}R_{\alpha\beta}$ is equivalent to ${}^{k}R_{\alpha\beta}$. Furthermore, the component for which $k=l$ satisfies the condition of the corollary of the last section, and must therefore have the form

$$H'_{k\alpha,k\beta} = H'(k)\delta_{\alpha\beta},$$

where $H'(k)$ is independent of α. The Eqs. (37) and (38) therefore reduce to

$$A'_{\beta}(W' - H'(k)) = 0,$$

and

$$(W' - H'(k))^{\nu} = 0.$$

The ν roots of the secular equation are therefore all equal and the energy level remains ν-fold degenerate, to this approximation at least. By discussing Eq. (32), etc., in the same manner, this result may be proven for each approximation in turn. It is therefore true of the rigorous solution of Eq. (28) also.

If the representations ${}^{k}R_{\alpha\beta}$ are not all irreducible, the situation is more complex. For simplicity, only that portion of the matrix H' for which $k=l$ will be discussed. The subscripts k may then be dropped, so that Eqs. (37) and (40) become

$$0 = A_{\beta}W' - A_{\alpha}\bar{H}'_{\alpha\beta}, \tag{37a}$$

and

$$R_{\alpha\beta}H'_{\beta\gamma} = H'_{\alpha\beta}R_{\beta\gamma}. \tag{40a}$$

The matrix $H'_{\alpha\beta}$ may be referred to as the perturbation matrix; it does not follow from Eq. (40a) that the perturbation matrix is diagonal, but these equations still impose definite restrictions upon its characteristic values. To bring these into evidence the basis v_{kn} may be supposed completely reduced by the transformation

$$u_{ipn} = a_{ipn,\alpha} v_{k\alpha}, \tag{41}$$

so that

$$Ru_{ipn} = {}^{i}R_{nm} u_{ipm}, \tag{42}$$

where ${}^{i}R_{nm}$ is an irreducible representation of $\mathfrak{G}$. The function ψ may now be written

$$\psi = c_{ipn} u_{ipn},$$

where

$$c_{ipn} = A_{\alpha} a^{-1}_{\alpha,ipn}.$$

Substituting this into Eq. (37a), one obtains

$$c_{jqm} W' - c_{ipn} H'_{ipn,jqm} = 0, \tag{43}$$

where

$$H'_{ipn,jqm} = a_{ipn,\alpha} H'_{\alpha\beta} a^{-1}_{\beta,jqm}.$$

The perturbation matrix thus undergoes a canonic transformation with the matrix a. Eq. (40a) becomes

$${}^{i}R_{nm} H'_{ipm,jqs} = H'_{ipn,jqm} {}^{j}R_{ms}$$

(which is not to be summed over i or j). The perturbation matrix has thus been broken up into smaller matrices (one for each set of values of i, p, j, q) each of which satisfies the hypothesis of one of Schur's lemmas. It therefore follows at once that

$$H'_{ipn,jqm} = H'_{pq} \delta_{ij} \delta_{n,m} \tag{44}$$

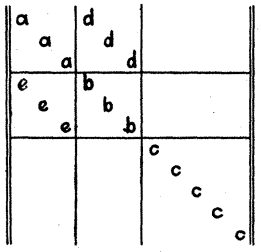
where H'_{pq} is dependent only on i, not on n.

This result may most readily be visuallized by writing the matrices out in tabular form. As an example, let $R_{\alpha\beta}$ have been reduced to the form

	11m	12m	21m
11n	${}^{1}R_{nm}$		
12n		${}^{1}R_{nm}$	
21n			${}^{2}R_{nm}$

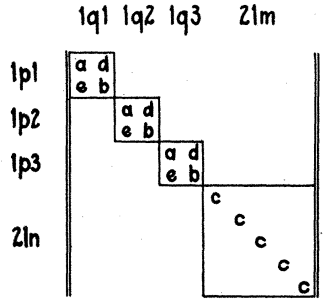
$$= {}^{i}R_{nm} \delta_{ij} \delta_{pq}. \tag{45}$$

The matrices occupying the first two squares along the diagonal are identical, but are inequivalent to the one in the third square. Then the matrix (44) must have the form



$$= H'_{pq}\delta_{ij}\delta_{nm}, \tag{46}$$

in which all elements not expressly indicated are zero, and all indicated by the same letter are equal. This matrix may be brought into another form by rearranging the order of its rows and columns:



$$= H'_{pq}\delta_{ij}\delta_{nm}. \tag{47}$$

The special form of the matrix may also be described in terms of the Eqs. (43). These consist of a number of distinct sets, one for each pair of values of j and m. Each set contains just q_j equations and involves only those q_j unknowns c_{jqm} for which j and m have this set of values. Furthermore, the coefficients of those sets having the same j (but different m) are identical. When one set has been solved, the solution of ν_j sets is thus known at once. It is obvious that the characteristic values W' are independent of m so that although the effect of the perturbation has been to remove some of the degeneracy of the unperturbed equation, the energy levels remain degenerate. This degeneracy is such that the functions belonging to a given energy level still form the basis of a representation of the group $\mathfrak{G}$, which will generally be irreducible. When the characteristic functions of a given energy level are the basis of an irreducible representation of $\mathfrak{G}$, the degeneracy is said to be *normal*. When the representation is reducible, the degeneracy is said to be *accidental*. It is seen that any perturbation whose operator commutes with the elements of $\mathfrak{G}$ (so that they are integrals of both the perturbed and unperturbed equation) can remove only the accidental degeneracy, but cannot affect the normal degeneracy. The study of the irreducible representations of $\mathfrak{G}$ is thus equivalent to the study of the normal degeneracy of the energy levels of all equations admitting these operations as integrals.

7. The multiplication table of a group. The rearrangement theorem. To use Schur's lemmas effectively in the study of irreducible representations, it is necessary to be able to calculate matrices $a_{\alpha n}$ satisfying Eq. (23) or (14). This is made possible by a theorem which depends on all three of the group postulates of Section 3, and which will be proven in this section.

Thus far, the number of distinct elements (operators) in the group $\mathfrak{G}$ has not entered into consideration. It may contain only a finite number, N, (N is then called the order of $\mathfrak{G}$) or it may contain an infinite number. This infinity may be denumerable or nondenumerable; an infinite group is called discrete or continuous according to these two alternatives. In the applications, a finite group (the permutation group) and an infinite continuous group (the rotations) will be encountered. These two cases will therefore be considered here; infinite discrete groups will not be treated.[14]

For the first, let the group be finite; its elements may then be distinguished by different letters: $I, A\ B\ C \cdots K\ L\ M\ N$ or by indices, $R_1 = I, R_2, R_3, \cdots R_N$.

The first notation will be preferred, since the number of indices to be considered is large as it is. When the elements are arranged in the order $I\ A\ B\ C \cdots M\ N$, they will be said to be in their "natural" order, though this order is as arbitrary as the symbols. It is then possible to construct a multiplication table of N rows and columns; the elements are entered in the first row and first column in their natural order, and the entry at the intersection of the c-th column and r-th row is R_rR_c, where R_r is the first element of the r-th row and R_c, the first element of the c-th column. Since $R_rR_c \neq R_cR_r$, this table will usually not be symmetric about the principal diagonal.

Example: The multiplication table of a certain group[15] of six elements $I\ A\ B\ C\ D\ E$ is

I	A	B	C	D	E
A	I	D	E	B	C
B	E	I	D	C	A
C	D	E	I	A	B
D	C	A	B	E	I
E	B	C	A	I	D

From it may be determined the various relations:

$$A^2 = B^2 = C^2 = I,\ BC = D,\ CB = E,\ \text{etc.}$$

This is not the most general multiplication table of six elements, for it will be noted that each element occurs once and only once in each row and each column. This is a direct consequence of the group postulates, and, for our purposes, their most important consequence.

[14] The theory of groups of finite order is treated by
Burnside, *l.c.* A. Speiser, Theorie der Gruppen endlicher Ordnung, (Berlin 1927).
G. A. Miller, H. F. Blichtfeldt, L. E. Dickson, Theory and applications of groups of finite order. (New York, 1916).
The Theory of continuous groups is found in Weyl, Chap. III.

[15] Speiser, *l.c.* p. 12.

It may be proven for the general case of a finite group as follows: suppose the element A does not occur in the column headed by R_c; then there is no element R_r in the first column (and therefore none in the group) for which $R_rR_c=A$. But, by Postulate III, R_r^{-1} is an element of the group and exists, and $X=AR_c^{-1}$ is also a member of the group by Postulate I. Since $XR_c=A$, this contradicts the supposition, which is therefore false. Each of the N elements of the group must thus appear at least once in each column. As there are only N entries in each column, it cannot appear more than once. Similar reasoning proves the theorem for the rows.

This result may be stated in terms of the elements of an invariant manifold. Let u be any member of $\mathfrak{M}$; then Au, Bu, etc. are also members, as is

$$v = \sum_R a(R)Ru. \tag{48}$$

In this formula, the $a(R)$ are ordinary constants, the notation indicating that a different number is associated to each element R of $\mathfrak{G}$; $a(R)$ may be called a numerical function of the elements of $\mathfrak{G}$, or, more simply, a function of R.[16] The summation indicates that R is to be given, successively, the significance of I, A, B, $C\cdots M$, N and the resulting N functions $a(R)Ru$, added together to form the function v. It is also possible to write

$$T = \sum_R a(R)R \tag{49}$$

where the operator T is defined by

$$v = Tu;$$

this defines the addition of operators in the usual manner. T is not always a member of $\mathfrak{G}$, however, though it is an integral of Eq. (1).

The importance of the theorem regarding the multiplication table is that it enables the expression

$$Sv = \sum_R a(R)RSu \tag{50}$$

to be rearranged into

$$Sv = \sum_R a(RS^{-1})Ru. \tag{51}$$

The significance of this equation may best be made clear in the case of an example based on the group I, A, B, C, D, E. Eq. (48) then reads

$$v = a(I)u + a(A)Au + a(B)Bu + a(C)Cu + a(D)Du + a(E)Eu.$$

If $S=D$, Eq. (50) becomes

$$Dv = a(I)Du + a(A)Bu + a(B)Cu + a(C)Au + a(D)Eu + a(E)Iu$$

[16] This terminology may conflict with that current in quantum theory, according to which a function of an operator or set of operators is itself an operator; where confusion is to be avoided, $a(R)$ will be called a numerical function of R. The most general operator function of the elements of a group is a linear one.

upon using the fifth column of the multiplication table for the expressions RD. The order of the terms may be rearranged into

$$Dv = a(E)u + a(C)Au + a(A)Bu + a(B)Cu + a(I)Du + a(D)Eu.$$

This is of the same form as the original expression except that the coefficient of, say Cu, is not $a(C)$ but $a(B)$, where $B = CE = CD^{-1}$. The action of the operator D on v results in a permutation of the coefficients $a(R)$. This is true only if D is a member of the group $\mathfrak{G}$, and only because each operator appears just once in each row and column of the multiplication table. If the first were not true, it would not be possible to represent Dv as a linear function of u, Au, $Bu \cdots$; if the first were true and the second not, then the coefficients of this linear function would not be so simply related to those of the original function v.

The equality of the sums (50) and (51) may be called the rearrangement theorem.[17] It may be extended without difficulty to the case of discrete infinite groups; the sum (48) is then an infinite series, and the theorem is certainly true if the series (50) converges absolutely. The extension to continuous groups requires some preliminary discussion of their multiplication table. The elements of a continuous group cannot be distinguished by integral indices, but must be considered as determined by continuously variable parameters. Let these be $\alpha \equiv (\alpha_1\alpha_2 \cdots)$; then the elements of the group may be represented by $R(\alpha_1\alpha_2 \cdots) \equiv R(\alpha)$. The significance of this notation is that

$$R(\alpha)u(x) = v(\alpha, x).$$

The result of operating on a function $u(x)$ is a function[18] of both α and x, which will be assumed to be sufficiently regular at all points of a certain range of α to make possible the analytic operations required in the following. For example, let the operations of $\mathfrak{G}$ be rotations of a plane geometrical figure about an axis perpendicular to its plane, and (x, y) the rectangular coordinates of a point in the figure. Then the parameter α may conveniently be taken to be the angle of the rotation. If $u = y^2$, then

$$R(\alpha)y^2 = (x \sin \alpha + y \cos \alpha)^2.$$

The association of the numerical values of the parameters to the elements of the group is to a large extent arbitrary (α, in the example, may be measured either in degrees or radians; $\beta = \alpha^2$ might also have been chosen as parameter). In order to be useful, however, this assignment must satisfy the condition of uniqueness: to every set of values $(\alpha_1\alpha_2 \cdots)$ in a certain range Γ there must correspond one and only one element of the group, and no element of the group may correspond to more than one set of values of α

[17] This theorem is usually proven in a more abstract form in textbooks on group theory and is not referred to by this name. See Speiser, p. 15.

[18] It may be noted that if $R_1, R_2 \cdots$ are elements of a finite group,

$$R_n u(x) = v(n,x),$$

so that this equation introduces nothing fundamentally new.

in Γ. In order that the angle α of the example may satisfy this requirement, the range Γ must be defined by $0 \leqq \alpha < 2\pi$ (if α is measured in radians). Γ is called the range of the group.

The multiplication table of the continuous group may now be given an analytic expression, even though a tabular array is impossible: if

$$R(\alpha'') = R(\alpha)R(\alpha'), \tag{52}$$

then the multiplication table is known as soon as α'' is known as a function of α and α':

$$\begin{aligned} \alpha_1'' &= \phi_1(\alpha_1\alpha_2 \cdots ; \alpha_1'\alpha_2' \cdots) \\ \alpha_2'' &= \phi_2(\alpha_1\alpha_2 \cdots ; \alpha_1'\alpha_2' \cdots) \\ &\cdots\cdots\cdots\cdots \end{aligned} \tag{53}$$

In the case of the example,

$$\alpha'' = \alpha + \alpha', \text{ mod } 2\pi.$$

The theorem regarding the multiplication table of the finite group may now be extended to continuous groups:

The equations $\alpha'' = \phi(\alpha, \alpha')$ possess the following properties:

1. If α' (or α) be held constant and α (or α') allowed to vary over the whole range Γ, then α'' also varies over the whole of Γ, nor does it assume the same value twice if α (or α') does not.

2. The equations may be solved both in the sense $\alpha = \phi'(\alpha', \alpha'')$, and $\alpha' = \phi''(\alpha, \alpha'')$, and the functions ϕ' and ϕ'' possess the same properties as ϕ.

The first of these statements may be proven by reasoning identical with that used for the finite group, the second is really implied in the first but also follows directly from the uniqueness of the reciprocal.

The analogue to Eq. (48) is now

$$v = \int_\Gamma a(\alpha)R(\alpha)u(x)d\Omega \text{ where } d\Omega = \sigma(\alpha)d\alpha_1 d\alpha_2 \tag{54}$$

The function $\sigma(\alpha)$ may be left arbitrary for the moment, except that it be single valued and finite in Γ. If $S = R(\alpha')$, then Eq. (50) becomes

$$R(\alpha')v = \int_\Gamma a(\alpha)R(\alpha)R(\alpha')u(x)d\Omega \tag{55}$$

or, on changing form α to α'' as the variables of integration,

$$R(\alpha')v = \int_\Gamma a(\alpha)R(\alpha'')u(x)\frac{\sigma(\alpha)}{\sigma(\alpha'')}\frac{\partial(\alpha_1\alpha_2 \cdots)}{\partial(\alpha_1''\alpha_2'' \cdots)}d\Omega'',$$

in which α is supposed to be expressed as a function of α' and α'' by means of proposition (2) above. *This integral is again to be extended once over the range*

Γ because of the properties of ϕ and ϕ'; this would not be the case if propositions (1) and (2) were not true. The expression may be simplified if $\sigma(\alpha)$ be determined so that

$$\frac{\sigma(\alpha)}{\sigma(\alpha'')}\,\frac{\partial(\alpha_1\alpha_2\cdots)}{\partial(\alpha_1''\alpha_2''\cdots)} = 1; \tag{56}$$

it then reads

$$R(\alpha')v = \int_\Gamma a(\alpha)R(\alpha'')u(x)d\Omega''. \tag{57}$$

The identity of the Eqs. (55) and (57) is the rearrangement theorem for continuous groups.

In the following, it will be convenient not to make any formal distinction between finite and continuous groups. This is possible if it be agreed that the notation of Eq. (48) shall also represent the integral of Eq. (54), with $a(R) \equiv a(\alpha)$, in the case of continuous groups.

The rearrangement theorem may obviously be extended to the matrices of a representation of $\mathfrak{G}$: if[19]

$$T_{nm} = \sum_R a(R)R_{nm},$$

then

$$\begin{aligned} T_{nm}S_{mk} &= \sum_R a(R)R_{nm}S_{mk} \\ &= \sum_R a(RS^{-1})R_{nk}. \end{aligned}$$

8. **The orthogonality relations.** Let ${}^1R_{nm}, {}^2R_{nm}, \cdots, {}^iR_{nm}, \cdots$, be irreducible representations, no two of which are equivalent; each is based on a closed manifold $\mathfrak{M}_i$ of ν_i dimensions. If $b_{\beta m}$ be any matrix of ν_i rows and ν_j columns, the matrix

$$a_{\alpha n} = \sum_R {}^iR^{-1}_{\alpha\beta}b_{\beta m}{}^iR_{mn}$$

satisfies the relation

$${}^iS_{\alpha\beta}a_{\beta m} = a_{\alpha n}{}^iS_{nm}$$

for all operators S in the group. This may be proven as follows:

$$a_{\alpha n}{}^iS_{nm} = \sum_R \sum_k c_{\alpha k}(R){}^iR_{kn}{}^iS_{nm},$$

where

$$c_{\alpha k}(R) = {}^iR^{-1}_{\alpha\beta}b_{\beta k}.$$

On applying the rearrangement theorem to the summation over R:

[19] Schur, *l.c.*, calls a matrix defined in this manner a "group matrix."

$$a_{\alpha n}{}^{i}S_{nm} = \sum_{R}\sum_{k} c_{\alpha k}(RS^{-1})\,{}^{i}R_{km}.$$

Now

$$c_{\alpha k}(RS^{-1}) = {}^{i}S_{\alpha\beta}c_{\beta k}(R),$$

and hence

$$\begin{aligned} a_{\alpha n}{}^{i}S_{nm} &= {}^{i}S_{\alpha\beta}\sum_{R}\sum_{k} c_{\beta k}(R)\,{}^{i}R_{km} \\ &= {}^{i}S_{\alpha\beta}a_{\beta m}, \end{aligned}$$

which was to be proven. Hence by Schur's lemmas, $a_{\alpha n}$ must vanish whenever $i \neq j$ and have the form $a(i)\delta_{\alpha n}$ when $i=j$. Since the matrix $b_{\beta m}$ is entirely arbitrary, this is equivalent to

$$\sum_{R} {}^{i}R^{-1}_{\alpha\beta}\,{}^{j}R_{mn} = \gamma_{\beta m}\delta_{ij}\delta_{\alpha n}, \tag{58}$$

where the matrix $\gamma_{\beta m}$ is independent of α and n. This matrix may readily be determined: setting $i=j$, $\alpha=n$ in Eq. (58) and summing over n from 1 to ν_i, the result is

$$\sum_{R} {}^{i}R^{-1}_{n\beta}R_{mn} = \sum_{R}\delta_{m\beta} = \nu_i\gamma_{\beta m}.$$

The second sum is equal to

$$N\delta_{m\beta}$$

where N is the order of $\mathfrak{G}$ if it is a finite group, or

$$N = \int_{\Gamma} d\Omega$$

if $\mathfrak{G}$ is continuous. Hence Eq. (58) becomes [20]

$$\sum_{R} {}^{i}R^{-1}_{\alpha\beta}\,{}^{i}R_{mn} = \frac{N}{\nu_i}\delta_{ij}\delta_{\alpha n}\delta_{\beta m}. \tag{59}$$

If these representations have been chosen so that their matrices are all unitary, $({}^{i}R^{-1}_{\alpha\beta} = {}^{i}R^{*}_{\beta\alpha})$ then Eq. (59) may also be written

$$\sum_{R} {}^{i}R^{*}_{\beta\alpha}\,{}^{i}R_{mn} = \frac{N}{\nu_i}\delta_{ij}\delta_{\alpha n}\delta_{\beta m}. \tag{59a}$$

In the case of a continuous group, each element ${}^{i}R_{\alpha\beta}$ (for given i, α, β) is a function of the group parameters only; Eq. (59a) states that these functions are all orthogonal to one another. In the case of a finite group, the elements ${}^{i}R_{\alpha\beta}$ (for fixed i, α, β again) are vectors in a certain space of N dimensions (the "group manifold"), and Eq. (59a) states that these vectors are orthogonal.

[20] The arrangement of the indices on the two sides of this equation is to be noted carefully. Reference: Weyl, *l.c.* Chap III.

For this reason the Eq. (59a), and the more general Eq. (59) as well, will be referred to as the orthogonality relations between the representations. They constitute a powerful analytic tool for the solution of complex problems.

An important corollary may be deduced from these equations in the case of finite groups: since there can be at most N orthogonal vectors in a space of N dimensions, and the ${}^iR_{\alpha\beta}$ constitute $\sum_{all\ i}\nu_i^2$ such vectors, it follows that this sum must be less than or at most equal to N. Since $\nu_i^2 \geqq 1$, it follows that *there are at most a finite number of inequivalent irreducible representations of a finite group.* The problem of finding *all* inequivalent irreducible representations of a finite group thus acquires a definite significance, and admits of solution. In the case of continuous groups, the orthogonality relations also materially restrict the irreducible representations, though their number is infinite. All inequivalent irreducible representations of the three dimensional rotation group will be determined in the special part of this review.

The converse of the theorem is also true: if a given sequence of representations ${}^iR_{\alpha\beta}$ satisfies Eq. (59) then they are irreducible and inequivalent. To prove this, let $A_{\alpha m}$ be a rectangular matrix satisfying

$$ {}^iR_{\alpha\beta}A_{\beta m} = A_{\alpha n}{}^iR_{nm} $$

for all R. On multiplying both sides by ${}^iR^{-1}_{fq}$ and summing over all R, Eq. (59) results in

$$ A_{pm}\delta_{q\alpha} = A_{\alpha q}\delta_{pm}\delta_{ij}. $$

If $i \neq j$, A_{pm} must be zero because the right side vanishes; if $i=j$, $A_{pm}=A\delta_{pm}$, where A is independent of p or m. Hence, the representation ${}^iR_{nm}$ is irreducible.

9. **The complete reduction of a given invariant manifold.** Let the sequence of representations ${}^1R_{nm}$, ${}^2R_{nm}\cdots$ of the previous section be known to contain all inequivalent irreducible representations of $\mathfrak{G}$. Then the task of completely reducing a given invariant manifold $\mathfrak{M}$ can be accomplished by a series of multiplications and summations (or integrations in the case of a continuous group). Let $\mathfrak{M}$ be given in terms of its basis $w_1, w_2, \cdots w_\nu$, so that

$$ Rw_\alpha = R_{\alpha\beta}w_\beta. \tag{60} $$

Then, if $a_{m\alpha}$ be any arbitrary matrix of ν_i rows and ν columns, the ν_i functions

$$ v_{in} = \sum_R {}^iR^{-1}_{nm}a_{m\alpha}R_{\alpha\beta}u_\beta \tag{61} $$

satisfy the relations

$$ Sv_{in} = {}^iS_{nm}v_{im} \tag{62} $$

if S is any member of $\mathfrak{G}$. To prove this, it is merely necessary to substitute from Eq. (61) on the left side of Eq. (62) and apply the rearrangement theorem.

Hence, *if the functions v_{in} do not vanish identically* they subtend an irreducible invariant submanifold of $\mathfrak{M}$. Repeating the process for various values of i and different a_{mn}, all irreducible submanifolds of $\mathfrak{M}$ can be found. Furthermore, that basis in each which is adapted to the matrices ${}^iR_{nm}$ of the standard representation is automatically obtained. This process smacks somewhat of trial and error, however, for often the functions v_{in} will vanish identically, and other times the same invariant submanifold will be obtained. A still more complicated situation arises when several nonvanishing bases for the same representation are obtained: let v_{in}, v'_{in}, v''_{in} all be bases for ${}^iR_{mn}$. Then it is quite possible that

$$v_{in} = av'_{in} + bv''_{in}$$

where a and b are independent of n. There is no general way of excluding such [21] linear relations, and therefore this method of reduction is not always the most practicable. Less laborious methods can often be found in special cases.

One general result may be deduced, however, for the completely reduced basis u_{ipn} which is determined in this manner may be written

$$u_{ipn} = \sum_R {}^iR^{-1}_{nm} a^p_{m\alpha} R_{\alpha\beta} w_\beta, \tag{61a}$$

where the matrices $a^p_{m\alpha}$ are arbitrary except for the restriction that the u_{ipn} be ν in number and linearly independent. The coefficients $a_{ipn,\alpha}$ of Eq. (19a) are therefore given by

$$a_{ipn,\alpha} = \sum_R {}^iR^{-1}_{nm} a^p_{m\beta} R_{\beta\alpha}.$$

10. **The character of a representation.** The element of trial and error in the method of reduction just outlined will be materially reduced if some general method can be found for determining those representations ${}^iR_{nm}$ which appear in the completely reduced form of a given representation, and the number of times q_i which each is repeated. (cf. Eq. (17a). This problem is solved by the theory of characters.[22]

This theory is based on the proposition that the sum of the diagonal elements or "spur" of a matrix is equal to the spur of any matrix arising from it by a canonic transformation: i.e., that

$$\sum_{n=1}^{\nu} R_{nn} = \sum_{\alpha=1}^{\nu} {}'R_{\alpha\alpha} \tag{63}$$

[21] Linear relations of the form

$$av_{in} + bv_{jm} = 0$$

cannot exist when $i \neq j$, $m \neq n$. For on operating with R,

$$a^iR_{nk}v_{ik} + b^jR_{ml}v_{jl} = 0;$$

on multiplying this by ${}^iR^{-1}_{nk}$ and summing over all R, Eq. (59) results in $a=0$.

[22] Weyl, l.c. p. 117.

whenever

$$'R_{\alpha\beta} = a_{\alpha n} R_{nm} a^{-1}_{m\beta}. \tag{64}$$

This proposition is obvious on setting $\alpha=\beta$ in Eq. (64) and summing, for

$$\sum_{\alpha=1}^{\nu} a_{\alpha n} a^{-1}_{m\alpha} = \delta_{nm}.$$

Hence, the numerical function of the operators R which is defined by

$$\chi(R) = \sum_{n=1}^{\nu} R_{nn} \tag{65}$$

is the same for all representations $'R_{\alpha\beta}$ which are equivalent to R_{nm}. The converse of this proposition is also true for irreducible representations, as will be shown. The function $\chi(R)$ is therefore called the character or characteristic of the representation R_{nm}. If R_{nm} is irreducible, its character is said to be primitive.

On setting $\alpha=\beta$, $m=n$ in Eq. (59) and summing over α and m, the following important property of primitive characters becomes evident:

$$\sum_R {}^i\chi(R^{-1})\,{}^i\chi(R) = N\delta_{ij}. \tag{66}$$

(The indices i indicate the representation of which ${}^i\chi(R)$ is the character.) Since every representation of the group $\mathfrak{G}$ is equivalent to a unitary one, for which ${}^iR^{-1}_{nm} = {}^iR^{*}_{mn}$, it follows that

$${}^i\chi(R^{-1}) = [{}^i\chi(R)]^*$$

so that Eq. (66) may be written

$$\sum_R {}^i\chi^*(R)\,{}^i\chi(R) = N\delta_{ij}. \tag{68}$$

This is called the orthogonality relation between primitive characters. The converse of the above proposition follows at once from this equation; for suppose $i\neq j$, but ${}^i\chi(R) = {}^j\chi(R) = \chi(R)$ for all R. Then

$$\sum_R |\,\chi(R)\,|^2 = 0$$

by Eq. (68), whence $\chi(R)=0$ for all R, including $R=I$. But $\chi(I) = \sum_n I_{nn} = \nu$, and hence the number of rows and columns in the matrices of the two representations would have to be zero—an absurd result.

The problem proposed in the first prargraph of this section can now be solved with considerable elegance: let $R_{\alpha\beta}$ be equivalent to a completely reduced representation of the form (17a), in which ${}^iR_{mn}$ appears q_i times. (If a particular irreducible representation j does not appear, $q_j=0$). Then

$$\chi(R) = \sum_\alpha R_{\alpha\alpha} = \sum_i q_i\,{}^i\chi(R);$$

on multiplying this equation by ${}^i\chi^*(R)$ and summing over R, Eq. (68) yields

$$\sum_R {}^i\chi^*(R)\chi(R) = Nq_i. \tag{69}$$

The computational work involved in calculating the character of $R_{\alpha\beta}$ and the sums (69) is usually less than that involved in determining by trial which of the sums (61) yields a nonvanishing set of v_{in}. Furthermore the knowledge of the number of independent sets of v_{in} to be expected is a definite advantage, and is sometimes sufficient to solve a given problem without actually carrying out the reduction of the manifold $\mathfrak{M}$.

The function $\chi(R)$ possesses one general property which materially reduces the amount of work required for its calculation; the numerical values of $\chi(R)$ and $\chi(S)$ are equal when R and S belong to the same "class" of the group $\mathfrak{G}$. A "class" of the elements belonging to a given group is defined as follows: if R and T are in the group, so is $S = TRT^{-1}$, and S is said to belong to the same class as R, or to be equivalent to R. It may also be said that S arises from R by means of a canonic transformation with a third element T of the group. Since the matrix S_{nm} is equal to $T_{nk}R_{kl}T^{-1}_{lm}$, the equation $\chi(R) = \chi(S)$ follows from the proposition regarding the invariance of the diagonal sum under a canonic transformation.

The number of classes, N_c, in a given finite group cannot be greater than the order of the group, and there are thus only $N_c \leqq N$ different numbers required to specify each character. As all primitive characters are linearly independent because of Eq. (69), it follows that there are at most N_c such characters, and hence at most N_c inequivalent irreducible representations. This restricts the number of irreducible representations of a finite group still more than Eq. (59).

11. **Miscellaneous theorems.** I. Any basis u_n of a unitary irreducible representation of $\mathfrak{G}$ is orthogonal, except for a normalizing factor. For

$$\begin{aligned}(u_n, u_m) &= (Ru_n, Ru_m)\\ &= R_{nk}R_{ml}{}^*(u_k, u_l);\end{aligned}$$

summing over R and using Eq. (59a) which applies to unitary irreducible representations,

$$N(u_n, u_m) = \delta_{nm} \sum_k (u_k, u_k)$$

or

$$(u_n, u_m) = C\delta_{nm},$$

where C is independent of n.

II. If two groups are given, which have no common element other than I, and if $QR = RQ$ whenever R belongs to one group and Q to the other, they can be combined into a single group whose elements are (1) all the elements R of one group; (2) all the elements Q of the other; (3) all products of the

form RQ. The resulting group is called the direct product of the two given groups.

If R_{nm} and $Q_{\alpha\beta}$ are irreducible representations of the two groups, the matrices $R_{nm}Q_{\alpha\beta}$ (whose rows are numbered with the two indices n,α and columns with m, β) form an irreducible representation of their direct product. This is most readily proven by noting that the elements of these matrices satisfy the orthogonality relations (59).

III. If a function u is given, which can be subjected to all the operations of a group $\mathfrak{G}$, it is possible to construct a linear manifold of functions invariant under $\mathfrak{G}$. For let R, S, be any operations of $\mathfrak{G}$, then the manifold $\mathfrak{M}$ may be defined by the statement that Ru and Su are always members, as is any linear function of its members. The most general function of this manifold is representable as

$$\psi = \sum_R a(R)Ru.$$

The linearly independent among the functions Ru may be chosen as its basis. It is said to be generated by u and $\mathfrak{G}$.

If all the functions Ru are linearly independent, it follows that v_{in} (Eq. (61)) cannot vanish for any i or any $a_{m\alpha}$. There will therefore be precisely ν_i different bases for the representation ${}^iR_{nm}$ contained in $\mathfrak{M}$.

IV. If the elements of two groups can be multiplicatively associated (cf. Section 2) in such a manner that every element of either group corresponds to one and only one element of the other, the groups are said to be isomorphic. Their multiplication tables are identical if one abstracts from superficial differences in notation and arrangement. It is even possible to abstract from more fundamental differences: the elements of one group may be operators, those of the other matrices. The matrices representing an operator group sometimes (but not always, cf. reference 7) form an isomorphic group. It is obvious that any representation of a group is also a representation of any group isomorphic to it. This theorem will be exceedingly useful in the following.

V. If R_n is a ν-dimensional irreducible representation and

$$\chi_{nm} = \sum_{class} R_{nm},$$

where the summation is not to be extended over the whole group, but only over a single class of elements, then

$$\chi_{nm} = \frac{N_c}{\nu}\chi(R)\delta_{nm},$$

R being any of the members of this class, and N_c the number of members in the class if the group is finite; if the group is continuous,

$$N_c = \int_{class} d\Omega.$$

Proof: if S is any element of the group

$$S_{nm}\chi_{mk}S^{-1}_{kl} = \sum_{class} S_{nm}R_{mk}S^{-1}_{kl}.$$

But $SRS^{-1} = T$ belongs to the same class as R.
Hence

$$S_{nm}\chi_{mk}S^{-1}_{kl} = \sum_{class} T_{nl} = \chi_{nl}.$$

The matrix χ_{nm} thus commutes with all members of the representation, and must therefore have the form $a\delta_{nm}$ where a is independent of n. The constant a may be evaluated by taking the diagonal sum:

$$\nu a = \sum_{class} \chi(R) = N_c \chi(R).$$

This theorem has been used by Dirac[23] as a definition of the characters $\chi(R)$.

II. Specification of the Groups to be Considered

12. **The integrals in the absence of an external field. The case of one electron.** In Section 3, the class of integrals under consideration was restricted by certain formal postulates. In this section, integrals are to be determined which satisfy these conditions. For the first, the spin variables s may be omitted from consideration,[24] and only the translatory variables considered; then if the system consists of single electron moving in a centra field of force, the wave equation must be invariant under a rotation of coordinate axes about an axis through the center of force. This fact is sufficient to determine some of the integrals of the group $\mathfrak{G}$.

This may be proven in a somewhat general form: let $H(x)$ and $f(x)$ be an arbitrary operator and function, respectively, involving the variables $x = x_1x_2\cdots$, and

$$x_i = \phi_i(x'_i) \tag{70}$$

a substitution of variables. Let R denote the operation of performing this substitution on any function of x; then

$$\begin{aligned} R[H(x)f(x)] &= H(\phi(x'))f(\phi(x')) \\ &= H'(x')f'(x'). \end{aligned} \tag{71}$$

If $H'(x) = H(x)$, the operator H is said to be invariant under the substitution R. Suppose this to be the case, and that $f(x)$ satisfies the equation

$$H(x)f(x) = Wf(x); \tag{72}$$

then

$$R[H(x)f(x)] = H(x')f'(x') = Wf'(x').$$

[23] P. A. M. Dirac, Proc. Roy. Soc. **A123**, 721 (1929).
[24] E. Wigner, Zeits. f. Physik **43**, 624; **45**, 601 (1927).

If the accent on the variables x be dropped after the substitution it is seen that $f'(x)$also satisfies Eq. (72) for the same value of W as $f(x)$. With a slight modification of the definition of the operation R, $f'(x)$ may be written $Rf(x)$, and R is then an integral of the Eq. (72).

The operation R may also be given a geometrical interpretation, for Eq. (70) relates every point x' (Fig. 1) to another point $x=\phi(x')$ of the same space

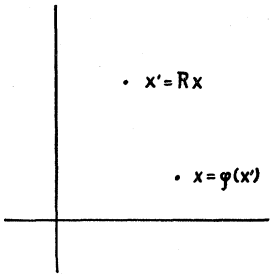


Fig. 1.

in such a manner that when either is given, the other may be determined. The function $f'(x)=Rf(x)$ assumes the same numerical value at $x=x'$ as $f(x)$ assumes at $x=\phi(x')$. The point x' may also be denoted by Rx, whereupon $f'(Rx)=f(x)$.

In the present case, the functions (70) are linear and homogeneous in the coordinates $x'y'z'$:

$$\begin{aligned} x &= a_1x' + a_2y' + a_3z' \\ y &= b_1x' + b_2y' + b_3z' \\ z &= c_1x' + c_2y' + c_3z' \end{aligned} \tag{73}$$

which represents a rotation of coordinate axes if

$$\begin{aligned} a_1b_1 + a_2b_2 + a_3b_3 &= 0, \quad \text{etc.}, \\ a_1^2 + a_2^2 + a_3^3 &= 1, \quad \text{etc.}, \\ \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} &= +1. \end{aligned} \tag{74}$$

The coefficients a_n, b_n, c_n, may all be expresseed in terms of the three Eulerian angles α, β, γ of the rotation (cf. Section 19). The operation R is to be defined as before and will be denoted by $R(\alpha, \beta, \gamma)$.

As a simple example, consider the Schrödinger equation for a free electron. It possesses the solution $f=\exp ikz$, where $h^2k^2/8\pi^2m=W$. Since it is invariant under the substitution (73), it therefore also possesses the solution $\exp ik(c_1x+c_2y+c_3z)$.

It is geometrically evident that the rotations form a group. They are also unitary operators when the inner product is defined as in Eq. (3) or (3a), as may be verified by performing the substitution of variables (73) in the integrals.

Before considering the general case of the wave equation involving all four coordinates x, y, z, s, it will be well to examine functions of s alone. The most general function of a variable whose only possible values are $+1/2$ and $-1/2$ may always be written in the form

$$\psi = A\xi(s) + B\eta(s) \tag{75}$$

where A and B are independent of s, and

$$\begin{aligned} \xi(\tfrac{1}{2}) &= 1, \quad \xi(-\tfrac{1}{2}) = 0, \\ \eta(\tfrac{1}{2}) &= 0, \quad \eta(-\tfrac{1}{2}) = 1. \end{aligned} \tag{76}$$

The most general linear operation, Q, which may be performed on such a function is a linear substitution of the quantities ξ and η. If a, b, c, d are arbitrary constants and

$$\begin{aligned} \xi &= a\xi' + b\eta' \\ \eta &= c\xi' + d\eta', \end{aligned} \tag{77}$$

then Q is defined by

$$Q\psi = A(a\xi + b\eta) + B(c\xi + d\eta). \tag{78}$$

If ψ is also a function of x, y, z, then it is merely necessary to consider A and B as functions of these variables. Such functions can be acted on by any operator T which is defined so as to be able to act on functions of x, y, z, and also by Q. The most general operation which may be peformed on them is a sum of products like TQ. It is to be noted that $TQ = QT$.

The wave equation of a spinning electron is then

$$H\psi = W\psi,$$

where H is an operator of this general form. Its specific expression need not concern us: its only property which is important is that it possesses integrals of the form $R(\alpha, \beta, \gamma)D(\alpha, \beta, \gamma)$, where $D(\alpha, \beta, \gamma)$ is a particular operation like Q. The form of D may be deduced directly from the postulates of the Pauli or Dirac spin theories.[25] The result may be stated without proof:

$$D: \qquad \begin{aligned} \xi &= a\xi' + b\eta' \\ \eta &= -b^*\xi' + a^*\eta' \end{aligned} \tag{79}$$

where

$$aa^* + bb^* = 1$$

and

$$\begin{aligned} a &= \cos\beta/2\, e^{i(\alpha+\gamma)/2} \\ b &= \sin\beta/2\, e^{i(-\alpha+\gamma)/2}. \end{aligned} \tag{80}$$

[25] P. A. M. Dirac, Proc. Roy. Soc. **A117**, 610; **118**, 351 (1928).
W. Pauli, Zeits. f. Phys. **43**, 601 (1927).
P. Jordan, Zeits. f. Phys. **44**, 21 (1927).
J. V. Neumann and E. Wigner, Zeits. f. Phys. **47**, 203 (1928), **49**, 73 (1928).
F. Möglich, Zeits. f. Phys. **48**, 852 (1928).

The quantities a and b have been used by Cayley and Klein[26] in the study of the rotation of a rigid body, and are often called the Cayley-Klein parameters.

The operations D (though not integrals) also satisfy the group postulates and are unitary operators. The last remark may again be verified by performing the substitution in Eq. (13a). To prove that they form a group, it may be noted that the expressions (80) for a and b are the most general consistent with $aa^*+bb^*=1$, provided only that α, β, γ are entirely arbitrary. Furthermore, if D_{st}, D'_{tr} be two substitutions of the form (79), their resultant substitution $D''_{sr}=D_{st}D'_{tr}$ is again of the form (79), with

$$\begin{aligned} a'' &= aa' - bb'^* \\ b'' &= ab' + ba'^*. \end{aligned} \tag{81}$$

These equations may also be written in terms of α, β, γ, etc. If there were no relation between the a and b of the operations D, and the α, β, γ of R, the operators $R(\alpha\beta\gamma)D(ab)$ would form the direct product (cf. Section 11, theorem II) of the groups R and D. However, the operators here considered do not include all the products $R(\alpha, \beta, \gamma)D(ab)$, but only those for which a and b are expressed in terms of α, β, γ by Eq. (80). It is not obvious from what has been said here that this subset forms a group, but this fact follows from the considerations of Neumann, l. c. An independent proof will be given in Section 19.

These are not the only members of the group $\mathfrak{G}$ in the absence of an external field. If the spin be again neglected for a moment, the wave equation must be invariant when the coordinate system is changed from right- to left-handed. This transformation may most readily be accomplished by the substitution

$$K: \qquad \begin{aligned} x &= -x', \\ y &= -y', \\ z &= -z'. \end{aligned} \tag{82}$$

Let K denote the operation of performing this substitution; then $K^2=\mathrm{I}$ and $KR=RK$, as is evident from the geometry involved. The operation K is then also a member of the group $\mathfrak{G}$ of the spin-free equation. It remains to be considered how K is to be extended to apply to the spin-variable s. It might be expected that it would be accompanied by a substitution like (78), but this is not the case. This may be made plausible by noting that the variables s are analogous to an angular momentum:

$$xp_y - yp_x = m(x\dot{y} - y\dot{x}).$$

Since this is an even function of the coordinates, the substitution (82) will not change it. The angular momentum is a rotor, not a vector. By analogy, then, it is defined that

$$Kf(x, y, z, s) = f(-x, -y, -z, s), \tag{83}$$

and this operator is also an integral of the Eq. (77).

[26] F. Klein and A. Sommerfeld, Theorie des Kreisels.

The complete group $\mathfrak{G}$ for the wave equation of a single spinning electron in a central field of force thus consists of the direct product of the groups RD and I, K.

13. **Extension to the case of N electrons.** The extension of the foregoing results to the case of N electrons moving under their mutual actions and that of a central field, is readily accomplished since all N electrons are identical and act on each other with forces dependent only on their relative positions and velocities. The operation R is to be defined as the substitution (73) applied to all N coordinate sets x_i, y_i, z_i simultaneously; the same applies to K. The operation D is also readily generalized, for the most general function involving N spin-variables s_i can always be expressed as a sum of 2^N terms, of which

$$\psi(x_i, y_i, z_i)\xi(s_1)\eta(s_2)\eta(s_3)\cdots\xi(s_N) \tag{84}$$

is typical. Each pair of functions $\xi(s_i)$, $\eta(s_i)$ will then undergo the substitution (79). That the operations RD as thus generalized are integrals of the Pauli wave equation will again be assumed without proof. The Dirac equation for one electron has not yet been generalized to the case of N electrons.

These operators obviously form a group which is isomorphic to that considered in the last section and thus possesses the same representations as that group (cf. Section 11, Theorem IV). It does not constitute the whole of the group of integrals in this case, however.

The identity of the N electrons introduces further integrals into the group $\mathfrak{G}$, for the wave equation must be invariant when the coordinates of two electrons are interchanged. If $\psi(x_1, y_1, z_1, s_1; x_2, y_2, z_2, s_2; \cdots,)$ is a solution, $\psi(x_2, y_2, z_2, s_2; x_1, y_1, z_1, s_1; \cdots,)$ will also be a solution. More generally, any permutation of the coordinates of the electrons will result in another solution. Denote these operations by P; they form a finite group of order $N!$. As the permutations commute with the operations of the groups already enumerated, the whole group $\mathfrak{G}$ is simply their direct product.

14. **The integrals in the presence of an external field.** When the system of electrons is acted on by an homogeneous field in addition to the central field of the nucleus and their mutual actions, the equation will no longer be invariant under all rotations of the coordinate system, but only under those rotations whose axes coincide in direction with the external field. This direction may be taken as the z-axis, and $S(\omega)$ may denote a rotation about this axis through the angle ω. The permutations P will still be integrals, since the external field acts on each electron according to the same law.

The operation K is no longer an integral either, but another operation may be found to replace it in the group $\mathfrak{G}$. This operation is different for electric or magnetic fields. In the latter case, the wave-equation is invariant under the transformation

$$K': \quad x_i = x_i', \quad y_i = y_i', \quad z_i = -z_i', \tag{85}$$

which represents a reflection of the coordinate system in the xy-plane. It will be evident that $K'^2=I$ and $K'S(\omega)=S(\omega)K'$. The group $\mathfrak{G}$ then consists of the direct products of the groups $PS(\omega)$ and I, K'.

In the presence of an electric field along the z-axis, neither K nor K' are integrals, but the operation

$$K'': \begin{aligned} x_i &= x_i', \\ y_i &= -y_i', \\ z_i &= z_i', \end{aligned} \tag{86}$$

which represents a reflection in the xz-plane, is an integral. It differs from K' in that it does not commute with $S(\omega)$. To see this, let I be the point whose coordinates are (x, y), K'' the point whose coordinates are $(K''x, K''y)$, S the point (Sx, Sy), etc. These points are plotted in Fig. 2. The

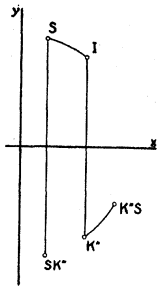


Fig. 2.

points SK'' and $K''S$ do not coincide, but it is seen that $SK''=K''S^{-1}$. This relation is obviously general. As $K''^2=I$, it follows that all products (e.g., $S(\omega)K''S(\omega_1)$) can be expressed either in the form $K''S(\omega)$ or else as $S(\omega)$. In the example, $S(\omega)K''S(\omega_1)=K''S^{-1}(\omega)S(\omega_1)=K''S(\omega_1-\omega)$. Therefore the operations, K'', $S(\omega)$, $K''S(\omega)$ form a group, even though K'' and $S(\omega)$ do not commute. The group $\mathfrak{G}$ is the direct product of this and the permutation group. When the spin of the electrons is taken into account, it is fairly obvious that the rotation $S(\omega)\equiv R(\omega, 0, 0)$ (cf. Section 19) is to be replaced by $S(\omega)D(\omega, 0, 0)$, where $D(\omega, 0, 0)$ is the unitary substitution associated to $R(\omega, 0, 0)$.

If the external field is not homogeneous, or consists of crossed electric and magnetic fields, no rotations are integrals, nor is there any operation analogous to K. The group $\mathfrak{G}$ then consists only of the permutations.

III. Construction of Irreducible Representations of the Rotation Group, etc.

15. **The group *I, K*.** The irreducible representations of the finite group of order $N=2$ which contains the identity I and[27] the element K, and whose multiplication table is $K^2=I$, can be found without difficulty. They are

[27] The element K may be replaced by K' or K'' without altering the validity of these considerations.

two in number (for there are two classes in the group) and because of Eq. (59b) each has but one row and column in its matrices. They are: $I=1$, $K=1$ and $I=1$, $K=-1$. The reduction of any given manifold $\mathfrak{M}$ by the method of Section 9 is simple: let u be any member of $\mathfrak{M}$, then the sums

$$\sum_R {}^i R_{nm}^{-1} Ru$$

are easily evaluated:

$$v = u + Ku, \quad w = u - Ku,$$

are functions satisfying

$$Kv = v, \quad Kw = -w,$$

respectively. The tests for linear independence will always be simple and there is little more to be said about this group. It is an almost trivial example of the application of the foregoing methods.

16. **The two-dimensional rotations.** The next simplest of the groups we have discussed in Part II is the group of rotations about the z-axis. Its operators are defined by

$$S\psi(xyz) = \psi(x'y'z'), \tag{87}$$

where

$$\begin{aligned} x &= x' \cos\omega + y' \sin\omega, \\ y &= -x' \sin\omega + y' \cos\omega, \\ z &= z'. \end{aligned}$$

The members of this group all commute with one another:

$$S(\omega)S(\omega_1) = S(\omega_1)S(\omega) = S(\omega_1 + \omega),$$

the sum $\omega+\omega_1$ being taken mod 2π. Any group whose members have this property is called Abelian, and its irreducible representations are necessarily ordinary numbers. For suppose $S_{nm}(\omega)$ were an irreducible representation whose matrices had more than one row and column, then

$$S_{nm}(\omega)S_{mk}(\omega_1) = S_{nm}(\omega_1)S_{mk}(\omega),$$

and therefore

$$S_{nm}(\omega) = a(\omega)\delta_{nm}$$

by the corollary to Schur's lemmas. This is in contradiction to the assumption of irreducibility.

Let $a(\omega)$ represent the operator $S(\omega)$. Then

$$a(\omega)a(\omega_1) = a(\omega + \omega_1),$$

which is a functional equation having as its solution

$$a(\omega) = e^{im\omega}. \tag{88}$$

Because $S(\omega+2\pi)=S(\omega)$, the constant m can take on only integral values (positive or negative). (Cf. the condition of unique correspondence, Section 2.)

The integral over the group is

$$\frac{1}{2\pi}\int_0^{2\pi}\cdots d\omega$$

and the orthogonality relations are

$$\frac{1}{2\pi}\int_0^{2\pi} e^{-im\alpha}e^{in\alpha} = \delta_{nm}.$$

The characters are in this case identical with the "matrices" $a(\omega)$.

17. **The representation of $\mathfrak{G}$ in the case of an external field.** The irreducible representations of the groups which dominate the theory of external fields are readily constructed. The case of a magnetic field may be treated first. The group $\mathfrak{G}$ then consists of the direct product of the groups I, K, and $S(\omega)$. Its representations are therefore two in number (cf. Section 11, theorem II) and can be written down immediately as

$$\begin{aligned} &S(\omega)\to e^{im\omega}, \quad K'\to +1,\\ &\text{and}\\ &S(\omega)\to e^{im\omega}, \quad K'\to -1. \end{aligned} \tag{89}$$

The group $\mathfrak{G}$ in the case of an homogeneous electric field is not Abelian, as has been seen in Section 14. Its representations will therefore not all be one-dimensional. To find them, let $\mathfrak{M}$ be a given manifold which is invariant under $\mathfrak{G}$, and suppose its basis given in such a form that any member of it, say u_m, satisfies the relation

$$S(\omega)u_m = e^{im\omega}u_m.$$

It is always possible to find such a basis, since $\mathfrak{M}$ is also invariant under the subgroup $S(\omega)$. The function $v_m=K''\,u_m$ will then satisfy the relation

$$S(\omega)v_m = e^{-im\omega}v_n,$$

which follows at once from $K''S=S^{-1}K''$. The submanifold subtended by the two functions u_m, v_m is invariant and, in general, irreducible. The representation based on it is two-dimensional:

$$S(\omega)\to\left\|\begin{matrix} e^{im\omega} & 0\\ 0 & e^{im\omega}\end{matrix}\right\|,\quad K''\to\left\|\begin{matrix}0&1\\1&0\end{matrix}\right\|. \tag{90}$$

In the particular case $m=0$, this submanifold is reducible and (90) breaks up into two one-dimensional representations:

$$\begin{aligned} &S(\omega)\to 1, \quad K''\to +1,\\ &\text{and}\\ &S(\omega)\to 1, \quad K''\to -1. \end{aligned} \tag{91}$$

18. **The unitary substitutions of two variables: $D(\alpha\beta\gamma)$.** It would be logical to consider the three-dimensional rotation group at this point, but it is expedient to consider the group of unitary substitutions of two variables $D(\alpha\beta\gamma)$ first. These two groups are intimately related, as will be shown in the next section, and the analysis of $D(\alpha\beta\gamma)$ is simpler.[28]

The operations $D(\alpha\beta\gamma)$ may be considered as defined by Eqs. (78), (79) and (80). For the present purpose, the special significance of ξ and η is immaterial and they may be taken to be any two complex variables which undergo this substitution. It is unitary, and hence

$$\xi\xi^* + \eta\eta^* = \xi'\xi'^* + \eta'\eta'^*, \tag{92}$$

as may be verified by performing the substitution (91) on the left hand side of (92). The reciprocal of the matrix (79) is therefore

$$\begin{Vmatrix} a^* & -b \\ b^* & a \end{Vmatrix}. \tag{93}$$

The substitution may be performed on any function of ξ and η, not only on linear ones:

$$Df(\xi\eta) = f'(\xi\eta), \tag{94}$$

where

$$f'(\xi'\eta') = f(a\xi' + b\eta', -b^*\xi' + a^*\eta').$$

The multiplication table of the group has already been discussed (cf. Eq. (81)); as a particular case, it will be verified that

$$D(\alpha\beta\gamma) = D(\gamma 00)D(0\beta 0)D(\alpha 00). \tag{95}$$

The group is obviously a three parameter continuous one, and it is necessary to determine the range Γ of its parameters. An inspection of Eq. (80) will convince the reader that if

$$0 \leqq \alpha < 4\pi, \qquad 0 \leqq \beta < \pi, \qquad 0 \leqq \gamma < 2\pi \tag{96}$$

the quantities a and b will assume all pairs of values satisfying $aa^* + bb^* = 1$, and assume each set of values only once. The inequalities (96) thus determine Γ (cf. Section 7).

The classes of elements $D(\alpha\beta\gamma)$ are readily determined. It is a well-known theorem, that if a unitary matrix D_{st} be given, another unitary matrix of determinant $+1$ can be found such that when D_{st} is subjected to a canonic transformation with the latter, it is transformed into a diagonal matrix. As the second matrix is a member of the group here under consideration, it follows that D_{st} belongs to the same class as a matrix

$$\begin{Vmatrix} a' & 0 \\ 0 & a'^* \end{Vmatrix}. \tag{97}$$

[28] Weyl, *l.c.* p. 107.

Since this matrix must be unitary, it follows that $a' = \exp i\omega/2$; because of the invariance of the spur of a matrix

$$a + a^* = a' + a'^* = 2\cos\omega/2,$$

or because of Eq. (80)

$$\cos\omega/2 = \cos(\beta/2)\cos(\alpha+\gamma)/2. \tag{98}$$

The class to which $D(\alpha\beta\gamma)$ belongs is thus determined uniquely by ω, and also contains $D(\omega, 0, 0)$. This result will be of value when the characters of the irreducible representations are to be calculated.

To obtain the irreducible representations, it is necessary to find invariant linear manifolds of functions of ξ and η. The simplest of these are the homogeneous polynomials of νth degree in ξ, η. They constitute a $\nu+1$ dimensional linear manifold $\mathfrak{M}_\nu$ which is subtended by the basis

$$u_p = \xi^p\eta^q \qquad (p+q=\nu,\ p=0, 1, 2\cdots\nu). \tag{99}$$

The manifold $\mathfrak{M}_\nu$ is invariant under the substitutions D, for

$$Du_p = (a\xi + b\eta)^p(-b^*\xi + a^*\eta)^q \tag{100}$$

is again a homogeneous polynomial of νth degree. In particular,

$$D(\omega, 0, 0)u_p = e^{i(p-q)\omega/2}u_p = e^{i(2p-\nu)\omega/2}u_p. \tag{101}$$

The manifold $\mathfrak{M}_\nu$ is irreducible; for suppose it contains an invariant submanifold $\mathfrak{M}_f$ and that

$$f = \sum_{p=0}^{\nu} a_p u_p$$

is a member of this submanifold. Then

$$D(\omega, 0, 0)f = \sum_{p=0}^{\nu} a_p e^{i(2p-\nu)\omega/2}u_p$$

is also a member, no matter what value ω may be given; hence the functions a_pu_p are individually members of $\mathfrak{M}_f$. The submanifold $\mathfrak{M}_f$ contains every u_p for which $a_p \neq 0$ in the polynomial f. Therefore there is certainly one value of p for which

$$D(0, \pi/2, 0)u_p = \left(\frac{\xi+\eta}{\sqrt{2}}\right)^p\left(\frac{-\xi+\eta}{\sqrt{2}}\right)^q$$

is in $\mathfrak{M}_f$; this is a polynomial in which the coefficient of $u_\nu = \xi^\nu$ is certainly not zero. Hence $\mathfrak{M}_f$ contains u_ν and therefore also

$$D(0, \pi/2, 0)u_\nu = \left(\frac{\xi+\eta}{\sqrt{2}}\right)^\nu,$$

which is a polynomial of νth degree *none of whose coefficients vanish.* Hence $\mathfrak{M}_f$ contains every one of the $\nu+1$ functions u_p and therefore coincides with

$\mathfrak{M}_\nu$. It is thus possible to find an irreducible representation for every integral value of $\nu \geqq 0$; it may be shown that there are no others.

The notation of Eq. (99) is not the most convenient for the following; let $\nu = 2l$, $p = l+m$, $q = l-m$, and let the functions

$$u^l_m = \xi^{l+m}\eta^{l-m}/[(l+m)!(l-m)!]^{1/2} \tag{102}$$

be chosen as a basis of $\mathfrak{M}_\nu$. Since ν and p take on integral values, it follows that

$$\begin{aligned} l &= 0, \tfrac{1}{2}, 1, 3/2, \cdots, \\ m &= l - p, \text{ where } p \text{ is an integer } \leqq 2l. \end{aligned} \tag{103}$$

The division brought about by the integral and half-integral values of l and m will prove of importance later. The quantity l will be used as an index for the representation:

$$D(\alpha\beta\gamma)u^l_m = {}^lD_{mn}(\alpha\beta\gamma)u^l_n \tag{104}$$

where the numbers ${}^lD_{mn}$ are defined by the algebraic equation

$$\frac{(a\xi + b\eta)^{l+m}(-b^*\xi + a^*\eta)^{l-m}}{[(l+m)!(l-m)!]^{1/2}} = \sum_{n=-l}^{l} {}^lD_{mn}\frac{\xi^{l+n}\eta^{l-n}}{[(l+n)!(l-n)!]^{1/2}}. \tag{105}$$

The reader will readily verify, for example, that the matrix ${}^1D_{mn}$ has the form

$$\begin{Vmatrix} a^2 & \sqrt{2}ab & b^2 \\ -\sqrt{2}ab^* & (aa^* - bb^*) & \sqrt{2}a^*b \\ b^{*2} & -\sqrt{2}a^*b^* & a^{*2} \end{Vmatrix}. \tag{106}$$

The matrix of Eq. (79), which was the starting point of the discussion, is properly denoted by ${}^{1/2}D_{mn}(\alpha\beta\gamma)$.

The matrices ${}^lD_{mn}$ are unitary. To prove this, it must be shown that

$$\sum_{m=-l}^{l} |\, u^l_m \,|^2$$

is invariant; but this is true, for the expression is identically equal to

$$(\xi\xi^* + \eta\eta^*)^{2l}/(2l)!;$$

hence

$${}^lD^{-1}_{mn} = {}^lD_{nm}{}^*. \tag{107}$$

The elements of these matrices, considered as functions of α, β, γ, have numerous properties which are readily deduced from their definition.[29] The most fundamental of these is (cf. Eq. (52) for notation)

$${}^lD_{mn}(\alpha''\beta''\gamma'') = {}^lD_{mk}(\alpha'\beta'\gamma'){}^lD_{kn}(\alpha\beta\gamma) \tag{108}$$

which expresses the fact that they represent the group D. Since the representation is irreducible, it follows from Eqs. (59) and (96) that

[29] They may also be obtained as solutions of a partial differential equation.

$$\int_{\alpha=0}^{4\pi}\int_{\beta=0}^{\pi}\int_{\gamma=0}^{2\pi} {}^{l}D_{mn}{}^{\lambda}D^{*}_{\mu\nu}\sigma d\alpha d\beta d\gamma = \frac{N}{2l+1}\delta_{l\lambda}\delta_{m\mu}\delta_{n\nu} \tag{109}$$

which is a most important orthogonality relation. The function σ can be determined from the considerations of Section 7, and is found to be

$$\sigma = (\sin\beta)/(4\pi)^2. \tag{110}$$

When σ is normalized with the factor $1/(4\pi)^2$, the value of N is unity. The characters are also readily obtained:

$${}^{l}\chi(\omega) = \sum_{m=-l}^{l} {}^{l}D_{mm}(\omega, 0, 0) = \sum_{m=-l}^{l} e^{im\omega} = [\sin(2l+1)\omega/2]/[\sin\omega/2]. \tag{111}$$

They satisfy the orthogonality relations:

$$\int_{\alpha=0}^{4\pi}\int_{\beta=0}^{\pi}\int_{\gamma=0}^{2\pi} {}^{l}\chi(\omega)^{\lambda}\chi(\omega)\sigma d\alpha d\beta d\gamma = \delta_{l\lambda}, \tag{112}$$

in which ω is to be expressed in terms of $\alpha\beta\gamma$ by means of Eq. (91). Two of the integrations of this triple integral may be performed, leaving

$$\int_{0}^{4\pi} {}^{l}\chi(\omega)^{\lambda}\chi(\omega)\rho(\omega)d\omega = \delta_{l\lambda} \tag{113}$$

where $\rho(\omega)$ is to be determined by integrating σ over all values of $\alpha\beta\gamma$ for which ω has a given value. The reader may verify that

$$\rho(\omega) = (\sin^2\omega/2)/4\pi. \tag{114}$$

On substituting Eq. (80) into Eq. (105), it is seen that

$${}^{l}D_{mn}(\alpha\beta\gamma) = e^{im\gamma}P^{l}{}_{mn}(\beta)e^{in\alpha}, \tag{115}$$

in which P^{l}_{mn} is independent of α and γ. These functions are generalizations of the spherical harmonics. This may be shown by setting $m=0$, $\xi=\eta=1$ in Eq. (105); it then reduces to

$$[\cos\beta + i\sin\beta\sin\alpha]^{l}/l! = \sum_{n} P^{l}{}_{0n}e^{in\alpha}/[(l+n)!(l-n)!]^{1/2}. \tag{116}$$

On multiplying by $\exp(-in\alpha)$ and integrating over α from zero to 2π, it is seen that

$$\frac{l!}{[(l+n)!(l-n)!]^{1/2}}P^{l}{}_{0n}(\beta) = \frac{1}{2\pi}\int_{0}^{2\pi}[\cos\beta + i\sin\beta\sin\alpha]^{l}e^{-in\alpha}d\alpha, \tag{117}$$

which is Laplace's integral for the tesseral harmonics.

19. **The three dimensional rotation group.** The rotation group may be defined analytically by the linear substitutions of Eqs. (73) and (74), the coefficients of which may be expressed in terms of the three Eulerian angles, $\alpha\beta\gamma$. These have a simple geometrical significance: consider a

sphere whose center is at the origin, and on which a quadrant AB is marked By the original orientation of the sphere is meant that in which AB occupies the position————(Fig. 3); when the sphere is rotated, AB will be moved to some position– – –. Let a point P be rigidly attached to the sphere, so that its coordinates in the original orientation are x, y, z, in the rotated orientation, x', y', z'. Then these two sets of coordinates will be related by an equation like (73).

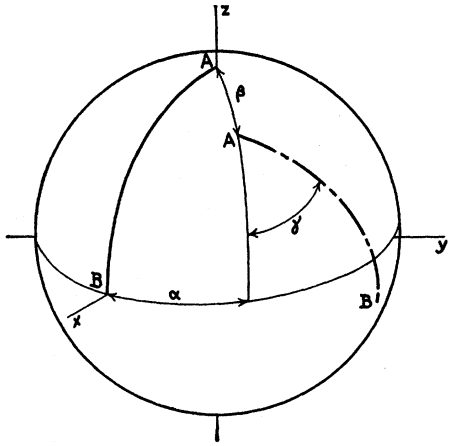


Fig. 3

The most general rotation of the sphere may be resolved into a succession of rotations about the y and z axes (which are fixed in space). Denote a rotation about the z axis through an angle α by $S(\alpha)$, a rotation about the y-axis through an angle β by $T(\beta)$. The angles are to be reckoned positive in the counter-clockwise direction when viewed from a point on the positive axis in question. The substitutions corresponding to these rotations are

$$S(\alpha): \quad \begin{aligned} x &= x' \cos\alpha + y' \sin\alpha \\ y &= -x' \sin\alpha + y' \cos\alpha \\ z &= z'; \end{aligned} \tag{118}$$

$$T(\beta): \quad \begin{aligned} x &= x' \cos\beta - z' \sin\beta \\ y &= y' \\ z &= x' \sin\beta + z' \cos\beta. \end{aligned} \tag{119}$$

Then it is obvious from the figure that the most general rotation is obtained if the sphere be rotated through $S(\gamma)$, then through $T(\beta)$, and finally through $S(\alpha)$:

$$R(\alpha\beta\gamma) = S(\gamma)T(\beta)S(\alpha). \tag{120}$$

The three angles thus specified are the Eulerian angles of the rotation. If they be restricted by the inequalities

$$0 \leqq \alpha < 2\pi, \quad 0 \leqq \beta < \pi, \quad 0 \leqq \gamma < 2\pi, \tag{121}$$

they are uniquely determined by the rotation, and conversely. The Eqs. (120) and (121) are very analogous to Eqs. (95) and (96) for the group of unitary substitutions, the only fundamental difference being that α in (96) is restricted only to lie between 0 and 4π, while in (121), it must lie between 0 and 2π.

The relationship between the two groups[30] which was mentioned at the beginning of Section 18 can now be made clear in a very simple manner: consider the three real functions of the two complex variables ξ, η

$$\begin{aligned} x &= \xi^*\eta + \xi\eta^* \\ y &= \frac{1}{i}(\xi^*\eta - \xi\eta^*) \\ z &= \xi\xi^* - \eta\eta^*. \end{aligned} \tag{122}$$

The identity $x^2+y^2+z^2=(\xi\xi^*+\eta\eta^*)^2$ will readily be verified. Now let ξ, η be subjected to the substitution $D(\alpha\beta\gamma)$ of Eq. (91), and x', y', z' calculated according to Eq. (122) from ξ' and η'. Then x', y', z' are related to x, y, z by a certain linear transformation which is readily calculated in terms of a, b. It will prove to be precisely $R(\alpha\beta\gamma)$. This is most readily seen by noting that $D(\alpha, 0, 0)$ induces the substitution $S(\alpha)$, $D(0, \beta, 0)$ the substitution $T(\beta)$. Furthermore, if two unitary transformations, D and D', induce two substitutions R and R' of the variables xyz, then the product DD' will induce RR'. The result then follows at once from Eqs. (95) and (120). From this it follows that a multiplicative correspondence exists between the elements $R(\alpha\beta\gamma)$ of the rotation group and the elements $D(\alpha\beta\gamma)$ of the unitary substitutions.[31] The correspondence is not unique, however, for if a rotation is induced by $D(\alpha, \beta, \gamma)$ it is also induced by $D(\alpha+2\pi, \beta, \gamma)$. This is the significance of the difference between (96) and (121). The irreducible representations of the rotation group can now be picked out of the set of representations of the group D, for all the matrices ${}^lD_{mn}(\alpha\beta\gamma)$ satisfy the condition of multiplicative correspondence with $R(\alpha\beta\gamma)$, and since (by Eq. (115))

$${}^lD_{mn}(\alpha + 2\pi, \beta, \gamma) = e^{+2\pi i n}\, {}^lD_{mn}(\alpha, \beta, \gamma),$$

all representations whose index l is integral satisfy the condition of unique correspondence. These are therefore the representations of the rotation group, those for which l is a half-integer are not.

The operations $R(\alpha, \beta, \gamma)$ $D(\alpha, \beta, \gamma)$ are seen to form a group which is isomorphic to $D(\alpha, \beta, \gamma)$. They therefore admit all matrices ${}^lD_{mn}(\alpha, \beta, \gamma)$ as representations.

IV. Applications to the Wave Equation

To complete the foregoing chapter, it would be logical to treat the permutation group as completely as the rotation group, etc. This would be a

[30] Weyl, *l.c.* p. 108.

[31] From this and Eq. (98) it may be deduced that every rotation belongs to the same class as a certain rotation $S(\omega)$ about the z-axis, a result which may also be proven geometrically.

lengthy and complex discussion, and is not necessary for present purposes. The brief discussion of the permutation group which is necessary will be given in Chapter V. The present chapter is to be devoted to applications which are independent of the permutation group.

20. **The wave equation without spin terms.** Consider an atom in a field-free region. It may contain any number N of electrons, but their spin is not to be taken into account. Let the rectangular coordinates of the i-th electron be x_i, y_i, z_i, as before. Then the group $\mathfrak{G}$ of the wave-equation contains a sub-group which is composed of the rotations $R(\alpha\beta\gamma)$, the reflection K, and the products $KR=RK$. Let u_m^l be a basis for an irreducible manifold of wave-functions belonging to the energy W, then

$$\begin{aligned} Ru_m^l &= {}^lD_{mn}u_n^l, \quad l = 0, 1, 2, \cdots \\ Ku_m^l &= \pm u_m^l. \end{aligned} \tag{123}$$

The basis will be called even or odd according as the positive or negative sign applies in the last equation.

The problem is to determine as much as possible regarding $u_m{}^l$ from Eq. (123). If the atom contained only one electron, it is known that its solutions have the form[32]

$$u_m^l = X(r)P_l{}^m(\theta)e^{-im\phi} \tag{124}$$

where r, θ, ϕ are the polar coordinates of the electron. Since the ${}^lD_{mn}$ are generalizations of the spherical harmonics, the presumption is that there exists a generalization of this formula which applies to any number of electrons.[33] To discover it, we transform from the variables $x_1y_1z_1 \cdots x_N y_N z_N$ to a new set which consists of three angles, A, B, Γ, specifying the orientation of the atom and $3(N-1)$ other variables specifying the configuration of the electrons relative to each other. This transformation can be accomplished in the following manner: suppose the electrons to be at the points $x_1y_1z_1 \cdots x_Ny_Nz_N$ and imagine them rigidly fastened to one another, so that the whole system can be rotated as a rigid body about any axis passing through the origin. Let it be so rotated until the electron 1 lies on the positive z-axis, and the electron 2 in the first or second quadrant of the xz-plane. Call this rotation $R^{-1}(\mathrm{AB}\Gamma)$; the coordinates of the electrons after the rotation will be

$$0, 0, \bar{z}_1;\ \bar{x}_2, 0, \bar{z}_2;\ \bar{x}_3, \bar{y}_3, \bar{z}_3;\ \cdots\ \bar{x}_N, \bar{y}_N, \bar{z}_N;\ (\bar{z}_1 \geqq 0, \bar{x}_2 \geqq 0). \tag{125}$$

These $3(N-1)$ quantities determine the relative positions, and together with A, B, Γ, are single valued functions of the x_i, y_i, z_i. They can therefore be used as coordinates in place of the latter, and will be denoted by A, B, Γ, $\bar{x}$.

In terms of these coordinates, certain functions satisfying the first of Eqs. (123) can easily be written down. They are*

[32] E. C. Kemble, Phys. Rev. Suppl. **1**, 195 (1929), also ftn. on p. 215, *ibid.*

[33] E. Wigner. Zeits. f. Physik **43**, 640 (1927).

* The Eq. (126) may also be deduced rigorously from Eq. (61a).

$$u^l_m = \sum_{n=-l}^{l} {}^lD^{-1}_{mn}(\mathrm{A}, \mathrm{B}, \Gamma) f_n(\bar{x}) = \sum_{n=l}^{l} e^{-i(m\mathrm{A}+n\Gamma)} P^l_{mn}(B) f_n(\bar{x}), \tag{126}$$

where the $2l+1$ functions f_n are any arbitrary functions dependent only on $\bar{x}$; this is the required generalization of Eq. (124). To prove this statement, note that a rotation $R(\alpha\beta\gamma)$ does not alter the relative coordinates, but changes $\mathrm{AB}\Gamma$ into $\mathrm{A'B'}\Gamma'$, where

$$R(\alpha, \beta, \gamma) R^{-1}(\mathrm{A'}, \mathrm{B'}, \Gamma') = R^{-1}(\mathrm{A}, \mathrm{B}, \Gamma). \tag{127}$$

Then if

$$R(\alpha, \beta, \gamma,) u^l_m = v^l_m,$$
$$v^l_m(\mathrm{A'}, \mathrm{B'}, \Gamma', \bar{x}) = u^l_m(\mathrm{A}, \mathrm{B}, \Gamma, \bar{x}) = {}^lD_{mn}(\alpha, \beta, \gamma) u^l_n(\mathrm{A'}, \mathrm{B'}, \Gamma', \bar{x}),$$

by Eqs. (127) and (126).

In this, it has been tacitly assumed that $N \geqq 2$; the particular case $N=1$ requires special consideration; here the special rotation $R^{-1}(0, \theta, \phi)$ (where θ and ϕ are the polar coordinates of Eq. (124)) brings the electron into the position 0, 0, r. It is then sufficient to set

$$u^l_m = {}^lD^{-1}_{m0}(0, \theta, \phi) f(r) = e^{-im\phi} P^l{}_{m0}(\theta) f(r);$$

the reasoning proving that u^l_m satisfies Eq. (123) is then the same as before.[34]

It remains to consider the second of Eqs. (123). Were it not for the requirements that $\bar{z}_1 \geqq 0$, $\bar{x}_2 \geqq 0$ the function Ku^l_m could be obtained from u^l_m by changing the signs of the variables $(\bar{x})$ and leaving A, B, Γ unaltered. These requirements are essential, however, since otherwise A, B, Γ, $\bar{x}$ would not be uniquely determined. Under these circumstances, it is necessary to resolve K into a rotation through π radians about the $\bar{y}$-axis (note the bar!) followed by a reflection K'' in the $\bar{x}\bar{z}$-plane. That the result is the operation K is geometrically evident. Hence

$$Ku^l_m = R^{-1}(\mathrm{A}, \mathrm{B}, \Gamma) R(0, \pi, 0) K'' R(\mathrm{A}, \mathrm{B}, \Gamma) u^l_m$$
$$= {}^lD^{-1}_{mn}(\mathrm{A}, \mathrm{B}, \Gamma) {}^lD_{kn}(0, \pi, 0) f_n(\bar{z}_1, \bar{x}_2, \bar{z}_2, \bar{x}_3, -\bar{y}_3, \bar{z}_3 \cdots \bar{x}_N, -\bar{y}_N, \bar{z}_N).$$

As

$${}^lD_{kn}(0, \pi, 0) = (-1)^{l+k} \delta_{k,-n},$$

[34] The expression ${}^lD^{-1}_{mn}(0,\theta,\phi)\, f_n(r)$ is only apparently more general. This becomes obvious if it be recalled that the surface spherical harmonics form a complete orthogonal set in terms of which any single valued function of θ and ϕ may be expanded.

this is equivalent to

$$Ku^l_m = {}^lD^{-1}_{mn}(\mathrm{A}, \mathrm{B}, \Gamma)(-1)^{l+n} f_{-n}(\bar{z}_1, \bar{x}_2, \bar{z}_2, \bar{x}_3, -\bar{y}_3, \bar{z}_3 \cdots \bar{x}_N, -\bar{y}_N, \bar{z}_N).$$

Comparing this with Eq. (126), it is seen that if the arbitrary functions f_n are restricted by the relations

$$(-1)^{l+n} f_n(\bar{z}_1, \bar{x}_2, \bar{z}_2, \bar{x}_3, \bar{y}_3, \bar{z}_3 \cdots) = \pm f_{-n}(\bar{z}_1, \bar{x}_2, \bar{z}_2, \bar{x}_3, -\bar{y}_3, \bar{z}_3 \cdots),$$

then the second of Eq. (123) may be satisfied for either sign.

In case $N=1$, however, the situation is otherwise, since here $n=0$, and the equation is then

$$(-1)^l f(r) = \pm f(r),$$

which can only be satisfied if $\pm 1 = (-1)^l$. The sign is thus determined by l, is $+$ if l is even, $-$ if l is odd. This exceptional property for $N=1$ will be found to be responsible for one of the selection principles of the hydrogen-like atoms.

21. **The infinitesimal rotations and the angular momentum operators.** The analogy between Eqs. (124) and (126) will lead the reader to suspect that l and m have the same physical significance in both expressions, namely that of azimuthal and magnetic quantum members. That this is true will be shown by calculating the matrices which represent the angular momentum operators.[35] This is very conveniently accomplished by considering the infinitesimal rotations, which rotate a given point x, y, z, into another, x', y', z', which is only infinitesimally distant. It is obvious that the vector $x'-x$, $y'-y$, $z'-z$, must be perpendicular to x, y, z. There are three such rotations, one about each of the coordinate axes:

$$\begin{aligned}
R_x: \quad & x = x' \\
& y = y' + \epsilon z' \\
& z = z' - \epsilon y', \\
R_y: \quad & x = x' - \epsilon z' \\
& y = y' \\
& z = z' + \epsilon x', \\
R_z: \quad & x = x' + \epsilon y' \\
& y = y' - \epsilon x' \\
& z = z',
\end{aligned} \tag{128}$$

in which ϵ is a real infinitesimal. When one of these substitutions is performed on a function of the coordinates of the N electrons, the result is, e.g.,

[35] Weyl. *l.c.* Chapter IV a.

$$\begin{aligned} R_x\psi(x_i, y_i, z_i) &= \psi(x_i, y_i + \epsilon z_i, z_i - \epsilon y_i) \\ &= \psi(x_i, y_i, z_i) + \epsilon \sum_{i=1}^{N} \left(z_i \frac{\partial \psi}{\partial y_i} - y_i \frac{\partial \psi}{\partial z_i} \right). \end{aligned} \tag{129}$$

The differential operators appearing in the sum will be familiar to the reader as (essentially) the operators associated to the angular momenta of the various electrons. If the operators associated[36] to the total angular momentum of the atom are designated by L_x, L_y, L_z, then

$$\begin{aligned} L_x\psi &= -\frac{1}{i} \sum \left(y_i \frac{\partial \psi}{\partial z_i} - z_i \frac{\partial \psi}{\partial y_i} \right), \\ L_y\psi &= -\frac{1}{i} \sum \left(z_i \frac{\partial \psi}{\partial x_i} - x_i \frac{\partial \psi}{\partial z_i} \right), \\ L_z\psi &= -\frac{1}{i} \sum \left(x_i \frac{\partial \psi}{\partial y_i} - y_i \frac{\partial \psi}{\partial x_i} \right), \end{aligned} \tag{130}$$

and the rotations R_x, R_y, R_z become

$$\begin{aligned} R_x &= I + i\epsilon L_x, \\ R_y &= I + i\epsilon L_y, \\ R_z &= I + i\epsilon L_z. \end{aligned} \tag{131}$$

It is a simple matter to calculate the matrices associated to R_x, R_y, R_z: the Cayley-Klein parameters are

$$\begin{aligned} R_x &: a = 1, \quad b = i\epsilon/2 \\ R_y &: a = 1, \quad b = \epsilon/2 \\ R_z &: a = 1 + i\epsilon/2, \quad b = 0. \end{aligned}$$

Giving u_m^l the significance of Eq. (102) for the moment, and using Eq. (105), it is seen that

$$R_x u_m^l = u_m^l + \frac{i\epsilon}{2} \{ [(l+m)(l-m+1)]^{1/2} u_{m-1}^l + [(l-m)(l+m+1)]^{1/2} u_{m+1}^l \}$$

$$R_y u_m^l = u_m^l + \frac{\epsilon}{2} \{ [(l+m)(l-m+1)]^{1/2} u_{m-1}^l - [(l-m)(l+m+1)]^{1/2} u_{m+1}^l \}$$

$$R_z u_m^l = u_m^l (1 + im\epsilon),$$

whence

$$\begin{aligned} (L_x - iL_y) u_m^l &= [(l-m)(l+m+1)]^{1/2} u_{m+1}^l \\ (L_x + iL_y) u_m^l &= [(l+m)(l-m+1)]^{1/2} u_{m-1}^l \\ L_z u_m^l &= m u_m^l, \end{aligned} \tag{132}$$

[36] E. C. Kemble, Phys. Rev. Suppl. **1**, 194 (1929).

and

$$L^2 u_m{}^l = (L_x{}^2 + L_y{}^2 + L_z{}^2) u_m{}^l = l(l+1) u_m{}^l. \tag{132a}$$

These equations show that l and m are really the total angular momentum and its component in the z-direction, each measured in units of $h/2\pi$.

For half-integral values of l, which acquire a significance when the spin is taken into account, the angular momentum matrices may still be defined by equations similar to (131). In the case $l=1/2$, they are

$$L_x = \tfrac{1}{2}\begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} \quad L_x = \tfrac{1}{2}\begin{Vmatrix} 0 & -i \\ i & 0 \end{Vmatrix}, \quad L_z = \tfrac{1}{2}\begin{Vmatrix} 1 & 0 \\ 0 & -1 \end{Vmatrix} \tag{133}$$

which are essentially the matrices which form the starting point of the Pauli spin theory.

22. **The vectorial addition of angular momenta as the reduction of a given manifold.** As a first example of the reduction of a given manifold, the case of a dynamical system which can be considered as composed of two parts loosely coupled together may be considered. This problem was treated with a high degree of success by Landé, who assigned an angular momentum to each part of the system. These were quantized and added together vectorially. The additional quantum condition that the resultant must also be integral or half integral was then introduced. This theory resulted in a rational classification of complex spectra. It will be shown that the Landé rules are obtained without difficulty by an application of the preceding methods.

The wave-equation of the system is in this case

$$(H_1 + H_2 + \epsilon H_{12})\psi = W\psi, \tag{134}$$

where the operators H_1 and H_2 are functions of different sets of coordinates. The solutions of

$$H_1 u = W_1 u$$

and

$$H_2 v = W_2 v \tag{135}$$

are assumed to be known, and each of the Eqs. (134) and (135) are supposed to admit the operations of the rotation group as integrals. For a given value of W_1 there will then be $2l+1$ solutions $u_m{}^l$, and for W_2, $2\lambda+1$ solutions $v_\mu{}^\lambda$, each set of which is the basis of a manifold invariant under the rotations. There will be six angular momentum operators L_x, L_y, L_z and Λ_x, Λ_y, Λ_z; the total angular momenta are $J_x = L_x + \Lambda_x$, etc. It is to be noted that the L's and Λ's commute, and that L^2 and Λ^2 commute with J_x, J_y, J_z. These last remarks are not essential to the following considerations, but it will be convenient to use them in interpreting the results.

The zero approximations to the solutions of Eq. (134) will be linear combinations of the functions $u_m{}^l v_\mu{}^\lambda$. These products subtend a manifold of $(2l+1)$ $(2\lambda+1)$ dimensions which is obviously invariant under the rotation group, for

$$Ru_m{}^l v_\mu{}^\lambda = {}^lD_{mn}{}^\lambda D_{\mu\nu} u_m{}^l v_\nu{}^\lambda. \tag{136}$$

The matrices ${}^lD_{mn}\ {}^\lambda D_{\mu\nu}$ (whose rows are designated by the two indices m and μ, and whose columns by n and ν) are thus a (reducible) representation of the rotation group. The problem of reducing this representation will be completely solved when $(2l+1)$ $(2\lambda+1)$ functions

$$w_M^J = \sum_m \sum_\mu A^{J\,l\lambda}_{M\,m\mu} u_m{}^l v_\mu{}^\lambda \tag{137}$$

are determined,[37] such that

$$Rw_M^J = {}^JD_{MN} w_N^J. \tag{138}$$

The first part of this problem[38] is to determine the values of J for which such functions exist, and the number of independent ones for each J and M. The theory of characters (Section 10) enables these to be determined at once: the character of the representation (136) is

$$\chi(\omega) = \sum_{m=-l}^{l} \sum_{\mu=-\lambda}^{\lambda} e^{i(m+\mu)\omega}$$

(cf. Eq. (111)), which may be transformed into

$$\begin{aligned} \chi(\omega) &= \sum_{J=|l-\lambda|}^{l+\lambda} \sum_{M=-J}^{J} e^{iM\omega} \\ &= \sum_{J=|l-\lambda|}^{l+\lambda} {}^J\chi(\omega) \end{aligned} \tag{139}$$

where ${}^J\chi(\omega)$ are the primitive characters of the rotation group. Hence there is one and only one w_M^J for each of the values $J=l+\lambda,\ J=l+\lambda-1, \cdots J=|l-\lambda|$ and none for any other values of J. This coincides with the result obtained by Landé.

It remains to determine the coefficients $A_{Mm\mu}{}^{J\,l\lambda}$ for these various values of J amd M. It is obvious that if this is accomplished for a particular set of functions $u_m{}^l\ v_\mu{}^\lambda$ satisfying Eq. (136), the same coefficients will apply to any other set satisfying this equation. For this reason it is permissable to take $u_m{}^l = \xi_1{}^{l+m}\eta_1{}^{l-m}/c(l,m)$ and $v_\mu{}^\lambda = \xi_2{}^{\lambda+\mu}\eta_2{}^{\lambda-\mu}/c(\lambda,\ \mu)$ where $\xi_1\eta_1$ and $\xi_2\eta_2$ are two pairs of variables which are both subjected to the substitution (79) and $c(l,\ m) = \sqrt{(l+m)!\ (l-m)!}$ The functions w_M^J will then be homogeneous polynomials in these four variables, of degree $2l$ in $\xi_1\eta_1$ and 2λ in $\xi_2\eta_2$. If such polynomials can be found which satisfy Eq. (138) for a given value of J, their coefficients will be, essentially, the $A_{Mm\mu}{}^{J\,l\lambda}$.

Now, if x is any arbitrary constant, the polynomials

$$f_M{}^J = (\xi_1 + x\xi_2)^{J+M}(\eta_1 + x\eta_2)^{J-M}/c(J,\ M) \tag{140}$$

[37] The coefficients A $\cdots$ have so many indices that it is expedient to exempt them all from the summation convention and to indicate all sums explicitly.

[38] The remaining considerations of this paragraph are purely mathematical and do not depend on J, l, λ being integers. They will thus still be applicable when the spin of the electron is taken into account (cf. Section 28) and R is replaced by D or RD.

obviously satisfy Eq. (138) for $\xi=\xi_1+x\xi_2$ and $\eta=\eta_1+x\eta_2$ undergo the same substitution as ξ_1, η_1 and ξ_2, η_2. They may be expanded in powers of x, and since x is arbitrary, the coefficients of x^p must also satisfy Eq. (138). Call these $f_M{}^{Jp}$; they are polynomials of degree $2J-p$ in $\xi_1\eta_1$ and degree p in $\xi_2\eta_2$. Furthermore, it is readily proven that

$$D(\alpha,\beta,\gamma)(\xi_1\eta_2-\xi_2\eta_1)^q=(\xi_1\eta_2-\xi_2\eta_1)^q; \tag{141}$$

hence the functions

$$w_M{}^J=Cf_M{}^{Jp}(\xi_1\eta_2-\xi_2\eta_1)^q, \tag{142}$$

where C is a constant independent of M, must satisfy Eq. (138); if q is a positive integer or zero, $w^J{}_M$ is a polynomial in the four variables and of degree $2J-p+q$ in $\xi_1\eta_1$ and $p+q$ in $\xi_2\eta_2$. If

$$p=J-(l-\lambda)$$
$$q=(l+\lambda)-J$$

it is therefore the polynomial required. Its coefficients can readily be calculated from Eqs. (140) and (142), and the A's from the equation

$$w_M{}^J=\sum_m\sum_\mu A^{Jl\lambda}_{Mm\mu}\frac{\xi_1{}^{l+m}\eta_1{}^{l-m}}{c(l,m)}\frac{\xi_2{}^{\lambda+\mu}\eta_2{}^{\lambda-\mu}}{c(\lambda,\mu)}. \tag{143}$$

This procedure yields precisely one set of A's for each value of J indicated by Eq. (139). The reduction of the basis $u^l_m v_\mu{}^\lambda$ is thus completed.[39]

As an example, let $\lambda=1$, $l>0$ then $J=l+1$, l, or $l-1$. For $J=l-1$, $p=0$, $q=2$ and

$$f_m{}^{l-1,0}=\xi_1{}^{l+m-1}\eta_1{}^{l-m-1}/c(l-1,m),$$

and therefore

$$w_m{}^{l-1}=C\xi_1{}^{l+m-1}\eta_1{}^{l-m-1}(\xi_1\eta_2-\xi_2\eta_1)^2/c(l-1,m),$$

or

$$w_m{}^{l-1}=\left[\frac{(l+m+1)(l+m)}{2l(2l+1)}\right]^{1/2}u^l{}_{m+1}v^1{}_{-1}-\left[\frac{(l+m)(l-m)}{l(2l+1)}\right]^{1/2}u^l_m v_0{}^1 + \left[\frac{(l-m)(l-m+1)}{2l(2l+1)}\right]^{1/2}u^l{}_{m-1}v_1{}^1. \tag{144a}$$

The arbitrary constant C has been chosen so that the sum of the squares of the coefficients is unity (normalization). For $J=l$, $p=1$ and $q=1$, while for $J=l+1$, $p=2$ and $q=0$. The reader will verify that

[39] This procedure is used by Weyl, *l.c.* p. 159.

$$w_m{}^l = \left[\frac{(l-m)(l+m+1)}{2l(l+1)}\right]^{1/2} u^l{}_{m+1} v_{-1}{}^1 + \frac{m}{[l(l+1)]^{1/2}} u^l{}_m v_0{}^1 - \left[\frac{(l+m)(l-m+1)}{2l(l+1)}\right]^{1/2} u_m{}^l{}_{-1} v_1{}^1$$

$$w_m{}^{l+1} = \left[\frac{(l-m+1)(l-m)}{2(l+1)(2l+1)}\right]^{1/2} u^l{}_{m+1} v_{-1}{}^1 + \left[\frac{(l+m+1)(l-m+1)}{(l+1)(2l+1)}\right]^{1/2} u^l{}_m v_0{}^1 + \left[\frac{(l+m+1)(l+m)}{2(l+1)(2l+1)}\right]^{1/2} u^l{}_{m-1} v_1{}^1. \tag{144b}$$

In the exceptional case, $l=0$, which was excluded above, the result is obviously $w_m{}^1 = u^0 v_m{}^1$.

The A's have numerous properties, which may be deduced from their definition. Since ${}^J D_{MN}$ is a unitary representation of the group, it follows from Section 11, Theor. I that the constant C may always be chosen so that

$$(w_M{}^J, w_N{}^K) = \delta_{MN}\delta_{JK};$$

it may also be assumed that

$$(u^l{}_m v_\mu{}^\lambda, u^l{}_n v_\nu{}^\lambda) = \delta_{mn}\delta_{\mu\nu}.$$

It then follows from Eq. (137) that

$$\sum_m \sum_\mu A^{Jl\lambda}_{Mm\mu} A^{Kl\lambda}_{Nm\mu} = \delta_{NM}\delta_{JK}, \qquad \sum_J \sum_M A^{Jl\lambda}_{Mm\mu} A^{Jl\lambda}_{Mn\nu} = \delta_{mn}\delta_{\mu\nu}; \tag{145}$$

the transformation (137) is thus orthogonal.

A large number of the coefficients are zero. Since ${}^l D_{mn}(\omega, 0, 0) = e^{im\omega}\delta_{nm}$, etc., it follows from Eq. (137) and (138) that

$$A^{Jl\lambda}_{Mm\mu} = 0 \text{ unless } m + \mu = M. \tag{146}$$

From this, it follows: that $w_{l+\lambda}{}^{l+\lambda}$ contains only one term, viz., $u^l_l v_\lambda{}^\lambda$; that $w^J_{l+\lambda-1}$ contains only two (for $J = l+\lambda$, and $l+\lambda-1$; it vanishes for all other J); etc. etc. This result can most conveniently be summarized by tabulating the A's in matrix array (Fig. 4). Each row of this matrix is characterised by the two indices M and J; they may conveniently be arranged so that rows having the same M are adjacent and $M-1$ is lower than M in the series. Inside these sets of rows, J varies from $l+\lambda$ to M, say again in descending order. Each column of the matrix is also determined by two indices, m and μ (for l and λ are fixed). These columns may be arranged so that all having the same $m+\mu = M$ are adjacent, and $M-1$ is farther to the right than M. Within these sets, $m-1$ may always be placed farther to the right than m. When the A's are arrayed in this manner, all entries not in the outlined squares of the figure are zero; the figure is drawn for $l=2$, $\lambda=1$. The first

square on the diagonal contains only one entry; the second, two rows and two columns; the third three rows and three columns. This increase by one continues until the square contains either $2l+1$ or $2\lambda+1$ rows and columns, then the size remains constant for a time, after which a decrease by one sets in, in such a way that the last square contains just one entry again. Since the matrix as a whole is orthogonal (Eq. (138)), it follows that each of the smaller matrices into which it reduces is also orthogonal. This result is of importance for the calculation of intensities and energy values by the method of sums.[40]

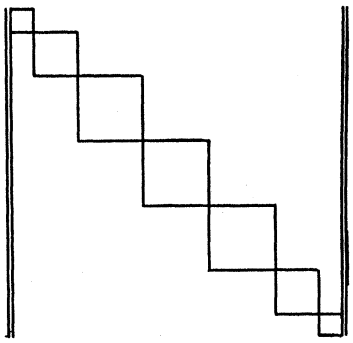


Fig. 4

From Eq. (137) and the properties of the operators L^2, Λ^2, J^2, etc., it follows that

$$\begin{aligned} L^2 w_M{}^J &= l(l+1) w_M{}^J, \\ \Lambda^2 w_M{}^J &= \lambda(\lambda+1) w_M{}^J, \\ J^2 w_M{}^J &= J(J+1) w_M{}^J, \\ J_z w_M{}^J &= M w_M{}^J. \end{aligned} \tag{147}$$

The matrices associated to these four operators by the functions $w_M{}^J$ are thus diagonal, a result which is often of importance. Interpreted in terms of the Landé theory, it means that the magnitudes of the vectors L and Λ remain constant in time (to this approximation) as does the magnitude of their resultant J. The separate components of the resultant are also a constant, but this is not true of the components of L and Λ. These vectors are to be considered as precessing about J as an axis. Since

$$L_x{}^2 = (J_x - \Lambda_x)^2, \text{ etc.},$$

it follows that

$$L^2 = J^2 - 2(J_x\Lambda_x + J_y\Lambda_y + J_z\Lambda_z) + \Lambda^2$$

whence it follows from (147) that

$$2(J_x\Lambda_x + J_y\Lambda_y + J_z\Lambda_z) w_M{}^J = [J(J+1) + \lambda(\lambda+1) - l(l+1)] w_M{}^J \tag{148}$$

[40] Cf. Sections 23 and 31.

which shows that the matrix associated to this operator is also diagonal. This means that the projection of Λ on J is a constant, in accordance with the last remark above. Similarly

$$2(L_x\Lambda_x + L_y\Lambda_y + L_z\Lambda_z)w_M{}^J = [l(l+1) + \lambda(\lambda+1) - J(J+1)]w_M{}^J. \quad (148a)$$

If the functions $u_m{}^l$ subtend the whole of the manifold of W_1 and $v_\mu{}^\lambda$ the whole manifold of W_2, so that these energy levels are just $(2l+1)$- and $(2\lambda+1)$-fold degenerate, the perturbation matrix

$$(H_{12}w_M{}^J, w_N{}^K)$$

is diagonal also. This follows at once from the considerations of Section 6, since there is no accidental degeneracy in J. If further degeneracy is present, this is not necessarily the case, but the non-diagonal elements will be relatively few in number and arranged as explained in that section.

23. **The method of sums.** The foregoing methods are capable of extension to the case of a given manifold subtended by a basis whose functions are products of, say N factors each of the form $u_m{}^l$. The manifold is certainly invariant under the rotation group and its basis can therefore be reduced to the form $w_M{}^{L\alpha}$ $(\alpha = 1 \cdots q(L))$. The values of L and the range of the index α which distinguishes different irreducible sub-bases of the same character may both be determined from the character of the representation. A slight generalization of the method of Eqs. (140)–(143) furnishes the coefficients of the transformation.

These calculations will not be carried out in detail, however, but another method which often enables the calculation of the characteristic values of the perturbation matrix without solving the secular equation will be explained.[41] Let the functions of the given basis be $v_m{}^k$ where m is the sum of the magnetic quantum numbers of the N factors and the index k is assigned in any manner which serves to distinguish functions having the same m. In other words, the explicit form of the functions is not to be utilized in the following considerations, but only the two properties:

1. The manifold subtended by $v_m{}^k$ is invariant under the group $R(\alpha\beta\gamma)$;

2. $$R(\omega, 0, 0)v_m{}^k = e^{im\omega}v_m{}^k.$$

Then, because of (1), it is known that a transformation matrix can be found such that

$$w_M{}^{L\alpha} = a_{ML\alpha;km}v_m{}^k, \qquad (\alpha = 1 \cdots q(L)),$$

where

$$Rw_M{}^{L\alpha} = {}^LD_{MN}w_N{}^{L\alpha}.$$

From this and (2), it follows that $w_M{}^{L\alpha}$ can be a function only of those $v_m{}^k$ for which $m = M$:

$$w_M{}^{L\alpha} = a_{L\alpha;k}(M)v_M{}^k. \quad (149)$$

[41] J. C. Slater, Phys. Rev. **34**, 1293 (1929).

In this equation, L must necessarily be greater than or at most equal to M, but not greater than the greatest value of m; whatever other restrictions it may be subject to need not be considered.

Now, let $H_{km;jn}$ be the perturbation matrix in terms of the $v_m{}^k$, and let it be known to commute with the representation of the rotation group. Then it follows from the property (2) that

$$H_{km;jn} = H_{kj}(m)\delta_{mn}. \tag{150}$$

On transforming this with the matrix $a_{L\alpha;k}(M)$, it must assume the form[42]

$$H'_{\alpha\beta}(L)\delta_{LK}\delta_{MN}$$

where $H'_{\alpha\beta}(L)$ is independent of M. It will be no loss of generality if we assume

$$H'_{\alpha\beta}(L) = H(L, \alpha)\delta_{\alpha\beta}, \tag{151}$$

for if this is not the case, a second canonic transformation, obtained by solving the perturbation equations, will always bring it to this form. The quantities $H(L, \alpha)$ are the characteristic values which are to be calculated. Now, for any given M, the matrix $H(L,\alpha)\ \delta\alpha\beta\delta_{Lk}$ $(L \geqq M)$ must be equivalent to the matrix $H_{kj}(M)$ because of (149). Because of the invariance of the diagonal sum, therefore,

$$\sum_{L\geqq M}\ \sum_{\alpha=1}^{q(L)} H(L, \alpha) = \sum_{k}\ H_{kk}(M). \tag{152}$$

Since $H(L, \alpha)$ is independent of M, the result of subtracting the analogous equation for $M+1$ from Eq. (152) (and then replacing M by L) is

$$\sum_{\alpha=1}^{q(L)} H(L, \alpha) = \sum_{k} H_{kk}(L) - \sum_{k} H_{kk}(L+1). \tag{153}$$

This equation permits the calculation of the sum of all energy levels associated to a given value of L. In two particular cases, viz.: when $q(L)=1$ or when the dependence of $H(L\alpha)$ on α is known from other considerations, it permits the calculation of the energy levels themselves *without the solution of the secular equation.* The importance of the method is derived from the fact that a large number of actual cases are included under the two special cases just mentioned.

24. **Selection principles.** The quantities $A_{Mm\mu}{}^{Jl\lambda}$ also govern the selection principles and intensities of spectral lines. For the first, consider an atom whose characteristic functions for the energy W are $u_m{}^l$, for the energy W' are[43] $u'_\mu{}^\lambda$. Then a virtual oscillator is associated to the transition $W, l, m \to W', \lambda, \mu$ whose vector amplitude is given by the three quantities

[42] It is suggested that the reader compare these considerations with the diagrams of Section 6.

[43] These are now characteristic functions of the *same* system, not as in Section 22, of *different* partial systems.

$$\begin{matrix}(Xu_m{}^l, u'_\mu{}^\lambda)\\ (Yu_m{}^l, u'_\mu{}^\lambda)\\ (Zu_m{}^l, u'_\mu{}^\lambda)\end{matrix} \tag{154}$$

where

$$X = \sum_i x_i, \quad Y = \sum_i y_i, \quad Z = \sum_i z_i \tag{155}$$

The expressions (154) can be partially evaluated in terms of the $A_{Mm\mu}{}^{J\,1}$. To see this, note that the three functions

$$\begin{aligned} v_1{}^1 &= \frac{1}{2^{1/2}}(-X + iY)\\ v_0{}^1 &= Z\\ v_{-1}{}^1 &= \frac{1}{2^{1/2}}(X + iY), \end{aligned} \tag{156}$$

satisfy the relation

$$Rv_p{}^1 = {}^1D_{pq}v_q{}^1. \tag{157}$$

The expressions (154) will be known if

$$(v^1{}_p u_m{}^l, u'_\mu{}^\lambda)$$

is known for each p. Now, by means of Eq. (144), functions $w_M{}^J$ can be calculated in terms of the products $v^1{}_p u^l{}_m$, and conversely, the products can be expressed in terms of $w_M{}^J$:

$$v_p{}^1 u_m{}^l = \sum_{J=l-1}^{l+1} A^{Jl1}_{m+p\ m\ p} w^J_{m+p}.$$

Hence

$$(v^1{}_p u^l{}_m, u_\mu{}'^\lambda) = \sum_{J=l-1}^{l+1} A^{Jl1}_{m+p\ m\ p}(w^J_{m+p}, u'_\mu{}^\lambda)$$

As was shown in Section 11,

$$(w_M{}^J, u'_\mu{}^\lambda) = a(J)\delta_{J\lambda}\delta_{M\mu},$$

where a is independent of μ. Hence

$$(v'_p u_m{}^l, u'_\mu{}^\lambda) = a(l\lambda)A^{\lambda l1}_{\mu m p}$$

if $\mu = m+p$, $|l-\lambda| \leqq 1$, and zero otherwise.

This result does not depend at all on the explicit form of the functions X, Y, Z, but merely on the fact that they are components of a vector and, as such, experience the same substitution as the coordinates of a point. They might even have been vector operators, i.e., operators which convert a scalar function into a vector. The matrices representing any vector quantity

all have the following elements, which are readily determined by picking out the proper A's from Eq. (144):

$$\lambda = l+1$$

$$\mu = m+1,\ \frac{1}{2^{1/2}}(-X+iY):\ a(l,l+1)\left[\frac{(l+m+2)(l+m+1)}{2(l+1)(2l+1)}\right]^{1/2};$$

$$\mu = m,\qquad Z:\ a(l,l+1)\left[\frac{(l+m+1)(l-m+1)}{(l+1)(2l+1)}\right]^{1/2}; \tag{159}$$

$$\mu = m-1,\ \frac{1}{2^{1/2}}(X+iY):\ a(l,l+1)\left[\frac{(l-m+2)(l-m+1)}{2(l+1)(2l+1)}\right]^{1/2};$$

$$\lambda = l$$

$$\mu = m+1,\ \frac{1}{2^{1/2}}(-X+iY):\ -a(l,l)\left[\frac{(l+m+1)(l-m)}{2l(l+1)}\right]^{1/2};$$

$$\mu = m,\qquad Z:\ a(l,l)\frac{m}{[l(l+1)]^{1/2}}; \tag{160}$$

$$\mu = m-1,\ \frac{1}{2^{1/2}}(X+iY):\ a(l,l)\left[\frac{(l-m+1)(l+m)}{2l(l+1)}\right]^{1/2};$$

$$\lambda = l-1$$

$$\mu = m+1,\ \frac{1}{2^{1/2}}(-X-iY):\ a(l,l-1)\left[\frac{(l-m-1)(l-n)}{2l(2l+1)}\right]^{1/2};$$

$$\mu = m,\qquad Z:\ -a(l,l-1)\left[\frac{(l+m)(l-m)}{l(2l+1)}\right]^{1/2}; \tag{161}$$

$$\mu = m-1,\ \frac{1}{2^{1/2}}(X+iY):\ a(l,l-1)\left[\frac{(l+m)(l+m-1)}{2l(2l+1)}\right]^{1/2};$$

all elements not explicitly indicated are zero. The angular momentum operators are vectors, and their matrices have this form, with $a(l,\,l+1)=a(l,\,l-1)=0$ and $a(l,\,l)=[l(l+1)]^{1/2}$ (cf. Eq. (132)).

The probability of the transition $W,\ l,\ m\to W',\ \lambda,\ \mu$ is proportional to the sum of the absolute squares of the matrix elements XYZ associated to it. The reader will readily verify that the sum of the probabilities for a transition from $W,\ l,\ m$ to $W',\ \lambda,\ m+1$, $W',\ \lambda,\ m$, and $W',\ \lambda,\ m-1$ is given by

$$|\,a(l,\lambda)\,|^2\,\frac{2\lambda+1}{2l+1} \tag{162a}$$

and is therefore *independent of* m. A similar formula holds for the sum of the transition probabilities from $W,\ l,\ \mu-1$, $W,\ l,\ \mu$, and $W,\ l,\ \mu+1$ to $W',\ \lambda,\ \mu$: it is equal to

$$|\,a(l,\lambda)\,|^2. \tag{162b}$$

The Eqs. (159), (160), and (161) embody the selection principles for the total angular momentum l and its component along the z-axis m. There is a further selection principle arising from the presence of the operation K in the group $\mathfrak{G}$: since X, Y, Z are odd functions, $(KX = -X)$, it follows that

$$(v^{l}{}_{p}u_{m}{}^{l}, u'_{\mu}{}^{\lambda}) = -(v^{l}{}_{p}Ku_{m}{}^{l}, Ku'_{\mu}{}^{\lambda});$$

hence the expression will vanish if $u_m{}^l$ and $u'_\mu{}^\lambda$ are both odd or both even. The result is Laporte's rule: *even terms combine only with odd terms and odd only with even.* In the case of the hydrogen atom, all terms with even l are even, and all terms with odd l are odd (cf. Section 20). The result is that $a(l, l) = 0$ in Eq. (160) so that $|\lambda - l| = 1$ for the hydrogen atom. How this rule applies in the case of more than one electron, will be seen in a later section.

It is to be noted that all elements (160) vanish for $l = \lambda = 0$, corresponding to the well-known special case of the selection principles.

25. **The addition of spin-vectors as the reduction of a manifold. The alternation rule for even and odd multiplets.** It is now necessary to consider the influence of the spin on the wave functions of an atom. The wave equation may be assumed to have the form

$$\{H(x_i y_i z_i) + H'(x_i y_i z_i s_i)\}\psi = W\psi. \tag{163}$$

It is indicated by experiment that the operator H' may offen be treated as a perturbation. If it be entirely neglected, the wave-function is

$$\psi(x_i, y_i, z_i, s_i) = u(x_i, y_i, z_i)\Xi(s_i) \tag{164}$$

where

$$Hu(x_i, y_i, z_i) = Wu(x_i, y_i, z_i) \tag{165}$$

and $\Xi(s_i)$ is any function of the spin variables.

The Eq. (163) is invariant under the group RD, and Eq. (165) under the group R. Hence the manifold of solutions u may be assumed to be completely reduced under this group: $u = u_m{}^l(x_i, y_i, z_i)$. Furthermore, the manifold of functions $\Xi(s_i)$ is invariant under the group of operations D. This manifold is 2^N dimensional (cf. Eq. (89)) and the functions

$$\begin{aligned} &\xi(s_1)\xi(s_2)\cdots\xi(s_N) \\ &\eta(s_1)\xi(s_2)\cdots\xi(s_N) \\ &\cdots\cdots\cdots\cdots \end{aligned} \tag{166}$$

may be taken as its basis. The problem of reducing this manifold is the one to be solved before any others are considered.[44] For typographical reasons, we introduce the notation

$$\begin{aligned} \xi_i &= \xi(s_i), \\ \eta_i &= \eta(s_i); \end{aligned} \tag{167}$$

[44] J. v. Neumann and E. Wigner, Zeits. f. Physik **47**, 200 (1928).

the problem is to find a polynomial $\Xi_\mu{}^S$ in the $2N$ quantities ξ_i, η_i, which shall be of first degree in each pair having a given index, and satisfy the equation

$$D\Xi_\mu{}^S = {}^SD_{\mu\nu}\Xi_\nu{}^S. \tag{168}$$

The values of S for which this problem possesses a solution and the number of independent solutions may be determined from the theory of characters. Since

$$D(\omega, 0, 0)\xi_i = e^{i\omega/2}\xi_i,\ D(\omega, 0, 0)\eta_i = e^{-i\omega/2}\eta_i,$$

the character of the representation based on the functions (166) is

$$\chi(\omega) = (2\cos\omega/2)^N.$$

Hence $q(S\ N)$, the number of times the irreducible representation ${}^SD_{\mu\nu}$ appears in the completely reduced form of this representation is (cf. Section 10 and Eq. (113)).

$$\begin{aligned} q(S, N) &= \frac{1}{4\pi}\int_0^{4\pi} (2\cos\omega/2)^N \sin[(2S+1)\omega/2]\sin[\omega/2]d\omega \\ &= \binom{N}{N/2 - S} - \binom{N}{N/2 - S - 1}, \end{aligned} \tag{169}$$

in which ( ) is the binomial coefficient. The values of S for which the problem possesses a solution are therefore

$$S = \frac{N}{2},\ \frac{N}{2} - 1,\ \frac{N}{2} - 2 \cdots \left(\frac{1}{2}\text{ or }0\right). \tag{170}$$

The values of S are integral or half integral according as N is an even or odd integer. This will prove to be the alternation rule for the occurrence of multiplets in the periodic system of elements.

When S has one of the values (170), the problem possesses $q(S, N)$ independent solutions. This function is seen to satisfy the recursion formula

$$q(S, N) = q(S + \tfrac{1}{2}, N - 1) + q(S - \tfrac{1}{2}, N - 1), \tag{171}$$

which corresponds to the earlier ideas regarding the addition of spin vectors: for $N=2$, the two vectors may either be parallel or antiparallel, resulting in $S=1$ or $S=0$; for $N=3$, $S=3/2$ may be realized in one way: by adding the third s parallel to two which are already parallel, while the value $S=1/2$ may be realized in either of two ways: by adding the third antiparallel to two which are parallel, or by adding it to two in the antiparallel orientation. If this process of successive addition be followed through, it will be found that the Eq. (171) is always satisfied. The values of the function $q(S, N)$ are given in Table I.

Polynomials satisfying Eq. (168) are readily constructed: if $a_1, a_2, \cdots a_{2S-1}$, a_{2S} are any $2S$ constants, the function

$$(a_1\xi_1 + \cdots + a_{2S}\xi_{2S})^{S+\mu}(a_1\eta_1 + \cdots + a_{2S}\eta_{2S})^{S-\mu}/c(S, \mu) \tag{172}$$

TABLE I

S \ N	0	1	2	3	4	5	6	7	8	9
0	1		1		2		5		14	
1/2		1		2		5		14		42
1			1		3		9		28	
3/2				1		4		14		48
2					1		5		20	
5/2						1		6		27
3							1		7	
7/2								1		8
4									1	
9/2										1

is obviously one of such a set, as is the coefficient of $a_1\, a_2 \cdots a_{2S}$. Call this $f_{\mu S}$; it is linear in each of the $2S$ pairs $\xi_1, \eta_1; \cdots \xi_{2S}, \eta_{2S}$. A polynomial which is of first degree in all pairs is

$$\Xi_{\mu S} = f_\mu{}^S(\xi_{2S+1}\eta_{2S+2} - \xi_{2S+2}\eta_{2S+1}) \cdots (\xi_{N-1}\eta_N - \xi_N\eta_{N-1}). \qquad (17\mathfrak{C})$$

Other polynomials of this kind may be obtained by permuting the indices of the quantities ξ_i, η_i. There are $N![/(2S)!\,(N/2-S)!\,2^{N/2-S}]$ *different* polynomials which arise in this way, but only $q(S, N)$ of them can be linearly independent. These will be indicated by a third index, p, which runs from $1 \to q(S, N)$: $\Xi_\mu{}^{Sp}$.

As an example, let it be required to find the irreducible submanifolds $S=1$, for $N=4$. They are three in number, since $q(1, 4)=3$. If $\mu=1$, the expression (172) becomes

$$(a_1\xi_1 + a_2\xi_2)^2/2^{1/2}.$$

The coefficient of a_1a_2 in this is $2^{1/2}\xi_1\xi_2$ and the function

$$\Xi_1{}^{11} = 2^{1/2}\xi_1\xi_2(\xi_3\eta_4 - \xi_4\eta_3);$$

the other functions may be taken to be

$$\Xi_1{}^{12} = 2^{1/2}\xi_2\xi_3(\xi_4\eta_1 - \xi_1\eta_4),$$
$$\Xi_1{}^{13} = 2^{1/2}\xi_3\xi_4(\xi_1\eta_2 - \xi_2\eta_1).$$

For $\mu=0$, the expression (172) is

$$(a_1\xi_1 + a_2\xi_2)(a_1\eta_1 + a_2\eta_2)$$

in which the coefficient of a_1a_2 is

$$(\xi_1\eta_2 + \xi_2\eta_1);$$

hence

$$\Xi_0{}^{11} = (\xi_1\eta_2 + \xi_2\eta_1)(\xi_3\eta_4 - \xi_4\eta_3),$$
$$\Xi_0{}^{12} = (\xi_2\eta_3 + \xi_3\eta_2)(\xi_4\eta_1 - \xi_1\eta_4),$$
$$\Xi_0{}^{13} = (\xi_3\eta_4 + \xi_4\eta_3)(\xi_1\eta_2 - \xi_2\eta_1).$$

The functions for $\mu=-1$ may be obtained from those for $\mu=1$ by replacing ξ by η in the first two factors of each term.

The basis of the manifold of the zero approximations to the solutions of Eq. (163) has therefore been brought to the form

$$u^l{}_m \Xi_\mu{}^{Sp}$$

which contains $(2l+1)\ 2^N$ functions. It is invariant under the operations RD:

$$RDu_m{}^l \Xi_\mu{}^{Sp} = {}^lD_{mn}{}^S D_{\mu\nu} u_n{}^l \Xi_\nu{}^{Sp}. \tag{174}$$

The problem of completely reducing this manifold may readily be solved by the method of Section 22. If

$$w_M{}^{Jp} = \sum_m \sum_\mu A^{Jl^S}_{M m\mu} u^l{}_m \Xi_\mu{}^{Sp}$$

then

$$RDw_M{}^{Jp} = {}^JD_{MN} w_N{}^{Jp}. \tag{175}$$

which shows that the complete reduction has been accomplished. J is the inner quantum number, which represents the total angular momentum of the atom.

It remains to consider the action of the operator K; obviously

$$Kw_M{}^{Jp} = \pm\ w_M{}^{Jp}$$

according as

$$Ku_m{}^l = \pm\ u_m{}^l.$$

There is thus no general relation between the value of J and the odd or even character of $w_M{}^{Jp}$

The effect of the perturbation H' will be to remove the degeneracy in p. It will not be able to remove the degeneracy in M, so that to *every approximation*, the energy level will be characterized by a certain value of J and $2J+1$ values of M. The selection principles for J and M are readily derived, for the probability of the transition from $u_M{}^J$ to $u'_N{}^L$ is determined by $(v_\mu^1 u_M{}^J, u'_N{}^L)$ where $v_\mu{}^1$ is still given by Eq. (156) and is independent of s_i. Since $RDv_\mu{}^1 = {}^1D_{\mu\nu} v_\nu{}^1$, all the reasoning of Section 24 is still valid and the selection principles are therefore

$$\begin{gathered} \Delta J = \pm\ 1 \text{ or } 0 \\ \Delta M = \pm\ 1 \text{ or } 0 \\ \text{odd} \rightleftarrows \text{even} \\ J=0 \text{ does not combine with } J=0 \end{gathered} \tag{176}$$

and the intensities of permitted transitions are given by Eqs. (159) – (161), with appropriate changes of notation.

V. The Permutation Group and the Pauli Exclusion Principle

26. **Physical considerations regarding the necessity and possibility of an exclusion principle.** It must be accepted as one of the most fundamental conclusions of modern physics that all electrons are identical. They have no distinguishing features, as have the planets of the solar system; each has identically the same charge, the same mass,—in short, when two electrons are placed in identical environments, they are acted on by the same forces and behave in the same manner. This fact may be expressed in terms of the wave-function: according to the quantum theory, $\psi\psi^*$ is the distribution function of a statistical ensemble of dynamical systems. If each of these contains N electrons, then $\psi' = P\psi$ (cf. Section 13) represents another ensemble differing from the first only in that the electrons have exchanged places. This second ensemble must be physically indistinguishable from the first ensemble.

For the purposes of the following pages it will be very convenient to denote the four variables of the i-th electron by the single symbol i. The wave function is then $\psi(1, 2, 3, \cdots N)$ and the wave equation is

$$H(1, 2, 3, \cdots N)\psi = W\psi. \tag{177}$$

The first consequence of the remarks of the preceding paragraph is that

$$H(1, 2, 3 \cdots N) = H(2, 1, 3 \cdots N) = \text{etc., etc.} \tag{178}$$

The operator may be said to be symmetric in the variables $1, 2, 3, \cdots N$, and

$$HP = PH. \tag{179}$$

The operations P are members of the group $\mathfrak{G}$ of the Eq. (177) as we have already seen. This is all that can be deduced from Eq. (177), but the consideration that the ensemble represented by ψ' must be indistinguishable from the ensemble represented by ψ results in more. The general principles of quantum-mechanics show that it can only be true if

$$\psi' = c\psi$$

where

$$cc^* = 1. \tag{180}$$

Solutions of Eq. (177) not having this property are therefore physically inadmissable. This enables an *a priori* restriction of the representations of the permutation group which need be considered: only those which are "matrices of one row and column" are required.

It can be shown that there are only two of these, and of them, one (the so-called "symmetric") does not correspond to any of the energy levels which are actually observed in atomic spectra. Only the "antisymmetric" one will therefore come into our consideration: the other governs the theory of photons, and may also be involved molecular spectra. Despite this fact, it will occasionally be convenient to consider other representations of the permutation group.

Before proceding to a study of these, it is well to convince ourselves that this "exclusion principle" is really consistent with the general principles of quantum dynamics. Is it not possible that if an atom is in a state represented by a function satisfying (180), it may, either spontaneously or otherwise, jump to a state which does not satisfy this condition? This may be answered in the negative, for first of all there can be no forced transitions to such a state for the perturbation causing a forced transition will be represented in Eq. (177) by an additional *symmetric* operator, H′ $(1 \cdots N)$. As the forbidden state will belong to the basis of an irreducible representation which is inequivalent to the permitted state, there will be no terms in the perturbation matrix corresponding to such a transition. The same reasoning leads to the conclusion that there will be no spontaneous transitions, for the functions X, Y, Z, (Eq. (155)) are also symmetric functions. The mathematical scheme is thus entirely consistent with the imposition of the additional exclusion principle.

27. **The permutation group and its representations.** The operation of replacing the natural sequence of objects $123 \cdots N$ by the sequence $i_1 i_2 \cdots i_N$ can be considered as the result of a number of interchanges. By an interchange is meant the operation which replaces the element i by the element j, and j by i leaving all others unmoved. Thus the sequence 1342 can be obtained from 1234 by first interchanging 2 and 3, then 2 and 4. This factorization of a given permutation into interchanges can be accomplished in more than one way. The number of interchanges required for one factorization may even be different than the number required for another; however the number will always be $\left\{\begin{matrix}\text{odd}\\ \text{even}\end{matrix}\right\}$ if it is $\left\{\begin{matrix}\text{odd}\\ \text{even}\end{matrix}\right\}$ in one particular case. Thus all permutations may be divided into two classes, odd and even. This classification is the basis of the theory of determinants.[45] It is seen that the product of two odd or two even permutations is even, the product of an odd and an even permutation is an odd permutation. The two one-dimensional representations of the permutation group are now readily discovered: the symmetric one is

$$P\psi = \psi, \tag{181a}$$

the antisymmetric

$$P\psi = \epsilon(P)\psi \tag{181b}$$

where $\epsilon(P) = +1$ if P is an even permutation, and $\epsilon(P) = -1$ if P is odd.

The antisymmetric representation is the only one which will be fundamental in the following. It will be convenient to know, however, that the functions $\Xi_\mu^{S\flat}$ (Section 25) constitute a basis for an[46] irreducible representation:

[45] For a convenient summary of the theory of determinants, cf. "Mathematical Monographs," No. 3. by L. G. Weld (*Ed.* Merriman and Woodward) Wiley & Sons. (1906).

[46] The representations ${}^{S}P_{pp}$, $S=0, 1/2 \cdots$, do not constitute a complete set of representations of the permutation group; there are others which cannot be brought into this form. Cf J. v. Neumann and E. Wigner, Zeits f. Physik **49**, 80 (1928), F. Hund, Zeits. f. Physik **43**, 788 (1927).

$$P\Xi_\mu{}^{Sp} = {}^{S}P_{pq}\Xi_\mu{}^{Sq}. \tag{182}$$

That the manifold of functions subtended by $\Xi_\mu{}^{Sp}$ is invariant under the permutations is obvious from its construction; that it is an irreducible manifold will not be proven. The matrices ${}^{S}P_{pq}$ are independent of μ, but dependent on S; the first statement follows at once from the fact that $PD=DP$. The elements ${}^{S}P_{pq}$ are real numbers.

Furthermore, if P_{nm} be any representation of the P's, so is

$$'P_{nm} = \epsilon(P)P_{nm}, \tag{183}$$

as follows at once from $\epsilon(P)\,\epsilon(Q)=\epsilon(PQ)$. If P_{nm} is irreducible, so is $'P_{nm}$, as may be seen at once from Schur's lemma. The two characters are related by

$$'\chi(P) = \epsilon(P)\chi(P); \tag{184}$$

the two irreducible representations can therefore be equivalent only if $\chi(P)=0$ for all odd permutations—a case which does occur. The representation $'P_{nm}$ is said to be reciprocal to P_{nm}. These are the only theorems regarding the representations of the permutation group which are required for the following.

28. **The Pauli exclusion principle and normal multiplets.** The Pauli exclusion principle was derived from a detailed consideration of the empirical material obtained from observation of the spectra of complex atoms. It states that in a given atom, no two electrons are in the same state. Two electrons may be in states having the same principal, azimuthal, and magnetic quantum numbers, but then their spin vectors must be oppositely directed. If their spins are parallel, at least one of the three other quantum numbers must be different. No real exception to this rule has ever been observed. It may be shown that this principle is equivalent to excluding the symmetric representation and retaining the antisymmetric representation as the only one having any physical significance.

To see this, it is convenient to consider the zero approximation to the wave equation, which neglects all interaction between electrons, considering only forces exerted on them by the nucleus. To this approximation the wave equation is

$$\sum_{i=1}^{N} H(i)\psi = W\psi, \tag{185}$$

where $H(i)$ is an operator depending only on the four coordinates of the ith electron. It is assumed that the solutions and characteristic values of

$$H(1)u(1) = Wu(1) \tag{186}$$

are known. Then a solution of Eq. (185) is

$$\psi(1, 2, 3, \cdots N) = u_1(1)u_2(2)u_3(3)\cdots u_N(N) \tag{187}$$

where $u_1, u_2, \cdots u_N$ are N solutions of the Eq. (186). The energy level W

is the sum $W_1+W_2+\cdots+W_N$ of the characteristic values to which u_1, $u_2, \cdots u_N$ belong. It will, for the moment, be assumed that these are not degenerate.

Then there will be at most $N!$ independent solutions of Eq. (185) for this value of W, which are obtainable from (187) by permuting the variables entering into the various u's: thus

$$\psi(3, 6, 2 \cdots n) = u_1(3)u_2(6)u_3(2) \cdots u_N(n)$$

is also a solution. From the manifold subtended by these functions, it is possible to select that function which satisfies Eq. (181a) or (181b) by the general method of Section 9. Thus

$$v = \sum_P P\psi(1, 2, \cdots N) \tag{188}$$

satisfies Eq. (181a) and

$$w = \sum_P \epsilon(P)P\psi(1, 2, \cdots N) \tag{189}$$

satisfies Eq. (181b). This last may also be written as a determinant:

$$w = \begin{vmatrix} u_1(1) & u_2(1) & \cdots & u_N(1) \\ u_1(2) & u_2(2) & \cdots & u_N(2) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ u_1(N) & u_2(N) & \cdots & u_N(N) \end{vmatrix}. \tag{190}$$

The Pauli exclusion principle states that if $u_1(1) = u_2(1)$ the state of the atom represented by v or w can have no physical significance. This is automatically fulfilled in the case of w, for in this case the determinant (190) will have two identical rows, and hence $w=0$. The symmetric function v will still exist, however, even in this case. If the atom were ever in a state represented by a symmetric function but obeying the Pauli exclusion principle, it would be possible for it to perform a transition to a state such as has just been described which does not obey the exclusion principle. The conclusion must therefore be that the state of atom is never represented by a symmetric wave-function. This result applies immediately only to the zero-approximation but there is no question but that it is the precise and general formulation of the exclusion principle.

It is then necessary to consider the effect of the interaction of the electrons on the antisymmetric wave-function only. This problem may be attacked in several ways, the most obvious being to consider the determinant (190) as the zero approximation to the solution of the complete wave equation and to treat the interaction of the electrons as a perturbation. This method will be applied in the next section; before going to this, it is illuminating to consider the perturbation problem presented by the Eq. (163) of Section 25,

in which the electrostatic interaction of the electrons is included in the operator H and the spin variables enter only into the perturbation operator H'. The rigorous solutions of the Eq. (165) are again assumed to be known. As this equation admits the permutations as integrals, it follows that the function $u^l{}_m$, will be a member of a certain manifold of functions of $x_i y_i z_i$ (not of s_i, however) which is invariant under the permutation group. This manifold need not be one-dimensional; let its basis be $u_m{}^{lr}$, where

$$Pu_m{}^{lr} = P_{rt} u_m{}^{lt}; \tag{191}$$

the representation P_{rt} will ordinarily be irreducible; it may depend on l, but is certainly independent of m.

The zero-approximations to the solutions of Eq. (163) are linear combinations of

$$u_m{}^{lr}(x_i, y_i, z_i)\Xi_\mu{}^{sp}(\rho_i); \tag{192}$$

it is necessary to find the antisymmetric function contained in the manifold subtended by this basis.[47] Since

$$Pu_m^{lr}\Xi_\mu{}^{Sp} = P^S_{rt}P_{pq}u_m^{lt}\Xi_\mu^{Sp},$$

the character of its representation is

$$\chi(P)^S\chi(P), \text{ where } \chi(P) = \sum P_{rr}, \text{ and } {}^S\chi(P) = \sum_q {}^SP_{qq}$$

and the manifold therefore contains

$$\frac{1}{N!}\sum_P \epsilon(P)\chi(P)^S\chi(P)$$

antisymmetric functions. According to the last paragraph of Section 26, this sum will be zero unless P_{rt} is equivalent to the representation reciprocal to ${}^SP_{pq}$, and equal to unity when this condition is fulfilled. The manifold thus contains at most one antisymmetric function, given by

$$w = \sum_P \epsilon(P)P\{u_m^{lr}\Xi_\mu{}^{Sp}\}, \tag{193a}$$

(which expression is essentially independent of the indices r and p). The number S was originally introduced as the total spin angular momentum (Section 23); the exclusion principle imposes a second role on S: it also determines the character of the representation P_{rt} based on the spin-free functions $u_m{}^{lr}$.

This result is of importance for the theory of normal multiplets: to each energy level of the spin-free equation there will be associated either zero, or else precisely $(2l+1)$ $(2S+1)$ characteristic functions satisfying the Pauli exclusion principle. To the zeroth approximation these have the form

$$w_{m\mu}{}^{ls} = \sum_r\sum_P a_{rp}u_m^{lr}\Xi_\mu{}^{Sp} \tag{193b}$$

[47] Neumann and Wigner, Zeits. f. Physik **49**, 80 (1928).

where the a_{rp} are independent of m and μ. These functions are a basis for the representation ${}^{l}D_{mn}{}^{s}D_{\mu\nu}$ of the group RD; the manifold subtended by them may be reduced by the method of Section 22 even though they are not single products. The basis of the reduced manifold is

$$w_M{}^J = \sum_m \sum_\mu A^{Jls}_{Mm\mu} w^{ls}_{m\mu} \tag{194}$$

where $J = l+s,\ l+s-1,\ \cdots |l-s|$. As there is no accidental degeneracy in the inner quantum number J, it follows at once that the perturbation matrix is diagonal, so that the functions $w_M{}^J$ are themselves the zero-approximations to the exact solutions. The energy levels will depend only on J and not on M. There will thus be $2S+1$ (or $2l+1$ if $l<S$) different energy levels arising from a single energy level of the spin-free equation. These levels form a unit, the normal multiplet;[48] the number $2S+1$ is called its multiplicity.

These energy levels may be characterised by the three quantum numbers J, l, S, for according to Eq. (147), the operators L^2 and S^2 are diagonal to the approximation here considered. The selection principles

$$\Delta J = \pm 1 \text{ or } 0, \text{ but } J = 0 \nrightarrow J = 0 \text{ excluded,}$$
$$\text{odd} \rightleftarrows \text{even}$$

follow immediately from the considerations of Section 24. The $w_M{}^J$ are odd or even according to the character of $u_m{}^{lr}$. These selection principles are valid to every approximation; two other approximate rules are easily deduced by noting that, because of Eq. (194), the transition amplitudes are approximately given by linear functions of

$$(X w^{lS}_{m\mu},\ w'^{l'S'}_{m'\mu'}), \text{ etc.} \tag{195}$$

Since X, Y, Z are independent of the spin variables (commute with $D(\alpha,\beta,\gamma)$) and $Dw_{m\mu}{}^{lS} = {}^{S}D_{\mu\nu} w_{m\nu}{}^{lS}$, it follows from Section 11, that each of these expressions will be zero unless

$$\Delta S = S - S' = 0.$$

Furthermore, it is known that the expressions (195) will vanish unless

$$\Delta l = \pm 1 \text{ or } 0.$$

These selection principles are identical with those deduced from the Landé theory. The quantum theory of the normal multiplet differs from the Landé theory in one important respect: namely, the dual rôle of the quantum number S. This makes itself apparent in the separations of the energy levels. According to the older theory, all levels having the same l and the same electronic configuration (cf. Section 29) should differ in energy only because of the magnetic properties of the electron. Such levels can differ in J and S, and the theory obviously demands that the energy-differences between levels of different J to be of the same order of magnitude as the

[48] Cf. F. Hund, Linien Spektren, Berlin (1927), sec. 22.

energy differences between levels of different S. It is well-known that this is not the case: in He I, for example, the levels of lines differing only in J are very closely spaced, but the energy differences between those levels of $S=0$ (singlet system) and $S=1$ (triplet system) are comparable to those between levels of different l ($nS-nP$, etc). This is readily accounted for by the foregoing analysis: even when the magnetic (spin) energies of the electrons are completely neglected, a quantum number S can be introduced which specifies the character of the representation P_{rt}. Levels having different S values will differ in energy even to this approximation (cf. Section 6). When the magnetic energy is then considered as a perturbation, the Pauli exclusion principle, Eq. (193a), requires that the total spin momentum shall be equal to this quantum number; it is thus comprehensible that the two sets of energy differences should have different magnitudes.[49]

It should be emphasized that the normal multiplet is an approximation to actual multiplets, as is obvious from the fact that only the zero approximation to the wave-function has been used in the discussion. The theory of non-normal multiplets may most readily be treated by considering both the electrostatic interaction of the electrons and their magnetic energy as perturbations.

29. **The enumeration of the terms arising from a given electronic configuration.** In Eqs. (185) and (186) the operators $H(i)$ depend on the four variables x_i, y_i, z_i, s_i. If the interaction of the spin and orbital motion be neglected, the solutions of Eq. (186) may be written

$$u_i(i) = v_i(x_i y_i z_i)\xi(\tau_i, s_i), \tag{195}$$

where $\xi(\frac{1}{2}, s_i)=\xi(s_i)$, $\xi(-\frac{1}{2}, s_i)=\eta(s_i)$. The number τ_i may be called the spin quantum number of the i-th electron, in distinction to s_i, which is the spin-variable. The N functions $v_1, v_2, \cdots v_N$ may be said to specify the electronic configuration of the atom. For the moment, it will be convenient to assume that the sequence of functions $v_1, v_2, \cdots v_N$ is given, but that the sequence of quantum numbers $\tau_1\tau_2\cdots\tau_N$ is arbitrary, subject only to the restriction that the determinant (190) must not vanish when the functions (195) are substituted into it. The first problem presenting itself is the enumeration of the sequences which satisfy this condition. As the τ's can take on only the values $\pm 1/2$, there are 2^N sequences in all, so that the number of sequences satisfying the condition cannot be greater than 2^N. For the purpose in hand, it may be assumed that the determinant will vanish only when two or more of the functions u_i are identical in form. This will certainly not be the case if all the functions v_i are different, but even if $v_i=v_j$ it may still be avoided if $\tau_i\neq\tau_j$. If more than two of the functions v_i are identical, however, then the determinant will certainly vanish.

In counting up the number of different nonvanishing determinants it will therefore be permissable to consider only the case where

[49] For an explicit treatment of this subject, see P. A. M. Dirac, Proc. Roy. Soc. **A123**, 714 (1929).

$$v_1 = v_2,\ v_3 = v_4,\ \cdots v_{2-1} = v_{2d}, \tag{196}$$

but

$$v_2 \neq v_4 \neq v_6 \cdots \neq v_{2d},$$

and

$$v_i \neq v_j$$

if either i or j is greater than $2d$. The integer d may be referred to as the number of pairs. Since τ_1 must be different from τ_2, τ_3 from τ_4, etc., it is allowed to take $\tau_1=\frac{1}{2}$, $\tau_2=-\frac{1}{2}$, $\tau_3=\frac{1}{2}$, $\tau_4=-\frac{1}{2}$, $\cdots \tau_{2d-1}=\frac{1}{2}$, $\tau_{2d}=-\frac{1}{2}$, leaving only the $N-2d$ indices $\tau_{2d+1}\cdots\tau_N$ arbitrary. For each sequence of these, the determinant will not vanish, nor will two sequences yield the same determinant. Hence there are just 2^{N-2d} antisymmetric solutions for each electronic configuration.

The next problem which presents itself is the addition of the spin vectors. The 2^{N-2d} antisymmetric functions just enumerated subtend a manifold, $\mathfrak{M}_D$, which is invariant under the group of unitary substitutions D. To see this, call the simple product of the N functions (195) $\psi(\tau_1\tau_2\cdots\tau_N)$; then certainly

$$D\psi(\tau_1\cdots\tau_N) = {}^{1/2}D_{\tau_1\sigma_1}\cdots{}^{1/2}D_{\tau_N\sigma_N}\psi(\sigma_1\cdots\sigma_N). \tag{197}$$

As the operations D commute with the permutations P, and the determinant obtained from this sequence may be written

$$w(\tau_1\cdots\tau_N) = \sum_P \epsilon(P)P\psi(\tau_1\cdots\tau_N), \tag{198}$$

the truth of the statement follows at once. The character of the representation based on $\mathfrak{M}_D$ is readily calculable:

$$D(\omega, 0, 0)w(\tau_1\cdots\tau_N) = e^{i\tau\omega}w(\tau_1\cdots\tau_N) \tag{199}$$

where

$$\tau = \tau_1 + \tau_2 + \tau_3 + \cdots + \tau_N$$

and is an integer or a half-integer according as N is even or odd; therefore

$$\chi(\omega) = \sum e^{i\tau\omega}.$$

The sum is to be extended over all sequences $\tau_1\cdots\tau_N$ which yield a non-vanishing determinant. It is most readily evaluated by first counting up the number of such sequences yielding a given value of τ; as the sum of the first 2d terms is necessarily zero, this number is

$$\binom{N-2d}{N/2-d-\tau}$$

and therefore

$$\begin{aligned}\chi(\omega) &= \sum_{\tau=-N/2+d}^{N/2-d}\binom{N-2d}{N/2-d-\tau}e^{i\tau\omega} \\ &= \sum_{S=1/2 \text{ or } 0}^{N/2-d} q(S,N-2d)^S\chi(\omega),\end{aligned} \tag{200}$$

by Eqs. (111) and (169). The possible values of the resultant spin are therefore

$$S = N/2 - d,\ N/2 - d - 1,\ \cdots\ (\tfrac{1}{2} \text{ or } 0).$$

Values of S greater than $N/2-d$ are thus excluded by the Pauli principle, as was evident from its physical statement. These considerations have led to a further conclusion, however: the representation S does not occur $q(S, N)$ times in the manifold of antisymmetric functions, but only $q(S, N-2d)$ times. The Pauli principle has also reduced the number of terms having a given value of S below that to be expected from the considerations of Section 25. This might also have been deduced from the vector model. The actual reduction of the basis (198) to the form w_μ^{Sp}, where $p=1, 2, \cdots q(S, N-2d)$, may readily be accomplished by the method already described.

It remains to consider the effect of the rotation group on the antisymmetric functions of $\mathfrak{M}_D$. This manifold is not invariant under the operations R, a circumstance which arises from a somewhat too narrow interpretation of the term "electronic configuration." The functions v_i of Eq. (195) are each characterized by three indices n, l, m, the principal, azimuthal, and magnetic quantum numbers. In assuming the sequence of functions $v_1 \cdots v_N$ to be given, it is assumed that all three indices of each function are specified. In the spectroscopic sense of the term "electronic configuration," only the n and l of each v_i is specified, the m being left arbitrary. This is the sense in which the term will be used hereafter: to each electronic configuration there will correspond a number of manifolds like $\mathfrak{M}_D$, one based on each of the possible sequences of functions v_i which can be obtained by different assignments of the magnetic quantum numbers. These possible sequences are restricted in two ways: the conditions (196) must be fulfilled, and $|m| \leqq l$ for each v. The antisymmetric functions obtained from each of these sequences will belong to the same zero-approximation energy level. The various manifolds $\mathfrak{M}_D$ will not all be identical, but may differ in the number of pairs in the sequence $v_1 \cdots v_N$ on which they are based. The manifold containing all of them will be invariant under R, as may be shown by a slight modification of the reasoning of Eqs. (197) and (198). This manifold is already partially reduced: the manifolds $\mathfrak{M}_\mu^s$ which contain all functions having a given S and μ are themselves invariant under the group R, since $RD=DR$. The $(2S+1)$ manifolds having a given S all have the same structure, so that it is necessary to consider the reduction of only one of them.

These considerations are very complex if general cases are considered, but become simple for any special case.[50] For this reason, only two examples will be treated: the first is that of an atom containing two electrons each having $n=2$ $l=1$. In spectroscopic notation, the configuration is p^2. The foregoing considerations are summarized in Table II.

[50] Cf. Wigner, Zeits. f. Physik **43**, 651 (1928) for other examples.

TABLE II. *Configuration* p^2.

m_1	m_2	d	$q(S, N-2d)$		m
			$S=1$	$S=0$	
1	0	0	1	1	1
1	−1				0
0	−1				−1
1	1	1	0	1	2
0	0				0
−1	−1				−2

The first two columns contain the six pairs of values m_1, m_2 which result in essentially different determinants. The second column contains d, the number of pairs. Each of these sequences results in a manifold $\mathfrak{M}_D$ containing 2^{2-2d} functions; these are reducible into submanifolds containing $(2S+1)$ $q(S, 2-2d)$ functions each, as is shown in the fifth and sixth columns. The submanifolds $\mathfrak{M}_\mu^S$ which are invariant under the rotation group are four in number: $\mathfrak{M}_1^1$, $\mathfrak{M}_0^1$, $\mathfrak{M}_{-1}^1$, and $\mathfrak{M}_0^0$, corresponding to the two values 1 and 0 of S. The characters of these are easily calculated: the action of $R(\omega, 0, 0)$ on any of the functions is to multiply it by $e^{im\omega}$, where m is the sum of the entries in the first two columns of the row corresponding to the function. The manifold $\mathfrak{M}_1^1$, for example, contains three functions for which $m=1$, 0, −1, respectively; its character is therefore ${}^1\chi$. The same is true of $\mathfrak{M}_0^1$ and $\mathfrak{M}_{-1}^1$. Hence each of these manifolds is irreducible under the group R and the nine functions contained in them may be designated by $w_{m\mu}^{11}$, having the properties

$$Rw_{m\mu}^{11} = {}^1D_{mn}w_{n\mu}^{11},$$
$$Dw_{m\mu}^{11} = {}^1D_{\mu\nu}w_{m\nu}^{11}.$$

The manifold $\mathfrak{M}_0^0$ is slightly more complex; it contains six functions, and its character is

$$\chi(\omega) = e^{2i\omega} + e^{i\omega} + 2 + e^{-i\omega} + e^{-2i\omega}$$
$$= {}^2\chi(\omega) + {}^0\chi(\omega).$$

It is therefore reducible into one manifold $L=2$ and one $L=0$. Its functions are thus w_{m0}^{20} and w_{00}^{00} for which

$$Rw_{m0}^{20} = {}^2D_{mn}w_{n0}^{20},$$
$$Dw_{m0}^{20} = w_{m0}^{20};$$

and

$$Rw_{00}^{00} = w_{00}^{00},$$
$$Dw_{00}^{00} = w_{00}^{00}.$$

These various functions will be linear combinations of the determinants calculated according to the procedure outlined in the earlier part of this

section. The coefficients of these combinations are again the A's of Section 22. On combining the S and L vectors, it is seen that the final result is one ${}^3P$, one ${}^1D$ and one ${}^1S$ multiplet. All of these arise from the electronic configuration p^2.

As a second example consider the configuration p^3s. The table (Table III) is arranged in the same manner as before. In connection with the first row,

TABLE III. *Configuration p^3s.*

$l=1$			$l=0$	d	2^{N-2d}	$q(N-2d)$			m
m_1	m_2	m_3	m_4			$S=2$	$S=1$	$S=0$	
1	0	−1	0	0	16	1	3	2	0
1	1	0	0						2
1	1	−1	0						1
1	0	0	0	1	4	0	1	1	1
−1	0	0	0						−1
1	−1	−1	0						−1
0	−1	−1	0						−2

a convenient check against mistakes in the preparation of such a table may be noted: the total number of functions in the $\mathfrak{M}_D$ corresponding to this row is 16, and this must be equal to

$$\sum_S (2S+1)q(S,N-2d) = 5\cdot 1 + 3\cdot 3 + 1\cdot 2 = 16.$$

The functions enumerated in this table fall into nine manifolds invariant under R: $\mathfrak{M}_\mu{}^2$, $\mathfrak{M}_\mu{}^1$, and $\mathfrak{M}_0{}^0$. The character of $\mathfrak{M}_\mu{}^2$ is primitive: ${}^0\chi$, but the characters of the others are reducible:

$$\mathfrak{M}_\mu{}^1\colon \chi(\omega) = e^{2i\omega} + 2e^{i\omega} + 3 + 2e^{-i\omega} + e^{-2i\omega} = {}^2\chi(\omega) + {}^1\chi(\omega) + {}^0\chi(\omega),$$

$$\mathfrak{M}_0{}^0\colon \chi(\omega) = e^{2i\omega} + 2e^{i\omega} + 2 + 2e^{-i\omega} + e^{-2i\omega} = {}^2\chi(\omega) + {}^1\chi(\omega).$$

Hence the forty[51] antisymmetric functions listed can be combined into linear combinations $w_{m\mu}{}^{LS}$ with the following values of L and S:

L	0	2	1	0	2	1
S	2	1	1	1	0	0

The coefficients of these linear functions can be calculated by a simple extension of the methods explained in Sections 22 and 25. On adding the L and S vectors, it is obvious that one multiplet of each of the types ${}^5S$, ${}^3D$, ${}^3P$, ${}^3S$, ${}^1D$, and ${}^1P$ will result.

One other case may be considered; that of an atom containing $2(2l+1)$ electrons of the same n and l. In this case the sequence of m-values is determined uniquely, for there are just $2l+1$ different values available and none may be used more than twice, hence each must be used exactly twice. Their

[51] It may not be out of place to remark that the examples are chosen from among the more complex, rather than from the simplest.

sum must then be zero, and since $N-2d=0$, $S=0$, and $q(0, 0)=1$, there is just one manifold, $L=S=J=0$. This is the explanation of the phenomenon of "closed shells."

It is obvious from the foregoing examples that the complete manifold of zero approximations for any unperturbed energy level will usually contain several irreducible submanifolds having the same character. In the first example, there are two manifolds for $J=2$, and two for $J=1$. These differ in the values of L and S from which they arise. The perturbation matrix will thus not be diagonal, but have the form $H_{\alpha\beta}(J)\,\delta(J, J')\,\delta(M, M')$ where α, β are indices referring to functions having the same J and M (cf. Section 6). The zero approximations to the characteristic functions will therefore differ from those considered at the end of Section 26 in that their terms do not all arise from the same L and S. The multiplets enumerated in the foregoing examples will not be "normal," and the two approximate selection principles $\Delta S=0$, $\Delta L=0, \pm 1$ will not be valid. This is an indication of the fact that this method of calculation is capable of yielding a better approximation than the former method.

30. **Evaluation of the perturbation matrix.** In actually carrying out the perturbation calculation outlined in the foregoing section, it is first necessary to evaluate inner products of the form

$$(Hw, w') \tag{201}$$

where w and w' are two N-rowed determinants and $PH=HP$. The evaluation of these is not difficult in the cases here under consideration. The determinants may conveniently be written in the form

$$w = \frac{1}{(N!)^{1/2}} \sum_P \epsilon(P) P\psi \tag{202}$$

$$w' = \frac{1}{(N)^{1/2}} \sum_P \epsilon(P) P\phi,$$

in which

$$\begin{aligned} \psi &= u_1(1)u_2(2)\cdots u_N(N) \\ \phi &= v_1(1)v_2(2)\cdots v_N(N) \end{aligned} \tag{203}$$

and $01/(N!)^{1/2}$ is a normalizing factor. The functions $u_i(i), v_i(i)$ depend on all four variables x_i, y_i, z_i, s_i, and all the u's are different from one another, as are all the v's.

Substituting (202) into (201),

$$\begin{aligned} (Hw, w') &= \frac{1}{N!} \sum_P \sum_Q \epsilon(P)\epsilon(O)(PH\psi, Q\phi) \\ &= \frac{1}{N!} \sum_P \sum_Q \epsilon(P)\epsilon(Q)(PHQ^{-1}\psi, \phi) \end{aligned}$$

since the permutations are orthogonal; since they commute with H, this is

$$= \frac{1}{N!} \sum_P \sum_Q \epsilon(PQ^{-1})(PQ^{-1}H\psi, \phi).$$

Applying the rearrangement theorem to $\sum_P$, this becomes

$$\begin{aligned}(Hw, w') &= \frac{1}{N!} \sum_P \sum_Q \epsilon(P)(PH\psi, \phi) \\ &= \sum_P \epsilon(P)(PH\psi, \phi),\end{aligned} \tag{204}$$

which contains at most $N!$ terms, as compared with the $(N!)^2$ terms initially to be considered. In actual cases, many of these remaining terms are zero; this arises from the fact that H always has the form

$$H = \sum_{i=1}^{N} H(i) \tag{205}$$

where $H(i)$ depends only on the four variables i, or else the form

$$H = \sum_{(i,j)} H(i, j) \tag{206}$$

where $H(i, j)$ depends only on the eight variables i, j.

In the first case, one is led to consider inner products of the type

$$(PH(i)\psi, \phi). \tag{207}$$

The functions entering into it will have the form

$$P\left\{H(i)u_i(i) \prod_{k \neq i} u_k(k)\right\}$$

and

$$v_i(i) \prod_{k \neq i} v_k(k).$$

If the inner product were merely a $3N$-fold integral (Eq. (3)), the expression (207) would be expressible as a product of N three-fold integrals; when it has the form of Eq. (3a), it can still be expressed as the product of N inner products each involving only four variables x_k, y_k, z_k, s_k, and defined by

$$[f(k), g(k)]_k = \sum_{s=-1/2}^{+1/2} \int\int\int f(x_k y_k z_k s) g^*(x_k y_k z_k s) dx_k dy_k dz_k.$$

Furthermore, the sequences $u_k(k)$, $v_k(k)$ will, in most actual cases, satisfy the following relations:

$$\begin{aligned}[u_k(1), u_j(1)]_1 &= \delta_{kj}, \\ [v_k(1), v_j(1)]_1 &= \delta_{kj}, \\ [u_k(1), v_j(1)]_1 &= c_k \delta_{kj}.\end{aligned} \tag{208}$$

In the last equation, c_k is either zero or one; this follows from the fact that the sequences $u_1 \cdots u_N$ and $v_1 \cdots v_N$ will usually differ only by a relatively few functions which are orthogonal to one another. Because of Eq. (208), it is seen that the expression (207) will certainly vanish unless $P = I$, and in that case, reduce to

$$(H(i)\psi, \phi) = [H(i)u_i(i), v_i(i)]_i \prod_{k\neq i} c_k.$$

Even this will vanish unless all the c_k, except possibly c_i, are different from zero. When this condition is fulfilled, with $c_1 = 0$

$$(Hw, w') = [H(1)u_1(1)v_1(1)]_1; \tag{209a}$$

in the particular case $w' = w$; however, none of the c's vanish and

$$(Hw, w) = \sum_{i=1}^{N} [H(i)u_i(i), u_i(i)]_i. \tag{209b}$$

When H has the form (206) inner products of the type

$$(PH(ij)\psi, \phi) \tag{210}$$

must be considered. These will certainly vanish unless P is either the identity or the (odd) permutation which interchanges i and j and leaves all other variables unchanged. In these particular cases, the expressions reduce to

$$[H(ij)u_i(i)u_j(j), v_i(i)v_j(j)]_{ij} \prod_{k\neq i,j} c_k$$

and

$$[H(ij)u_i(j)u_j(i), v_i(i)v_j(j)]_{ij} \sum_{k\neq i,j} c_k$$

where $[\ \ ]_{ij}$ is an inner product involving only the eight variables i and j. These still vanish unless all c_k, except possibly c_i and c_j, are equal to 1. There are now three cases to distinguish:

1. $c_1 = c_2 = 0$, all other $c_k = 1$; then

$$\begin{aligned}(Hw, w') = {} & [H(12)u_1(1)u_2(2), v_1(1)v_2(2)]_{12} \\ & - [H(12)u_1(2)u_2(1), v_1(1)v_2(2)]_{12};\end{aligned} \tag{211a}$$

2. $c_1 = 0$, all other $c_k = 1$;

$$\begin{aligned}(Hw, w') = {} & \sum_{i=2}^{N} \{ [H(1, i)u_1(1)u_i(i), v_1(1)v_i(i)]_{1i} \\ & - [H(1, i)u_1(i)u_i(1), v_1(1)v_i(i)]_{1i} \};\end{aligned} \tag{211b}$$

3. $w = w'$, all $c_k = 1$;

$$\begin{aligned}(Hw, w') = {} & \sum_{(i,j)} \{ [H(ij)u_i(i)u_j(j), u_i(i)u_j(j)]_{ij} \\ & - [H(ij)u_i(j)u_j(i), u_i(i)u_j(j)]_{ij} \}.\end{aligned} \tag{211c}$$

The calculation of the perturbation matrix in terms of the determinants w which formed the starting point of the last section is thus reduced to a relatively simple form. By carrying out the canonic transformations involved in the reduction of the manifold of antisymmetric functions, this matrix can be reduced to the form of Eq. (47). The system of equations thus obtained is in its simplest form and its solution cannot be avoided if the characteristic functions are to be determined. Its characteristic values may often be determined by the method of sums without actual solution of the secular equation, as has been explained.

VI. The Influence of External Fields

31. **The Zeeman effect. The sum rules for the g-factors.** The energy levels of an atom when not acted on by an external force are, as has been seen, always degenerate. The manifold of wave functions corresponding to each is $2J+1$ dimensional[52] and its basis may be supposed given as $u_M{}^J$. Each of these functions is the basis of one of the irreducible representations of the group S, K', $K'S$, which dominates the problem when an homogeneous magnetic field $\mathfrak{H}$ is applied to the atom.

The perturbation operator which must be added to the operator of the wave equation in the absence of the field is

$$\frac{eh\mathfrak{H}}{4\pi mc}(L_z + 2S_z) = \frac{eh\mathfrak{H}}{4\pi mc}(J_z + S_z) \tag{212}$$

where J_z, L_z, S_z are the z-components of the operators associated to the total, orbital, and spin angular momenta, respectively. The perturbation matrix arising from this operator must be diagonal, since there is no accidental degeneracy and the representation of the manifold is completely reduced. This result may be made more precise by noting that J_z+S_z is the z-component of a vector operator, and that therefore, according to the results of Section 24,

$$(J_z + S_z)u_M{}^J = gMu_M{}^J \tag{213}$$

where g is a factor independent of M, but dependent on J. The action of the magnetic field will therefore be to split up the degenerate energy level into $2J+1$ single levels, which are equally spaced.

The approximate values of the g-factors can be determined in the case of normal multiplets, and partially determined in the more general cases. In the former case, the $u_M{}^J$ are approximately of the form

$$u_M{}^J = \sum_m \sum_\mu A^{JLS\ LS}_{M m\mu} v_{m\mu}, \tag{214}$$

so that

$$g(J) - 1]Mu_M{}^J = S_z u_M{}^J = \sum_m \sum_\mu A^{JLS\ LS}_{M m\mu\mu} v_{m\mu}. \tag{215}$$

This equation may be solved for $g(J)$ in several ways, the simplest depending on Eq. (148).[53] The method which will be used here is analogous to the method of sums, and is preferred because it may be extended to the case of a general multiplet. In Eq. (214) J takes on the values $L+S$, $L+S-1$, $\cdots |L-S|$, and $|M| \leqq J$; the equations may also be written, according to Section 22:

$$u^{L+S}_{L+S} = v^{LS}_{LS};$$

$$u^{L+S}_{L+S-1} = a v^{LS}_{(L-1)S} + b v^{LS}_{L(S-1)},$$

$$u^{L+S-1}_{L+S-1} = c v^{LS}_{(L-1)S} + d v^{LS}_{L(S-1)}; \quad \text{etc.}$$

[52] For the special case $J=0$, this degeneracy disappears, but it is not necessary to treat this case separately.

[53] Weyl, *l.c.* p. 164.

From the first of these and Eq. (215), it follows that

$$\text{or} \qquad \begin{aligned} (L+S)[g(L+S)-1] &= S \\ g(L+S)-1 &= S/(L+S). \end{aligned} \tag{216a}$$

From the second, using the fact that the diagonal sum of the matrix associated to S must be invariant under the canonic transformation,

$$\text{or} \qquad \begin{aligned} (L+S-1)\{[g(L+S)-1]+[g(L+S-1)-1]\} &= 2s-1 \\ g(L+S-1)-1 &= (2S-1)/(L+S-1)-S/(L+S). \end{aligned} \tag{216b}$$

The three equations which follow those written out above determine u^{L+S}_{L+S-2}, u^{L+S-1}_{L+S-2}, and u^{L+S-2}_{L+S-2} in terms of $v^{LS}_{L(S-2)}$ $v^{LS}_{(L-1)(S-1)}$, and $v^{LS}_{L(S-2)}$. Reasoning in the same manner, it is found that

$$g(L+S-2)-1 = (3S-3)/(L+S-2)-(2S-1)/(L+S-1). \tag{216c}$$

Continuing, it is readily found by induction, that

$$g(J)-1 = \frac{J(J+1)-L(L+1)+S(S+1)}{2J(J+1)} \tag{216}$$

which is Landé's formula.

This method is capable of yielding some information regarding the approximate values of $g(J)$ in the case of the general multiplets. As an example, let it be required to discuss the g-factors of the levels arising from the configuration *sd*. In this case, the considerations of Section 29 indicate that two values of S, namely 0 and 1, are associated to each energy level when the mutual interaction is not considered. The basis of the manifold of this level may be taken to be v^{20}_{m0}, $v^{21}_{m\mu}$. The effect of the mutual interaction is then, approximately, merely to reduce this basis to $u_M{}^3$, $u'_M{}^2$, $u''_M{}^2$ and $u_M{}^1$, these functions being linear combinations of the original v's. The characteristic to be noted is that there are two sets of u's for which $J=2$; if the coupling were "normal," these functions would correspond to a ^3D and a ^1D normal multiplet. The same procedure is to be followed as before:

$$u_3{}^3 = v^{21}_{21},$$

hence

$$3[g(3)-1] = 1, \text{ or } g(3) = 4/3.$$

The next step encounters difficulties, however:

$$\begin{aligned} u_2{}^3 &= av^{20}_{20} + bv^{21}_{20} + cv^{21}_{11} \\ u'_2{}^2 &= a'v^{20}_{20} + b'v^{21}_{20} + c'v^{21}_{11} \\ u''_2{}^2 &= a''v^{20}_{20} + b''v^{21}_{20} + c''v^{21}_{11} \end{aligned}$$

and hence there is only one equation for two unknowns:

$$2\{[g(3)-1]+[g'(2)-1]+[g''(2)-1]\} = 0+0+1$$

whence all that can be deduced is that

$$g'(2)+g''(2) = 13/6.$$

For $g(1)$, the result is determinate again:

$$g(1) = \tfrac{1}{2}.$$

On comparing these results with the g-factors for a normal ^3D, viz., $g(3) = 4/3$, $g(2) = 7/6$, $g(1) = 1/2$, and a normal ^1D, $g(2) = 1$, the general rule becomes obvious: If several levels having the same J-value arise from the same electronic configuration, the sum of their g-factors is equal to the sum of the g-factors calculated from Eq. (216) for the normal multiplets. It may also be remarked that this rule is independent of the field strength and thus governs the Paschen-Back effect.

32. **Stark effect.**[54] The manifold of functions associated to each energy level in the absence of an external field when specified in terms of a basis u_{-M}^{J} is also completely reduced to its submanifolds which are invariant under the group S, K'', SK'' which governs the problem when an homogeneous electric field is applied to the atom. The basis of the irreducible representation of this group spcified by the number M is

$$u_M{}^J,\ u_{-M}{}^J$$

or $u_0{}^J$ if J is integral and $M=0$. The action of the electric field, if J is integral, is thus to split up the energy level into J levels each two-fold degenerate, and one non-degenerate level. If J is half integral, there will be $J+1/2$ two-fold degenerate levels.

The perturbation energy in this case is

$$-eEZ,$$

E being the strength of the field and Z the sum of the rectangular coordinates of the electrons in the direction of E. Since

$$(Zu_M{}^J,\ u_M{}^J),$$

and

$$(Zu_M{}^J,\ u_{-M}{}^J),$$

are both zero (Z is an odd function) the perturbation matrix vanishes identically in this case. Hence the splitting of the energy level will be zero to the first approximation[55] The separation between the different levels will be proportional to E^2. There will thus be no linear Stark-effect unless there is accidental degeneracy—i.e., unless two levels of different J have the same energy value. This is the case only in one actual example—that of hydrogen.

[54] E. Wigner, Zeits. f. Physik **43**, 643 (1928).

[55] Cf. Neumann and Wigner, l.c.