GROUP-THEORETICAL METHODS IN THE QUANTUM PHYSICS OF SOLIDS (SPATIAL SYMMETRY)

A. V. SOKOLOV and V. P. SHIROKOVSKIĬ

Usp. Fiz. Nauk 71, 485-513 (July, 1960)

TABLE OF CONTENTS

THE purpose of the present paper is to present the fundamental ideas of the application of group theoretical methods to quantum physics of solids in a form accessible to a wide circle of physicists. As yet there has not been in the Soviet literature any sufficiently systematic presentation of this problem. The paper covers theoretical material concerning spatial symmetry and is a logical extension of the paper of the authors concerning point symmetry.[1]

1. INTRODUCTION

The study of space groups and their application to the quantum theory of solids has played a very important role in the development of our knowledge of metallic and semiconductor crystals. The successes achieved in the theory of semiconductors and their rapid inclusion in technology during the last ten years is to a large extent due to the application of the method of group theory to the study of electron energy spectra and wave functions.

We may mention that the application of the theory of space groups to the physics of solids is of decisive importance, since in solids we cannot neglect the crystal structure of the material and therefore we cannot avoid the use of the mathematical apparatus of group theory.

Bethe[2] first succeeded in applying group theory to the quantum physics of solids. In his paper he investigated theoretically the splitting in a crystalline field of levels which were degenerate in the free atom. The next very important step was a series of papers by Seitz[3] concerning the theory of space groups. In these papers of Seitz and the fundamental paper of Bouckaert, Smoluchowski and Wigner[4] we have the beginning of a whole series of papers concerning the band theory of solids. The authors of reference 4, using the concept of the group of the wave vector, obtained tables of characters of the irreducible representations of the symmetry groups of simple lattices, while in the papers of Herring[5] and of Döring and Zehler[6] this method was extended and applied to more complicated lattices. Of the later investigations we should mention the papers of Herring[7] and Elliott,[8] which were concerned with an investigation of the effect on the energy spectra of electrons in crystals, of time reversal, and of the spin-orbit interaction.

We should also mention the very recent paper of Koster[9] on the theory of space groups, in which a detailed and quite clear presentation of their properties is given.

The treatment of band theory using the rigorous apparatus of group theory enables one to clarify certain delicate questions which escape consideration in the usual methods of treatment. For example, when we include not only the translational but also the rotational symmetry of the crystal lattice, we get an overlapping of the energy bands. Such an overlapping may have an essential effect, for example, on infrared absorption in semiconductors and on other phenomena.

In the present paper we discuss the physical basis for the application of the method of group theory to the physics of solids. Fundamental attention is given to the theory of space groups and their most important properties. In particular, we consider the representations of the space groups and give methods for con-

structing them from the known representations of the 32 point groups. The fundamentals of the band theory of solids are described from the point of view of group theory. The general method developed by us for finding conserved quantities and a complete set[1] is applied to the investigation of the energy spectrum and the classification of states for electrons in a crystal lattice. The change in the electron energy spectrum and states when spin-orbit interaction is included is treated. Finally, in the last paragraph a discussion is given of the question of invariance of the Hamiltonian with respect to time reversal and rules are formulated for the discovery of additional degeneracy both for particles without spin and for particles with spin $\frac{1}{2}$.

2. DESCRIPTION OF SPATIAL SYMMETRY OPERATIONS

Ideal crystal lattices always possess a certain spatial symmetry which is a combination of point symmetry and translational symmetry. Therefore, in the application of the theory of groups to the physics of solids, the theory of the symmetry properties of crystalline systems is of basic importance.

Before we treat the rigorous theory of space groups, it is necessary to describe the fundamental geometrical concepts.

If a system τ is to be taken from position τ_1 to position τ_2, this can be done in various ways. However, all the motions by which this transition is accomplished are considered to be equivalent and the simplest one is always selected. As such simple motions one can choose translations and rotations. It can be shown that an arbitrary motion is equivalent in the above sense to one or a set of several of these simplest motions.

If we subject the system τ to the action of some translation $\mathbf{T}$, all its points describe trajectories which are equal in magnitude and direction; therefore, a translation is completely determined by giving the path of one point.

A rotation in space is characterized by giving the position of some axis a; its magnitude is determined by the angle of rotation α. If we rotate not through the angle α, but through the angle $\alpha + 2\pi$, the moving system will occupy the same final position as it does after the rotation through angle α. Consequently, these two rotations are equivalent. Therefore in our discussion we shall limit ourselves to those rotations having angles less than 2π. The transition from the initial position τ_1 to the final position τ_2 can be achieved by a rotation of the system τ around the axis a through the angle $2\pi - \alpha$ in the opposite direction. We call such a rotation negative. It is obvious that every rotation of the system τ around the axis a can be carried out by either a positive or a negative rotation. But since we always treat just one of the whole set of equivalent motions, we shall for simplicity work only with positive rotations. We denote the rotation through angle α around the axis a by $\mathbf{A}(\alpha)$. If the body coincides with itself after rotation around the axis a through the angle $2\pi/n$, this axis is said to be an n-fold symmetry axis.

A screw motion consists of a rotation through angle α around a certain axis a and translation by the amount t along it. The rotation and translation are carried out simultaneously. The transition of the system τ from position τ_1 to position τ_2 accomplished by a screw motion can always be carried out so that the translation and rotation follow one another in arbitrary succession. A screw axis with angle of rotation α around the axis a and a translation of magnitude t will be denoted by $\mathbf{A}(\alpha, \mathbf{t})$. If the body coincides with itself after rotation around the axis through angle $2\pi/n$ and simultaneous translation by t along the same axis, we say that the body has an n-fold screw axis. If we carry out the rotation and translation n times with respect to an n-fold screw axis, we will as a result move the body along the axis through a distance equal to nt. Consequently, when there is a screw axis, the body must always have a periodicity of the usual type along this axis with a period no greater than nt. This means that an n-fold screw axis can be associated only with translation through distances

$$\mathbf{t} = \frac{p}{n}\mathbf{a} \qquad (p = 1, 2, \ldots, (n-1)),$$

where $\mathbf{a}$ is the smallest period along the direction of the axis.

Let us now consider, together with the system τ, a system $\bar{\tau}$ obtained from the first by mirror reflection in some plane σ. The system and its mirror image cannot be brought into coincidence with one another by means of ordinary motions; this can be accomplished only by applying a mirror reflection. If τ_1 is some position of the body τ, and $\bar{\tau}_2$ is some other position of the body $\bar{\tau}$, we can first reflect τ_1 in some plane and then make the resultant mirror image $\bar{\tau}_1$ coincide with $\bar{\tau}_2$ by means of a Euclidean motion. Such an operation is called an operation of the second kind. The simplest operations of the second kind are reflections.

A reflection is characterized by giving the position of some plane σ; we say that we have a reflection in a glide plane, when we add to the reflection in σ a translation parallel to it through t. We denote the reflection in the plane σ by Σ and the reflection in a glide plane by $\Sigma(\mathbf{t})$, where t is the magnitude of the translation parallel to σ. We say that a body has a plane of symmetry σ if it coincides with itself after reflection in this plane. The body has a glide plane of symmetry σ if it coincides with itself after reflection in the plane and simultaneous translation through a distance t in a direction parallel to the plane. Twofold reflection in a glide plane results in a simple translation through distance 2t. Therefore a body can have only those glide planes of symmetry in which the magnitude of

the translation t is equal to $\mathbf{a}/2$ where a is the length of the smallest period in the direction of this translation.

A mirror rotation consists of a rotation around some axis a through angle α and subsequent reflection in the plane σ perpendicular to this axis. We say that a body has an n-fold mirror-rotation axis if it coincides with itself after rotation about this axis through angle $2\pi/n$ and subsequent reflection in the plane perpendicular to the axis. We use the notation $\mathbf{S}(\alpha)$ for mirror reflections.

If the angle of rotation for a mirror reflection is $\alpha = \pi$, this operation results in inversion with respect to the point of intersection of the axis a and the plane σ.

The symmetry of a crystal is very naturally subdivided into macroscopic and microscopic. Macroscopic symmetry determines those properties of a crystal which depend only on directions in it, so that the crystal behaves like a homogeneous continuous body. Here the word "homogeneous" emphasizes that the dependence of physical properties on direction is the same at all points of the crystal. From the point of view of structural crystallography, macroscopic symmetry is given, as we know, by the 32 crystal classes. These are symmetry groups made up of point symmetry elements: rotations and reflections.

By microscopic symmetry we mean the full intrinsic symmetry of crystal lattices. Microscopic symmetry determines those properties of a crystal which depend on the arrangement of the atoms in its lattice. We shall first treat translation groups, which express the translational symmetry of the possible space lattices and then the space groups which express the combined rotational and translational symmetry.

3. SOME RESULTS OF GROUP THEORY

Suppose we have a finite or infinite set $\mathfrak{G}$ of elements $g_1, g_2, \ldots, g_k$. This set forms a group if the following conditions are satisfied:

1. The product $g_i g_j$ of any two elements of the set taken in a definite order is an element of the same set

$$g_i g_j = g_h.$$

2. There is an element e in the set $\mathfrak{G}$ satisfying the relation

$$eg_k = g_k e = g_k .$$

This element is called the unit element.

3. To every element g_k of the set $\mathfrak{G}$ there corresponds in the same set another element g_k^{-1} defined by the relation

$$g_h g_h^{-1} = g_h^{-1} g_h = e$$

and called the inverse of g_k.

4. The product of the elements satisfies the associative law

$$(g_i g_j) g_h = g_i (g_j g_h).$$

In general the commutative law does not hold, i.e., in the general case

$$g_i g_j \neq g_j g_i.$$

If the number g of elements of the group $\mathfrak{G}$ is finite, we say that we have a finite group and that its order is g. When all the elements of the group commute with one another, the group is said to be commutative, while if this is not the case, it is said to be noncommutative.

A set $\mathfrak{H}$, made up of an arbitrary number of elements of the group $\mathfrak{G}$ is called a sub-group of the group $\mathfrak{G}$ if this set itself is a group with respect to the operation defined in $\mathfrak{G}$. The number of elements h of this subgroup is called the order of the sub-group $\mathfrak{H}$.

Suppose we are given a sub-group $\mathfrak{H}$ in a group $\mathfrak{G}$. If g_k is any element of $\mathfrak{G}$, the product $g_k\mathfrak{H}$ is called a left residue class of the group $\mathfrak{G}$ with respect to the subgroup $\mathfrak{H}$, defined by the element g_k. It is clear that the element g_k itself is contained in the class $g_k\mathfrak{H}$, since the subgroup $\mathfrak{H}$ contains the unit element. If g'_k is an arbitrary element of the class $g_k\mathfrak{H}$, the left residue classes $g_k\mathfrak{H}$ and $g'_k\mathfrak{H}$ coincide, i.e. every left residue class is defined by any one of its elements. In fact, since $g'_k \in g_k\mathfrak{H}$, $g'_k = g_k h_k$. Then $g'_k\mathfrak{H} = g_k(h_k\mathfrak{H}) = g_k\mathfrak{H}$. From this it follows that any two left residue classes of the group $\mathfrak{G}$ with respect to the subgroup $\mathfrak{H}$ either coincide or have no element in common. As a result we find that the whole group $\mathfrak{G}$ splits into non-overlapping left residue classes with respect to the subgroup $\mathfrak{H}$. This expansion is called the expansion into left residue classes of the group $\mathfrak{G}$ with respect to the subgroup $\mathfrak{H}$. It will have the form

$$\mathfrak{G} = \mathfrak{H} + g_2\mathfrak{H} + \ldots + g_m\mathfrak{H}. \qquad (3.1)$$

Instead of this left-sided expansion, we could also obtain a right-sided expansion of the group $\mathfrak{G}$ with respect to the subgroup $\mathfrak{H}$. It can be stated that both expansions of any group $\mathfrak{G}$ with respect to an arbitrary subgroup $\mathfrak{H}$ consist of the same number of residue classes. If we remember also that each residue class consists of precisely h elements, it follows immediately from the expansion (3.1) that

$$g = hm, \qquad (3.2)$$

i.e., the order of a subgroup is a divisor of the order of the group.

A subgroup $\mathfrak{N}$ of the group $\mathfrak{G}$ is said to be a normal divisor of this group, or an invariant subgroup, if the left-sided expansion of the group $\mathfrak{G}$ with respect to the subgroup $\mathfrak{N}$ coincides with the right-sided expansion. In other words, $\mathfrak{N}$ will be a normal divisor in $\mathfrak{G}$ if the left-sided and right-sided residue classes of $\mathfrak{G}$ with respect to $\mathfrak{N}$ defined by the element g_k coincide for all g_k:

$$g_k\mathfrak{N} = \mathfrak{N}g_k. \tag{3.3}$$

This shows that the necessary and sufficient condition for a subgroup $\mathfrak{N}$ to be a normal divisor of a group $\mathfrak{G}$ is that it commute with any element of the group $\mathfrak{G}$.

Let us now give the following definition, which enables us to construct groups starting from a given group $\mathfrak{G}$. Let $\mathfrak{N}$ be a normal divisor of the group $\mathfrak{G}$. Furthermore let N_i and N_j be residue classes with respect to $\mathfrak{N}$: $N_i = g_i\mathfrak{N}$, $N_j = g_j\mathfrak{N}$. We form the product N_iN_j; we have

$$N_iN_j = g_i\mathfrak{N}g_j\mathfrak{N} = g_ig_j\mathfrak{N}\mathfrak{N} = g_k\mathfrak{N} = N_k,$$

i.e., the product N_iN_j is also a residue class with respect to $\mathfrak{N}$. Thus there is established a law of multiplication in the set of residue classes. Let us show that it satisfies the group axioms 2 — 4. Associativity is obvious since it holds in $\mathfrak{G}$. The unit element of the group is $\mathfrak{N}$. In fact, if N_k is any residue class,

$$\mathfrak{N}N_k = \mathfrak{N}g_k\mathfrak{N} = g_k\mathfrak{N}\mathfrak{N} = g_k\mathfrak{N} = N_k.$$

The element reciprocal to $N_k = g_k\mathfrak{N}$ is $g_k^{-1}\mathfrak{N}$, since

$$(g_k\mathfrak{N})(g_k^{-1}\mathfrak{N}) = g_kg_k^{-1}\mathfrak{N} = \mathfrak{N}.$$

The group of residue classes obtained in this way is called the factor group of the group $\mathfrak{G}$ with respect to the subgroup $\mathfrak{N}$ and is denoted by $\mathfrak{G}/\mathfrak{N}$.

4. SPACE GROUPS AND THEIR PROPERTIES[9]

From the purely group-theoretical point of view, space groups are a special case of more general groups of linear transformations which preserve length. The general form of such transformations can be written as follows:

$$\left.\begin{aligned} x_1' &= \alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3 + t_1,\\ x_2' &= \alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3 + t_2,\\ x_3' &= \alpha_{31}x_1 + \alpha_{32}x_2 + \alpha_{33}x_3 + t_3, \end{aligned}\right\} \tag{4.1}$$

or, abbreviated,

$$\mathbf{x}' = \alpha\mathbf{x} + \mathbf{t}. \tag{4.2}$$

In order that transformations of the type (4.1) preserve length, it is necessary to require that the vectors $\mathbf{t}$ have real components and that the matrices α be real orthogonal matrices. From the last requirement it follows that by a unitary transformation the matrix α can always be brought to the form

$$\alpha = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}. \tag{4.3}$$

In this form these matrices allow of a simple interpretation, namely, the matrices with the + sign correspond to rotations around the x_1 axis through angle φ (proper rotations), while the matrices with the − sign can be regarded as rotations around the x_1 axis through angle φ with subsequent reflection in the x_2x_3 plane (improper rotations).

Thus the transformations (4.1) can be understood to be rotations around some axis with a succeeding translation by the vector $\mathbf{t}$. Such a coordinate transformation will be denoted in operator form as

$$\{\alpha \mid \mathbf{t}\}. \tag{4.4}$$

In this notation the operator $\{\epsilon \mid 0\}$ is the identity transformation, the operators $\{\epsilon \mid \mathbf{t}\}$ are translations, while the operators $\{\alpha \mid 0\}$ are rotations (proper or improper).

It is easily shown that in operator form the product of two transformations of type (4.1) is written as follows:

$$\{\alpha \mid \mathbf{t}\}\{\beta \mid \mathbf{u}\} = \{\alpha\beta \mid \alpha\mathbf{u} + \mathbf{t}\}, \tag{4.5}$$

It is also easy to find the operator which is inverse to a given operator,

$$\{\alpha \mid \mathbf{t}\}^{-1} = \{\alpha^{-1} \mid -\alpha^{-1}\mathbf{t}\}. \tag{4.6}$$

Thus the operators of type (4.4) form a group. We mention two obvious properties of this group:

1. The operators $\{\alpha \mid 0\}$ form a subgroup in this group;

2. The operators $\{\epsilon \mid \mathbf{t}\}$ also form a subgroup in this group and this is an invariant subgroup of the original group since

$$\{\alpha^{-1} \mid -\alpha^{-1}\mathbf{u}\}\{\varepsilon \mid \mathbf{t}\}\{\alpha \mid \mathbf{u}\} = \{\varepsilon \mid \alpha^{-1}\mathbf{t}\}, \tag{4.7}$$

i.e., every element of this subgroup which is conjugate to a given element $\{\epsilon \mid \mathbf{t}\}$ also belongs to it.

The space groups are characterized by the fact that they have an invariant subgroup of translations of a special kind, namely: all the pure translations of the space group have the form

$$\{\varepsilon \mid \mathbf{R}_n\}, \tag{4.8}$$

where

$$\mathbf{R}_n = n_1\mathbf{t}_1 + n_2\mathbf{t}_2 + n_3\mathbf{t}_3,$$

the n_i are integers and $\mathbf{t}_i$ are three linearly independent translations which are called primitive translations. Consequently, the pure translations in a space group are linear combinations with integer coefficients of the three fundamental primitive translations. The periodically repeating set of points generated by the vectors $\mathbf{R}_n$ is called a lattice.

All the properties of space groups can be derived from the fact that they contain an invariant subgroup of this type. Thus, for example, it is immediately obvious that if $\mathbf{R}_n$ is an admissible translation, then $\alpha\mathbf{R}_n$, where $\{\alpha \mid \mathbf{t}\}$ is an element of the space group, will also be an admissible translation. This follows immediately from the relation

$$\{\alpha \mid \mathbf{t}\}\{\varepsilon \mid \mathbf{R}_n\}\{\alpha^{-1} \mid -\alpha^{-1}\mathbf{t}\} = \{\varepsilon \mid \alpha\mathbf{R}_n\}.$$

One of the consequences of the fact that space groups have an invariant subgroup of admissible

translations of the form (4.8) is the limitation imposed on rotation operators. It turns out that one can have rotations around an axis only through angles which are multiples of 60 and 90°; the improper rotations are products of such rotations with the inversion. In the classification of possible groups which are formed by the rotation operators contained in space groups, one finds that there are only 32 (32 crystal classes or 32 point groups). The rotational part of every space group corresponds to one of these 32 point groups.

On the other hand, knowing to which crystal class a particular space group corresponds, one can obtain information concerning the possible invariant subgroups of admissible translations. We have already seen that if $\mathbf{R}_n$ is an admissible translation and $\{\alpha|\mathbf{t}\}$ is an element of the space group, then $\alpha\mathbf{R}_n$ will also be an admissible translation. Consequently the lattice generated by the admissible translations of the space group must remain invariant under the action of the operations of the point group. This proves to be sufficient for imposing completely definite limitations on the fundamental translation vectors $\mathbf{t}_1$, $\mathbf{t}_2$, $\mathbf{t}_3$, from which all the $\mathbf{R}_n$ are obtained. Investigations show that there are 14 different lattices (Bravais lattices).

It should be mentioned that a space group is only partially characterized by its point group and the type of lattice. Space groups having the same crystal class and the same lattice can differ still in the form of the operators of the translational part (in the form of $\mathbf{t}$). However, investigations show that all the operators of a given space group can always be represented in the form

$$\{\varepsilon|\mathbf{R}_n\}\{\alpha|\mathbf{v}(\alpha)\}, \tag{4.9}$$

where $\mathbf{R}_n$ are the vectors of admissible translations, and $\mathbf{v}(\alpha)$ is a vector which is characteristic for the particular rotation α and is either equal to zero or is an admissible translation [we note that $\mathbf{v}(\epsilon) = 0$].

All space groups are divided into two types with respect to the vectors $\mathbf{v}(\alpha)$. The first type contains those groups in which $\mathbf{v}(\alpha)$ is equal to zero for all α. These are the so-called simple space groups, of which there are 73. To each operator α of the point group in a simple space group there corresponds the operator $\{\alpha|0\}$. Making use of this and the equation

$$\{\alpha|0\}\{\beta|0\} = \{\alpha\beta|0\}, \tag{4.10}$$

we can assert that the operators $\{\alpha|0\}$ form a group isomorphic to the point group. In other words, a simple space group contains the whole point group as a subgroup.

In the other 157 space groups $\mathbf{v}(\alpha)$, for at least one α, cannot be chosen equal to zero. Motions containing a shift by an inadmissible translation which follows a proper or improper rotation usually correspond to glide planes and screw axes. In this case the point group is not a subgroup of the space group.

However, we have seen that in every space group the admissible translations form an invariant subgroup. Let us denote the space group by $\mathfrak{G}$, and the invariant subgroup of translations by $\mathfrak{T}$. We can then form the factor group $\mathfrak{G}/\mathfrak{T}$. It is easy to show that this factor group is isomorphic to the point group consisting of the rotational part of the operators of the space group.

5. REPRESENTATIONS OF SPACE GROUPS (GENERAL THEORY)[9]

The finding and classification of the irreducible representations of space groups is most conveniently begun with a description of the irreducible representations of the translation groups $\mathfrak{T}$ consisting of elements of the form $\{\epsilon|\mathbf{R}_n\}$. In order to make this group finite, we assume that

$$\{\varepsilon|\mathbf{t}_1\}^N = \{\varepsilon|\mathbf{t}_2\}^N = \{\varepsilon|\mathbf{t}_3\}^N = \{\varepsilon|0\}. \tag{5.1}$$

Since different translations $\{\epsilon|\mathbf{t}_i\}$ commute with one another, the group $\mathfrak{T}$ can be regarded as the direct product of groups generated by each of these elements. It is therefore sufficient to find the representations of one of these groups, for example, the group generated by the element $\{\epsilon|\mathbf{t}_1\}$, which obviously is an abelian group of N'th order. Then, according to the general theory, its representations must have the form

$$\exp\left(i\frac{2\pi p_1}{N}\right), \qquad p = 0, 1, \ldots, N-1. \tag{5.2}$$

The products of all the irreducible representations of the one-dimensional translation groups with one another give the irreducible representations of the full group $\mathfrak{T}$. Thus we will have

$$\exp i(2\pi\mathbf{k}\mathbf{R}_n), \tag{5.3}$$

where

$$\begin{aligned}\mathbf{R}_n &= n_1\mathbf{t}_1 + n_2\mathbf{t}_2 + n_3\mathbf{t}_3,\\ \mathbf{k} &= k_1\mathbf{b}_1 + k_2\mathbf{b}_2 + k_3\mathbf{b}_3,\end{aligned} \tag{5.4}$$

in which

$$\begin{aligned}&\mathbf{t}_i\mathbf{b}_j = \delta_{ij}, \qquad i, j = 1, 2, 3,\\ &k_i = \frac{p_i}{N}, \quad p_i = 0, \ldots, (N-1).\end{aligned} \tag{5.5}$$

Consequently the vector $\mathbf{k}$ completely determines the irreducible representations of the translation group. When N tends to infinity, there corresponds, to each vector $\mathbf{k}$ with components lying in the interval $0-1$, some irreducible representation. We should mention only that the vector $\mathbf{k}$ is defined to within an arbitrary vector $\mathbf{h}$ of the form

$$\mathbf{h} = h_1\mathbf{b}_1 + h_2\mathbf{b}_2 + h_3\mathbf{b}_3, \quad h_i - \text{integers} \tag{5.6}$$

(vector of the reciprocal lattice), i.e., the vector $\mathbf{k}' = \mathbf{k} + 2\pi\mathbf{h}$ characterizes the same representation as $\mathbf{k}$.

We now proceed to study the irreducible representations of the space groups. We shall denote the space

group by $\mathfrak{G}$, its elements by $\{\alpha \mid \mathbf{a}\}$, and the matrices of the irreducible representations of the space group by $D(\alpha \mid \mathbf{a})$. (Without loss of generality we can assume that these matrices are unitary.)

Those matrices of the irreducible representations of a space group which correspond to pure translations give a representation of the translation group $\mathfrak{T}$, where again without loss of generality we may assume that all the $D(\epsilon \mid \mathbf{R_n})$ are diagonal and have the form

$$D(\varepsilon \mid \mathbf{R}_n) = \begin{pmatrix} \exp(i\mathbf{k}_1\mathbf{R}_n) & 0 & \dots & 0 \\ 0 & \exp(i\mathbf{k}_2\mathbf{R}_n) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \exp(i\mathbf{k}_q\mathbf{R}_n) \end{pmatrix}. \tag{5.7}$$

In the preceding paragraph it was already stated that if $\mathbf{R_n}$ is some admissible translation, then $\alpha^{-1}\mathbf{R_n}$ is also an admissible translation. This followed immediately from the equation

$$\{\alpha \mid \mathbf{a}\}^{-1}\{\varepsilon \mid \mathbf{R}_n\}\{\alpha \mid \mathbf{a}\} = \{\varepsilon \mid \alpha^{-1}\mathbf{R}_n\}. \tag{5.8}$$

The matrix corresponding to $\alpha^{-1}\mathbf{R_n}$ is written as

$$D(\varepsilon \mid \alpha^{-1}\mathbf{R}_n) = \begin{pmatrix} \exp(i\alpha\mathbf{k}_1\mathbf{R}_n) & \dots & 0 \\ \dots & \dots & \dots \\ 0 \dots \exp(i\alpha\mathbf{k}_j\mathbf{R}_n) \dots & & 0 \\ \dots & \dots & \dots \\ 0 \dots 0 \dots & & \exp(i\alpha\mathbf{k}_q\mathbf{R}_n) \end{pmatrix} \tag{5.9}$$

[Here we have used the relation $\mathbf{k}(\alpha^{-1}\mathbf{R_n}) = \alpha\mathbf{k}\mathbf{R_n}$.] From the general properties of unitary representations, we have the relation

$$D(\varepsilon \mid \alpha^{-1}\mathbf{R}_n) = D^{+}(\alpha \mid \mathbf{a})\, D(\varepsilon \mid \mathbf{R}_n)\, D(\alpha \mid \mathbf{a}). \tag{5.10}$$

It is easy to see that if a given diagonal matrix is transformed into another diagonal matrix by means of some unitary transformation, the two matrices can differ only in the order of their diagonal elements. Therefore, if the principal diagonal of the matrix (5.7) includes element $\exp(i\mathbf{k}_1\mathbf{R_n})$, there must also be an element $\exp(i\alpha\mathbf{k}_1\mathbf{R_n})$ for all α of the point group. We can even assert that each diagonal element in (5.7) must have the form $\exp(i\alpha\mathbf{k}_1\mathbf{R_n})$, where α is an element of the point group.*

Then the matrix (5.7) is written as follows:

$$D(\varepsilon \mid \mathbf{R}_n) = \begin{pmatrix} \exp(i\mathbf{k}_1\mathbf{R}_n) & 0 & \dots & 0 \\ 0 & \exp(i\alpha_2\mathbf{k}_1\mathbf{R}_n) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 \dots & \dots 0 & \dots & \exp(i\alpha_q\mathbf{k}_1\mathbf{R}_n) \end{pmatrix}, \tag{5.11}$$

i.e., the n-dimensional matrix $D(\epsilon \mid \mathbf{Rn})$ splits into q diagonal blocks; all other elements are equal to zero. The diagonal blocks themselves are diagonal (in fact, scalar) matrices of order $d = n/q$; $\alpha_1 = \epsilon$, α_2, ... α_q is a set of elements of the point group which take $\mathbf{k}_1$ into $\mathbf{k}_1$, $\mathbf{k}_2$, ..., $\mathbf{k}_q$ respectively, i.e.,

$$\alpha_i\mathbf{k}_1 = \mathbf{k}_i, \tag{5.12}$$

*We omit the proof of this last assertion.

The $\mathbf{k_i}$ correspond to different representations of $\mathfrak{T}$.

For what follows it is convenient to write the matrices of arbitrary elements of a space group in the form

$$D(\alpha \mid \mathbf{a}) = \begin{pmatrix} D_{11}(\alpha \mid \mathbf{a}) & D_{12}(\alpha \mid \mathbf{a}) & \dots & D_{1q}(\alpha \mid \mathbf{a}) \\ D_{21}(\alpha \mid \mathbf{a}) & D_{22}(\alpha \mid \mathbf{a}) & \dots & D_{2q}(\alpha \mid \mathbf{a}) \\ \dots & \dots & \dots & \dots \\ D_{q1}(\alpha \mid \mathbf{a}) & D_{q2}(\alpha \mid \mathbf{a}) & \dots & D_{qq}(\alpha \mid \mathbf{a}) \end{pmatrix}, \tag{5.13}$$

where $D_{ij}(\alpha \mid \mathbf{a})$ (i, j = 1, 2, ..., q) are d-dimensional matrices. In the notation of (5.13) the matrices (5.11) will be

$$D_{ij}(\varepsilon \mid \mathbf{R}_n) = \exp(i\alpha_i\mathbf{k}_1\mathbf{R}_n)\, E_d\delta_{ij}. \tag{5.14}$$

Let us now proceed to study the matrices representing elements which differ from pure translations. First we consider any element $\{\beta \mid \mathbf{b}\}$ of $\mathfrak{G}$ having the property

$$\exp(i\mathbf{k}_1\mathbf{R}_n) = \exp(i\beta\mathbf{k}_1\mathbf{R}_n) \tag{5.15}$$

for all $\mathbf{R_n}$, in other words this means that the relation

$$\beta\mathbf{k}_1 = \mathbf{k}_1 + 2\pi\mathbf{h} \tag{5.16}$$

is valid. We note that the elements of the group $\mathfrak{G}$ which have this property form a group (i.e., they form a subgroup of the group $\mathfrak{G}$). We shall call this group the group of $\mathbf{k}_1$ and denote it by $\mathfrak{K}$. In particular this group contains the whole group $\mathfrak{T}$.

According to (5.8) the element $\{\beta \mid \mathbf{b}\}$ must satisfy the relation

$$D(\varepsilon \mid \mathbf{R}_n)\, D(\beta \mid \mathbf{b}) = D(\beta \mid \mathbf{b})\, D(\varepsilon \mid \beta^{-1}\mathbf{R}_n) \tag{5.17}$$

for all $\mathbf{R_n}$. From this, using (5.13) and (5.14) and comparing the elements in the first columns on right and left, we find

$$\exp(i\alpha_j\mathbf{k}_1\mathbf{R}_n)\, D_{j1}(\beta \mid \mathbf{b}) = D_{j1}(\beta \mid \mathbf{b}) \exp(i\mathbf{k}_1\beta^{-1}\mathbf{R}_n)$$
$$= D_{j1}(\beta \mid \mathbf{b}) \exp(i\mathbf{k}_1\mathbf{R}_n). \tag{5.18}$$

Since the expression (5.18) can be represented as

$$D_{j1}(\beta \mid \mathbf{b})\,[\exp(i\alpha_j\mathbf{k}_1\mathbf{R}_n) - \exp i\mathbf{k}_1\mathbf{R}_n] = 0$$

and if, according to the definition, we replace $\alpha_j\mathbf{k}_1$ by $\mathbf{k}_j$, it follows immediately that $D_{j1}(\beta \mid \mathbf{b})$ is equal to zero if $j \neq 1$. From the unitarity of the transformations it is easy to show also that $D_{ij}(\beta \mid \mathbf{b}) = 0$ for $j \neq 1$. Thus for all $\{\beta \mid \mathbf{b}\}$ belonging to the group $\mathfrak{K}$ we have

$$D(\beta \mid \mathbf{b}) = \left(\begin{array}{c|c} D_{11}(\beta \mid \mathbf{b}) & 0 \\ \hline 0 & | \end{array}\right), \tag{5.19}$$

i.e., the matrices $D_{11}(\beta \mid \mathbf{b})$ form a representation of the group $\mathfrak{K}$.

Now let us consider the matrix corresponding to the element $\{\alpha_j \mid \mathbf{a}_j\}$, for which $\alpha_j\mathbf{k}_1 = \mathbf{k}_j$. From the relation

$$D(\varepsilon \mid \mathbf{R}_n)\, D(\alpha_j \mid \mathbf{a}_j) = D(\alpha_j \mid \mathbf{a}_j)\, D(\varepsilon \mid \alpha_j^{-1}\mathbf{R}_n) \tag{5.20}$$

it follows that

$$\exp(i\alpha_i \mathbf{k}_1 \mathbf{R}_n) D_{i1}(\alpha_j \mid \mathbf{a}_j) = D_{i1}(\alpha_j \mid \mathbf{a}_j) \exp(i\alpha_j \mathbf{k}_1 \mathbf{R}_n) \quad (5.21)$$

for all $\mathbf{R}_n$, i.e., we find that the only non-zero block in the first column of $D(\alpha_j \mid \mathbf{a}_j)$ is the j'th. The matrices representing $D(\alpha_j \mid \mathbf{a}_j)$ can be chosen in such a form that $D_{j1}(\alpha_j \mid \mathbf{a}_j) = \exp(i\alpha_j \mathbf{k}_1 \mathbf{R}_n) E_d \delta_{j1}$.

Thus we have characterized the first row and first column of the matrices representing elements of the type $\{\beta \mid \mathbf{b}\}$, and the first column of matrices representing $\{\alpha_j \mid \mathbf{a}_j\}$, where $\alpha_j \mathbf{k}_1 = \mathbf{k}_j$. This proves to be sufficient to characterize the form of the whole representation $D(\alpha \mid \mathbf{a})$.

First of all we note that by using the subgroup $\mathfrak{K}$ and the elements $\{\alpha_j \mid \mathbf{a}_j\}$ we can make an expansion of $\mathfrak{G}$ into its residue classes, i.e.,

$$\mathfrak{G} = \mathfrak{K} + \{\alpha_2 \mid \mathbf{a}_2\}\mathfrak{K} + \ldots + \{\alpha_q \mid \mathbf{a}_q\}\mathfrak{K}. \quad (5.22)$$

Let us consider the l'th column of $D(\alpha \mid \mathbf{a})$. According to (5.22), for each α there is some α_m of $\alpha_1, \alpha_2, \ldots, \alpha_q$, such that the equation

$$\exp(i\alpha \mathbf{k}_l \mathbf{R}_n) = \exp(i\alpha_m \mathbf{k}_1 \mathbf{R}_n). \quad (5.23)$$

is valid. Multiplying (5.10) on the left by $D(\alpha \mid \mathbf{a})$ and comparing the l'th columns of both sides for the jl'th block, we find

$$D_{jl}(\alpha \mid \mathbf{a}) \exp(i\alpha \mathbf{k}_l \mathbf{R}_m) = \exp(i\mathbf{k}_j \mathbf{R}_m) D_{jl}(\alpha \mid \mathbf{a}). \quad (5.24)$$

It then follows that $D_{jl}(\alpha \mid \mathbf{a}) = 0$ for $j \neq m$. Consequently, the only non-zero block in the l'th column is the m'th, where m is determined by (5.23).

Now we can find an explicit expression for $D_{lm}(\{\alpha \mid \mathbf{a}\})$. On the basis of (5.23) we can write

$$\{\alpha \mid \mathbf{a}\}\{\alpha_l \mid \mathbf{a}_l\} = \{\alpha_m \mid \mathbf{a}_m\}\{\beta \mid \mathbf{b}\},$$

or, considering the matrices, we have:

$$D(\{\alpha \mid \mathbf{a}\}) = D(\{\alpha_m \mid \mathbf{a}_m\}) D(\{\beta \mid \mathbf{b}\}) D^+(\{\alpha_l \mid \mathbf{a}_l\}). \quad (5.25)$$

From this we find for the ml'th block

$$D_{ml}(\{\alpha \mid \mathbf{a}\}) = \sum_{i,j} D_{mi}(\{\alpha_m \mid \mathbf{a}_m\}) D_{ij}(\{\beta \mid \mathbf{b}\}) D^+_{lj}(\{\alpha_l \mid \mathbf{a}_l\})$$

$$= \sum_{i,j} E_d \delta_{1i} D_{ij}(\{\beta \mid \mathbf{b}\}) E_d \delta_{1j} = D_{11}(\{\beta \mid \mathbf{b}\}).$$

Thus we have finally:

$$D_{ml}(\{\alpha \mid \mathbf{a}\}) = D_{11}(\{\beta \mid \mathbf{b}\}). \quad (5.26)$$

Summarizing, we have the following results: Every irreducible representation $D(\alpha \mid \mathbf{a})$ of the space group $\mathfrak{G}$ can be brought to a form in which the invariant subgroup of translations $\mathfrak{T}$ is represented by diagonal matrices. If n is the dimension of the representation, the elements of the diagonal matrices can be arranged so that the first d elements of the matrix $D(\epsilon \mid \mathbf{R}_n)$ have the form $\exp(i\mathbf{k}\mathbf{R}_n)$ for all $\mathbf{R}_n$, while the remaining elements split into $\left(\frac{n}{d} - 1\right)$ groups in each of which there are d elements of the form $\exp(i\alpha j \mathbf{k}\mathbf{R}_n)$ $(j = 1, \ldots, \frac{n}{d} = q)$. Here α_j is the element of the point group which corresponds to $\{\alpha_j \mid \mathbf{a}_j\}$ of the space group. To this block arrangement of the matrices for translations there corresponds a block arrangement of the matrices $D(\alpha \mid \mathbf{a})$ in the irreducible representation of $\mathfrak{G}$, where the matrices $D(\alpha \mid \mathbf{a})$ break up into q blocks of dimension d, which are denoted by $D_{ij}(\alpha \mid \mathbf{a})$. The elements $\{\beta \mid \mathbf{b}\}$ which have the property

$$\exp(i\beta \mathbf{k}\mathbf{R}_n) = \exp(i\mathbf{k}\mathbf{R}_n), \quad (5.27)$$

form a group $\mathfrak{K}$, which contains the group $\mathfrak{T}$. The matrices $D_{11}(\beta \mid \mathbf{b})$ give an irreducible representation of $\mathfrak{K}$.

The elements $\{\alpha_i \mid \mathbf{a}_i\}$ and the subgroup $\mathfrak{K}$ can be used for the expansion of $\mathfrak{G}$ in left residue classes.

For any element $\{\alpha \mid \mathbf{a}\}$ of the group $\mathfrak{G}$ and any α_l we can find an α_m such that

$$\exp(i\alpha\alpha_l \mathbf{k}\mathbf{R}_n) = \exp(i\alpha_m \mathbf{k}\mathbf{R}_n). \quad (5.28)$$

We then find that the only non-zero block in $D(\alpha \mid \mathbf{a})$ is the m'th block in the l'th column and that the matrix which stands at this place is $D_{11}(\beta \mid \mathbf{b})$. The only non-zero block in the first column of blocks for $\{\alpha_j \mid \mathbf{a}_j\}$ is the j'th.

Finally, it can be shown that any representation of $\mathfrak{G}$ of the form considered above is an irreducible representation of the group $\mathfrak{G}$. In this proof one makes use of the fact that a representation is irreducible if the only matrix which commutes with it is a scalar matrix.

6. REPRESENTATIONS OF SPACE GROUPS (POSSIBLE SIMPLIFICATIONS)[9]

Representations are most simply constructed by using basis functions. Suppose that the d orthogonal functions $u'_1, \ldots, u'_d$, which on translation through $\mathbf{R}_n$ are multiplied by $\exp(i\mathbf{k}\mathbf{R}_n)$, form an irreducible representation $\mathfrak{K}$ of the group of elements $\{\beta \mid \mathbf{b}\}$, for which $\exp(i\beta\mathbf{k}\mathbf{R}_n) = \exp(i\mathbf{k}\mathbf{R}_n)$ for all $\mathbf{R}_n$. Then the n = dq functions

$$u^i_j = \{\alpha_i \mid \mathbf{a}_i\} u_1, \qquad i = 1, \ldots, q,$$
$$\{\alpha_1 \mid \mathbf{a}_1\} = \{\varepsilon \mid 0\}, \qquad j = 1, \ldots, d \quad (6.1)$$

form the basis of an irreducible representation of the space group $\mathfrak{G}$. Here $\{\alpha_i \mid \mathbf{a}_i\}$ are elements of $\mathfrak{G}$ for which the expansion (5.22) holds. Using the results stated above, we can understand the procedure for finding all irreducible representations of the space group. We first select a vector $\mathbf{k}$ within or on the boundary of the Brillouin zone. The rotational part of the operator $\{\beta \mid \mathbf{b}\}$ of the group $\mathfrak{G}$ must satisfy the condition $\beta\mathbf{k} = \mathbf{k} + 2\pi\mathbf{h}$. Using this relation we construct all the irreducible representations of this group of elements which have the property that the diagonal elements of the matrices representing pure translations have the form $\exp(i\mathbf{k}\mathbf{R}_n)$. From the preceding treatment it follows that this will lead to all irreducible

representations of $\mathfrak{G}$ which are associated with the vector $\mathbf{k}$. We then obtain all the irreducible representations of the group $\mathfrak{G}$ since we assume that the vector $\mathbf{k}$ runs through all values in the Brillouin zone.

In finding the corresponding irreducible representations of $\mathfrak{R}$ we can make some simplifications. Let the rotational parts α of the space group whose elements are $\{\alpha|\mathbf{a}\}$ form a point group $\mathfrak{G}_0$. This group is one of the 32 admissible point groups, whose irreducible representations are well known. It is obvious that the rotational parts of the operators $\{\beta|\mathbf{b}\}$ in $\mathfrak{R}$ also must form a subgroup of the group $\mathfrak{G}_0$. We shall denote it $\mathfrak{G}_0(\mathbf{k})$. First we consider points inside the Brillouin zone. For any point inside the Brillouin zone the rotational operators in $\mathfrak{R}$ satisfy the condition $\beta\mathbf{k} = \mathbf{k}$. We denote one of the irreducible representations of the group $G_0(\mathbf{k})$ by $\Gamma(\beta)$. We show that the irreducible representation of $\mathfrak{R}$ is determined by the expression

$$D_{11}(\{\beta|\mathbf{b}_j\}) = \exp i\mathbf{k}\mathbf{b}\Gamma(\beta), \tag{6.2}$$

if $\{\beta|\mathbf{b}\}$ belongs to $\mathfrak{R}$. The product of two operators $\{\beta|\mathbf{b}\}$ and $\{\beta'|\mathbf{b}'\}$ in $\mathfrak{R}$ is equal to $\{\beta\beta'|\beta\mathbf{b}' + \mathbf{b}\}$. Multiplying the matrices representing these operators, we find

$$D_{11}\{\beta|\mathbf{b}\}D_{11}\{\beta'|\mathbf{b}'\} = \exp i\mathbf{k}\mathbf{b}\exp i\mathbf{k}\mathbf{b}'\Gamma(\beta)\Gamma(\beta')$$
$$= \exp i\mathbf{k}(\mathbf{b}+\mathbf{b}')\Gamma(\beta\beta'). \tag{6.3}$$

The matrix representing the product of these operators is given by

$$D_{11}(\{\beta\beta'|\beta\mathbf{b}'+\mathbf{b}\}) = \exp i\mathbf{k}(\beta\mathbf{b}'+\mathbf{b})\Gamma(\beta\beta')$$
$$= \exp(i\beta^{-1}\mathbf{k}\mathbf{b}')\exp(i\mathbf{k}\mathbf{b})\Gamma(\beta\beta') = \exp i\mathbf{k}(\mathbf{b}+\mathbf{b}')\Gamma(\beta\beta'). \tag{6.4}$$

From these relations it follows that (6.2) forms a representation of the group $\mathfrak{R}$, and this representation is irreducible as a consequence of the irreducibility of the representation $\Gamma(\beta)$.

Thus, knowing all the representations of the 32 point groups, we can find all the irreducible representations of all the space groups which are associated with the vector $\mathbf{k}$ in the interior of the Brillouin zone.

Now we consider a point on the surface of the Brillouin zone. At these points the relation $\beta\mathbf{k} = \mathbf{k} + 2\pi\mathbf{h}$ may hold. In this case the results just obtained for finding irreducible representations of space groups will in general not be true. However, they remain true for the simple space groups in which $\mathbf{v}(\alpha) \equiv 0$. In these groups $\mathbf{a}$ and $\mathbf{b}$ are admissible translations. Again, we choose $D_{11}(\{\beta|\mathbf{b}\})$ in the form $\exp i\mathbf{k}\mathbf{b}\Gamma(\beta)$, where now $\{\beta|\mathbf{b}\}$ is an operator for which $\beta\mathbf{k} = \mathbf{k} + 2\pi\mathbf{h}$. We again obtain an irreducible representation of $\mathfrak{R}$. Equation (6.3) remains unchanged, but (6.4) is proved differently.

In this case we have

$$D_{11}(\{\beta\beta'|\beta\mathbf{b}'+\mathbf{b}\}) = \exp i\mathbf{k}(\beta\mathbf{b}'+\mathbf{b})\Gamma(\beta\beta')$$
$$= \exp i\beta^{-1}\mathbf{k}\mathbf{b}'\exp i\mathbf{k}\mathbf{b}\Gamma(\beta\beta') = \exp i(\mathbf{k}+2\pi\mathbf{h}_j)\mathbf{b}'\exp i\mathbf{k}\mathbf{b}\Gamma(\beta\beta'$$
$$= \exp i\mathbf{k}(\mathbf{b}+\mathbf{b}')\Gamma(\beta\beta'). \tag{6.5}$$

Here we have used the fact that $\mathbf{b}'$ is an admissible translation, and that therefore $\exp(i\mathbf{b}'2\pi\mathbf{h}_j) = 1$. Consequently we can find all the irreducible representations of $\mathfrak{G}$ for simple space groups at points on the surface of the Brillouin zone. To do this we must know the irreducible representations of the 32 point groups.

For points in $\mathbf{k}$ space on the boundaries of the Brillouin zone whose group $\mathfrak{R}$ contains operators having non-primitive translations, the situation is more complicated. However, in this case one can find the irreducible representations of the simplest space groups of this type associated with points on the surface of the Brillouin zone by using the special properties of each group.

7. BAND THEORY OF SOLIDS FROM THE POINT OF VIEW OF GROUP THEORY

The theory of groups enables us to treat the band theory of solids from a general point of view. Such a treatment based on the consideration of all symmetry properties of a crystal lattice leads to a clarification of fine features in the energy spectrum of the electron (for example, degeneracy), which is impossible when one uses the simple band theory which considers only the translational symmetry of the crystal.

We know that the wave functions and energy levels corresponding to different quantum states of an electron in a crystal are determined from the solutions of the Schrödinger equation, which in the one-electron approximation, in the usual notation, has the form

$$[-\Delta + V(\mathbf{r})]\psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{7.1}$$

The most characteristic feature of the potential energy $\mathbf{V}(\mathbf{r})$ for an electron in a crystal is its symmetry: the potential of the electron must have the same symmetry as the crystal itself. In particular, the function $\mathbf{V}(\mathbf{r})$ must be invariant under the action of any translation which brings the crystal into coincidence with itself. Consequently,

$$V(\mathbf{r}+\mathbf{R}_n) = V(\mathbf{r}), \tag{7.2}$$

where

$$\mathbf{R}_n = n_1\mathbf{t}_1 + n_2\mathbf{t}_2 + n_3\mathbf{t}_3 \tag{7.3}$$

and $\mathbf{t}_1$, $\mathbf{t}_2$ and $\mathbf{t}_3$ are fundamental lattice vectors.

In order to avoid the complications which arise in an attempt to impose real boundary conditions on the solutions, one usually uses a so-called "cyclic crystal." A cyclic crystal is nothing other than part of an infinite crystal, contained within a parallelepiped with edged $N\mathbf{t}_1$, $N\mathbf{t}_2$, $N\mathbf{t}_3$, whose center coincides with the origin of coordinates. The number N is an arbitrarily large integer. It is easy to show that a cyclic crystal contains N^3 unit cells each with volume $\mathbf{t}_1[\mathbf{t}_2 \times \mathbf{t}_3]$. The cyclic boundary conditions require that any two points in space which differ by a vector $N\mathbf{t}$, be regarded as physically equivalent, i.e.,

$$\psi(\mathbf{r}) \equiv \psi(\mathbf{r}+N\mathbf{t}). \tag{7.4}$$

Thus, the mathematical problem is formulated as follows. We are required to find solutions of Eq. (7.1) with the potential $\mathbf{V}(\mathbf{r})$, having the complete spatial symmetry of the lattice, and satisfying the boundary conditions (7.4).

As was first shown by Bloch, every solution of the problem thus formulated must have the form:

$$\psi(\mathbf{k}, \mathbf{r}) = \exp(i\mathbf{kr})\, u(\mathbf{k}, \mathbf{r}), \tag{7.5}$$

where $u(\mathbf{k}, \mathbf{r})$ is a periodic function with the period of the lattice

$$u(\mathbf{k}, \mathbf{r}) = u(\mathbf{k}, \mathbf{r}+\mathbf{R}_n) \tag{7.6}$$

and $\mathbf{k}$ is the wave vector or quasimomentum of the electron in the crystal. Condition (7.4) can be satisfied only when

$$\mathbf{k} = \frac{2\pi}{N}(\kappa_1\mathbf{b}_1 + \kappa_2\mathbf{b}_2 + \kappa_3\mathbf{b}_3), \tag{7.7}$$

where κ_1, κ_2, κ_3 are integers and $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ are the fundamental vectors of the reciprocal lattice, defined by the equations

$$\mathbf{t}_i\mathbf{b}_j = \delta_{ij}, \qquad i, j = 1, 2, 3.$$

It can be shown that two Bloch functions whose wave vectors differ by 2π times a vector of the reciprocal lattice

$$\mathbf{h} = h_1\mathbf{b}_1 + h_2\mathbf{b}_2 + h_3\mathbf{b}_3 \qquad (h_i - \text{integers}$$

are physically equivalent. In order not to treat physically equivalent solutions, we should limit the range of variation of $\mathbf{k}$ in the reciprocal space. This is most simply done by requiring that the vector $\mathbf{k}$ lie in the central unit cell of the reciprocal lattice. For this it is sufficient to set

$$\mathbf{k} = k_1\mathbf{b}_1 + k_2\mathbf{b}_2 + k_3\mathbf{b}_3 \qquad \left(-\frac{\pi}{a} \leqslant k_i \leqslant \frac{\pi}{a}\right). \tag{7.8}$$

Wave vectors whose values are in the central unit cell of the reciprocal lattice are called reduced wave vectors, and the central unit cell itself is called the reduced Brillouin zone. Consequently, every Bloch function can be characterized by a certain wave vector $\mathbf{k}$. It then follows immediately that the eigenvalues of (7.1) will also be functions of the wave vector $\mathbf{k}$. Actually the wave equation (7.1) does not have just a single eigenvalue and eigenfunction for each admissible value $\mathbf{k} = \mathbf{k}_0$. We shall number the various quantum states corresponding to the same value of the reduced wave vector by an index n, where this index n can be assigned to the various quantum states in order of increasing energy, i.e., the states will be numbered by values of n so that the equation

$$E_1(\mathbf{k}_0) \leqslant E_2(\mathbf{k}_0) \leqslant \ldots \leqslant E_j(\mathbf{k}_0) \leqslant \ldots$$

holds. Having done this for all values of $\mathbf{k}$, we obtain a system of energy functions

$$E_1(\mathbf{k}) \leqslant E_2(\mathbf{k}) \leqslant \ldots \tag{7.9}$$

As $\mathbf{k}$ varies within the reduced Brillouin zone, each of the functions determines a certain hypersurface in four-dimensional space. The projections of these surfaces onto the coordinate planes give the so-called energy bands. Thus each state is completely defined by the assignment of two quantities: the reduced wave vector and the number of the band, i.e.,

$$\psi(\mathbf{r}) = \psi_n(\mathbf{k}, \mathbf{r}) = \exp(i\mathbf{kr})\, u_n(\mathbf{k}, \mathbf{r}), \tag{7.10}$$

$$E = E_n(\mathbf{k}). \tag{7.11}$$

Now let us consider in a little more detail the group-theoretical treatment of the problem. As we have already said, the Schrödinger equation for the determination of electron states in a crystal must be invariant with respect to transformations which bring the crystal into coincidence with itself. In the case of an infinite crystal, such transformations form an infinite group which is called the spatial symmetry group. From this full space group we can always select the infinite abelian subgroup of translations. The basis functions for irreducible representations of this subgroup are the Bloch functions (7.10), and the action of a translation operator reduces to multiplication by $\exp(-i\mathbf{kt})$, i.e.,

$$\{\varepsilon \mid \mathbf{t}\}\, \psi_n(\mathbf{k}, \mathbf{r}) = \exp(-i\mathbf{kt})\, \psi_n(\mathbf{k}, \mathbf{r}). \tag{7.12}$$

Using these functions we can construct irreducible representations for the whole space group. In the following we shall temporarily disregard complicated symmetry elements, i.e., we shall assume that every operation of the group, $\{\alpha \mid \mathbf{t}\}$, can be represented in the form $\{\epsilon \mid \mathbf{t}\}\ \{\alpha \mid 0\}$ where $\{\epsilon \mid \mathbf{t}\}$ and $\{\alpha \mid 0\}$ are operations of the group. Consequently in studying space groups of this type we need consider, in addition to the translations, only the operations of the corresponding point group. We know that the action of such operators has the effect of taking a function with the wave vector $\mathbf{k}$ into a wave function with the wave vector $\mathbf{k}'$ which is obtained from $\mathbf{k}$ by applying this operator, i.e.,

$$\{\alpha \mid 0\}\, \psi_n(\mathbf{k}, \mathbf{r}) = \psi_n(\alpha\mathbf{k}, \mathbf{r}). \tag{7.13}$$

In fact, let us choose a function $\psi_n(\mathbf{k}, \mathbf{r})$ and apply the element $\{\alpha \mid 0\}$ to it:

$$\{\alpha \mid 0\}\, \psi_n(\mathbf{k}, \mathbf{r}) = \exp(i\mathbf{k}\alpha^{-1}\mathbf{r})\, u_n(\mathbf{k}\alpha^{-1}\mathbf{r}) = \exp(i\alpha\mathbf{kr})\, u_n'(\mathbf{k}, \mathbf{r}).$$

On the other hand,

$$\psi_n(\alpha\mathbf{k}, \mathbf{r}) = \exp(i\alpha\mathbf{kr})\, u_n(\alpha\mathbf{k}, \mathbf{r})$$

and consequently formula (7.13) is proven.

If the vector $\mathbf{k}$ lies in a general direction, it then follows that in addition to each function $\psi_n(\mathbf{k}, \mathbf{r})$, the function $\psi_n(\alpha\mathbf{k}, \mathbf{r})$ belongs to the same energy value. This on the one hand determines the symmetry of $E_n(\mathbf{k})$ and on the other hand shows the degeneracy of

the particular energy value, since

$$E_n(\mathbf{k}) \equiv E_n(\alpha\mathbf{k}). \tag{7.14}$$

It should be mentioned that since we are considering the problem of the spectrum of a multidimensional system we always have infinite degeneracy, since the equation

$$E_n(\mathbf{k}) = \text{const}$$

is satisfied by an infinite number of states, in fact by a whole surface in $\mathbf{k}$ space. However, even for a very slight change in the potential $\mathbf{V}(\mathbf{r})$ we obtain in general a different surface

$$E'_n(\mathbf{k}) = \text{const},$$

for which, if the symmetry remains the same, the equation

$$E'_n(\mathbf{k}) \equiv E'_n(\alpha\mathbf{k})$$

will still be satisfied. We must therefore distinguish between degeneracy which is associated with the multidimensionality of the problem, which we shall call nonessential, and degeneracy which is caused by the symmetry of the system. Here it also necessary to state that if the vector $\mathbf{k}$ lies in a general direction, the states which are degenerate with one another have different values of the reduced wave vector, and therefore the assignment of $E_n(\mathbf{k})$ and $\mathbf{k}$ completely determines the state of the system.*

If, however, the vector $\mathbf{k}$ does not lie in an arbitrary direction, but there are elements of symmetry which leave $\mathbf{k}$ invariant, the picture becomes somewhat more complicated. The degeneracy associated with the symmetry of the system is lifted partially in this case, i.e., the functions $\psi_n(\mathbf{k}, \mathbf{r})$ and $\psi_n(\alpha\mathbf{k}, \mathbf{r})$ with

$$\alpha\mathbf{k} - \mathbf{k} = 2\pi\mathbf{h}$$

may belong to different energy values. Nevertheless, since the degeneracy is not lifted completely, it turns out that functions degenerate with one another have the same reduced wave vectors $\mathbf{k}$. Therefore, for a complete characterization of states of this type it is necessary to assign, in addition to the values $E_n(\mathbf{k})$ and $\mathbf{k}$, the values of certain other quantum numbers. These arguments can be carried through with a few changes also in the case where the group contains complex elements.

8. GENERAL METHOD FOR INVESTIGATING THE ENERGY SPECTRUM AND CLASSIFICATION OF STATES (LINEAR CHAIN)[10]

When an electron is in the field of a crystal, the Schrödinger equation for it is left invariant only by the symmetry operations of this field. In other words, the infinitesimal transformations will no longer be included in the motions leaving the Schrödinger equation invariant, and this in turn will lead to a breakdown of the usual conservation laws. Now the conserved quantities will be certain quasiquantities, associated with the symmetry of the particular crystalline field. A knowledge of the symmetry group of the lattice enables us to determine these quasiquantities, to select from them a set which are simultaneously measurable, investigate in detail the problem of degeneracy, state selection rules, etc. The whole treatment can be carried out in general form and results are obtained which supplement and clarify the material of the preceding paragraph, but we shall restrict ourselves to some examples and present them in the greatest possible detail in order best to make clear the method of investigation.

The equation for the determination of electronic states in a one-dimensional periodic field is written in the form

$$\psi''(x) + [E - V(x)]\psi(x) = 0, \tag{8.1}$$

where the potential $\mathbf{V}(x)$ has the properties $\mathbf{V}(x + a) = \mathbf{V}(x)$ and $\mathbf{V}(-x) = \mathbf{V}(x)$. Consequently the symmetry group for this problem is the space group consisting of translations along the x-axis: $\{\epsilon \mid t\}$ (here $t = ap$ with $p = 0, \pm 1, \ldots$), the inversion $\{i \mid 0\}$ in the origin, and the products of these elements $\{i \mid 0\} \times \{\epsilon \mid t\}$. We can select as the generators of this space group the elements

$$\{\varepsilon \mid a\} \text{ and } \{i \mid 0\}, \tag{8.2}$$

whose eigenvalues are

$$\exp(ika) \text{ and } \pm 1. \tag{8.3}$$

The first operator in (8.2) is easily expressed in terms of the infinitesimal displacement $\frac{\partial}{\partial x}$ along the x axis. It has the form $\exp\left(a\frac{\partial}{\partial x}\right)$ or, since $\frac{\partial}{\partial x} = i\hat{p}$,

$$\exp(ia\hat{p}). \tag{8.4}$$

Thus we can conclude that the operator (8.4) represents some physical quantity closely related to the electron momentum. Since we may write

$$\exp(ia\hat{p})\psi = \exp(iak)\psi, \tag{8.5}$$

the vector $\mathbf{k}$ appearing in this equation is called the quasimomentum of the electron in the crystal. We note that replacing k by $k + \frac{2\pi n}{a}$ does not change the eigenvalue of the operator (8.4), so that one can introduce the reduced wave vector k which varies between $-\frac{\pi}{a}$ and $+\frac{\pi}{a}$. For the present, we shall not do this.

The eigenvalues of the inversion operator characterize

*This is equivalent to specifying the following quantum numbers: the number n of energy band and the components of quasimomentum k_x, k_y, k_z.

the parity of the states. It should be noted that the operators (8.2) are not commutative in general, so that the quasimomentum of the electron and the parity of the state cannot be determined simultaneously.

Now we separate the elements of the group into classes of conjugate elements. To do this we must transform each element of the group by all others. It is easy to see that one obtains the following types of conjugate elements:

I $\{\varepsilon \mid \mathbf{t}\}$, $\{\varepsilon \mid i\mathbf{t}\}$,
II $\{i \mid 0\}\ \{\varepsilon \mid i\mathbf{u} - \mathbf{u} + \mathbf{t}\}$, $\{i \mid 0\}\ \{\varepsilon \mid i\mathbf{u} - \mathbf{u} + i\mathbf{t}\}$.

From this we easily find the classes of conjugate elements: $\hat{K}_1 = \{\epsilon \mid 0\}$, $\hat{K}_2$ consists of elements of the form $\{i \mid 0\}\ \{\epsilon \mid 2qa\}$, $\hat{K}_3$ of elements $\{i \mid 0\} \times \{\epsilon \mid (2q+1)a\}$, and, in addition, there are also an infinite number of classes of the type $\hat{K}_p = \{\epsilon \mid \pm pa\}$ with $p = 1, 2, \ldots$.

We form the operators for the various classes:

$$\hat{K}_1 = \{\varepsilon \mid 0\},$$

$$\hat{K}_2 = \{i \mid 0\} \sum_q \{\varepsilon \mid 2qa\} / \sum_q 1,$$

$$\hat{K}_3 = \{i \mid 0\}\ \{\varepsilon \mid a\} \sum_q \{\varepsilon \mid 2qa\} / \sum_q 1,$$

$$\hat{K}_p = \frac{1}{2}[\{\varepsilon \mid pa\} + \{\varepsilon \mid -pa\}]$$

and determine their eigenvalues:

$$\hat{K}_1 \psi(k, x) = \psi(k, x),$$

$$\hat{K}_2 \psi(k, x) = \{i \mid 0\}\, \delta\left(k, \frac{\pi n}{a}\right) \psi(k, x) = \varkappa \psi(k, x),$$

$$\hat{K}_3 \psi(k, x) = \{i \mid 0\} \exp(ika)\, \delta\left(k, \frac{\pi n}{a}\right) \psi(k, x)$$

$$= \varkappa \exp(ika)\, \psi(k, x),$$

$$\hat{K}_p \psi(k, x) = \cos pka\, \psi(k, x),$$

where $\kappa = 0$ for $k \neq \frac{\pi n}{a}$ and $\kappa = \pm 1$ for $k = \frac{\pi n}{a}$.

Although we use the fact that the eigenfunctions have the Bloch form when we calculate the eigenvalues of the operators for the classes, this is not necessary, since the eigenvalues are determined by the characters of the representations and consequently do not depend on the form of the basis functions.

In accordance with the general rule, the energy will depend on the eigenvalues of the operators for the classes, i.e.,

$$E = E(\varkappa, \cos ka). \tag{8.6}$$

Since the operators (8.2) do not commute with one another, the energy terms will in general be degenerate. To enumerate the degenerate states we can use either the eigenvalues of the operator $\{\epsilon \mid a\}$, or the eigenvalues of the operator $\{i \mid 0\}$; in the first case we will fix the vector $\mathbf{k}$, and in the second the parity of the state.

Analyzing the table of eigenvalues of the class operators we can draw the following conclusions:

1. When $k \neq \frac{\pi n}{a}$ the energy is a function of $\cos(ka)$ and, of course, also depends on the specific form of the potential:

$$E = E(\cos ka). \tag{8.7}$$

States with a given energy are doubly degenerate since there correspond to each level two functions which, in particular, can be chosen in the form:

$$\psi^{(1)}(k, x) = \exp(ikx)\, u(k, x),$$
$$\psi^{(2)}(k, x) = \exp(-ikx)\, u(-k, x).$$

2. When $k = \frac{\pi n}{a}$ the energy jumps discontinuously since at such points its dependence on the eigenvalues κ makes itself felt. In other words, at these points we can simultaneously determine the eigenvalues of both of the operators, the quasimomentum $\mathbf{k}$ and the parity of the state. This can be shown most simply in matrix form. If we assume that the eigenfunctions are of the Bloch type, the operator for translation through distance a along the x axis is written in matrix representation in the form

$$\begin{pmatrix} \exp(ika) & 0 \\ 0 & \exp(-ika) \end{pmatrix},$$

and the inversion operator as

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

When $k = \frac{\pi n}{a}$ these operators commute and consequently can be brought simultaneously to principal axes.

Thus when $k \neq \frac{\pi n}{a}$ the energy is a continuous function of $\mathbf{k}$ and jumps at $k = \frac{\pi n}{a}$, i.e., the energy spectrum has a band structure. Introducing the reduced wave vector $\mathbf{k}$, which varies between $-\frac{\pi}{a}$ and $+\frac{\pi}{a}$, and the number of the energy band n, we arrive at the usual scheme of the band theory, i.e.,

$$E = E_n(\cos ka),$$
$$\psi_n^{(1)}(k, x) = \exp(ikx)\, u_n(k, x),$$
$$\psi_n^{(2)}(k, x) = \exp(-ikx)\, u_n(-k, x).$$

As a supplement to the usual band theory, we obtain the result that states corresponding to edges of bands should be given a definite parity.

9. ELECTRONS IN A FIELD OF CUBIC SYMMETRY

The investigations of the energy spectrum of electrons in a crystal, when we take account of the full spatial symmetry of the lattice, can be carried out exactly as for the two preceding cases. It is easy to obtain the result that the conserved quantities for the

electron will be the quasi-angular momentum, the quasimomentum and the parity.

The energy spectrum will have a banded structure; within a band the energy is a function of the quantities:

$$\begin{aligned} &\cos k_1 a + \cos k_2 a + \cos k_3 a, \\ &\cos k_1 a \cos k_2 a + \cos k_2 a \cos k_3 a + \cos k_3 a \cos k_1 a, \\ &\cos k_1 a \cos k_2 a \cos k_3 a. \end{aligned} \tag{9.1}$$

The states which are degenerate with one another can be enumerated, for example, by the values of the wave vector $\mathbf{k}$.

It is of considerable interest to study the phenomenon of overlapping of energy bands which occurs in this case. Before we proceed to the study of particular problems let us make some general comments. We already know that symmetry operations satisfying the equation $\alpha\mathbf{k} - \mathbf{k} = 2\pi\mathbf{h}$ form a group which we call the group of the wave vector. A wave function $\psi_n(\mathbf{k}, \mathbf{r})$ with wave vector $\mathbf{k}$ either remains invariant under the action of the transformations of the group of $\mathbf{k}$, or is transformed into some other wave function with the same wave vector. In the first case, there will be only one wave function with wave vector $\mathbf{k}$; in the second there will be several wave functions which, under the action of the transformations of the group of $\mathbf{k}$, transform according to some irreducible representation of that group. When the group of the wave vector contains transformations other than the identity, the operators corresponding to elements of the group will commute, for functions with a given $\mathbf{k}$, with the translation operators, and therefore their eigenvalues can be determined simultaneously with $\mathbf{k}$.

When several (for example, s) wave functions have the same wave vector $\mathbf{k}$, and if we choose the wave vector κ so that $\mathbf{k} + \kappa$ is in a general direction, there will appear s wave functions with energies close to $E_n(\mathbf{k})$. But, since for wave vectors which lie in general directions it can not happen that two different wave functions with the same wave vector belong to the same energy, they all must belong to different energy bands which for small values of κ are close to one another and touch at the point $\mathbf{k}$. In general, one should also remember two further points. First κ may be such that $\mathbf{k} + \kappa$ still has the group of $\mathbf{k}$. In this case, the touching of the bands will occur as before. Secondly the group of $\mathbf{k} + \kappa$ may be a subgroup of the group of $\mathbf{k}$, but still contain transformations other than the identity. Then, in the transition from $\mathbf{k}$ to $\mathbf{k} + \kappa$, the energy bands will be partially split.

This same argument can be carried out in a different form. The Schrödinger equation for the electron in a crystal (7.1) remains invariant under all the symmetry transformations of the crystal, translations, rotations, reflections, etc. The eigenfunctions can always be chosen in the Bloch form (7.5) and if we substitute them in equation (7.1), we have for the periodic part $U(\mathbf{k}, \mathbf{r})$

$$[-(\nabla - i\mathbf{k})^2 + V(\mathbf{r})]\, u(\mathbf{k}, \mathbf{r}) = E(\mathbf{k})\, u(\mathbf{k}, \mathbf{r}). \tag{9.2}$$

The operator

$$\hat{H}(\mathbf{k}, \mathbf{r}) = [-(\nabla - i\mathbf{k})^2 + V(\mathbf{r})] \tag{9.3}$$

is still invariant under translations, but the application of the point symmetry operators changes $\hat{H}(\mathbf{k}, \mathbf{r})$ to $\hat{H}(\alpha\mathbf{k}, \mathbf{r})$. It is obvious that

$$\hat{H}(\mathbf{k}, \mathbf{r}) \equiv \hat{H}(\alpha\mathbf{k}, \mathbf{r}) \tag{9.4}$$

only when $\{\alpha \mid 0\}$ belongs to the group of the particular wave vector $\mathbf{k}$.

Suppose now that the group of the wave vector $\mathbf{k}$ is the group $\mathfrak{G}(\mathbf{k})$. Then the operator (9.3) is left invariant by the elements of $\mathfrak{G}(\mathbf{k})$ so that the eigenfunctions $u(\mathbf{k}, \mathbf{r})$ and eigenvalues $E(\mathbf{k})$ can be classified according to the irreducible representations of this group. In particular, the dimensionality of the representation is equal to the degree of degeneracy of the particular energy value $E(\mathbf{k})$.

The shift from the point $\mathbf{k}$ to the point $\mathbf{k} + \kappa$ corresponds to a shift from the group $\mathfrak{G}(\mathbf{k})$ to the group $\mathfrak{G}(\mathbf{k} + \kappa)$, and if the shift occurs from a point with higher symmetry to a point with lower symmetry, the group $\mathfrak{G}(\mathbf{k} + \kappa)$ is a subgroup of $\mathfrak{G}(\mathbf{k})$. In this case, the irreducible representations of $\mathfrak{G}(\mathbf{k})$ will in general be reducible for $\mathfrak{G}(\mathbf{k} + \kappa)$ and can be expanded into irreducible representations of $\mathfrak{G}(\mathbf{k} + \kappa)$, which shows that there will be a splitting of the energy bands touching at the point $\mathbf{k}$ when we shift to the point $(\mathbf{k} + \kappa)$. Such a splitting can either be absent if the symmetry is not reduced, or can be complete if the symmetry is lowered substantially.

Let us now analyze this in more detail for the example of a simple cubic lattice. In this case, the Brillouin zone is a cube with edge $\frac{2\pi}{a}$, and those wave vectors $\mathbf{k}$ which lie along symmetry elements or end on the surface of the Brillouin zone will have symmetry groups different from the identity.

Obviously, the groups of the wave vectors must be subgroups of the full cubic group $\mathfrak{O}_h$. It is therefore sufficient to find all the subgroups of this group and to see for which $\mathbf{k}$ the equation

$$\alpha\mathbf{k} - \mathbf{k} = 2\pi\mathbf{h},$$

is valid as $\{\alpha \mid 0\}$ runs through all the elements of the particular subgroup.

It is easy to see, for example, that the group of the vector $\mathbf{k} = (000)$ coincides with $\mathfrak{O}_h$. Analyzing its character table (Table I), we see that for a given $\mathbf{k}$ we can have non-degenerate, doubly-generate and triply-degenerate states. Let us see what happens to these levels if we shift from the point $\mathbf{k} = (000)$ to the point $\mathbf{k} = (00z)$. The group of this wave vector is the group $\mathfrak{C}_{4v}$ with characters given in Table II. Resolving the irreducible representations of $\mathfrak{O}_h$ into

TABLE I

$\mathfrak{O}_h$	E	$8C_3$	$3C_2$	$6C_2$	$6C_4$	I	$8C_3I$	$3C_2I$	$6C_2I$	$6C_4I$
Γ_1	1	1	1	1	1	1	1	1	1	1
Γ_2	1	1	1	−1	−1	1	1	1	−1	−1
Γ_3	2	−1	2	0	0	2	−1	2	0	0
Γ_4	3	0	−1	−1	−1	3	0	−1	1	−1
Γ_5	3	0	−1	1	1	3	0	−1	−1	1
Γ_1'	1	1	1	1	1	−1	−1	−1	−1	−1
Γ_2'	1	1	1	−1	−1	−1	−1	−1	1	1
Γ_3'	2	−1	2	0	0	−2	1	−2	0	0
Γ_4'	3	0	−1	1	−1	−3	0	1	−1	1
Γ_5'	3	0	−1	−1	1	−3	0	1	1	−1

TABLE II

$\mathfrak{C}_{4v}$	E	C_2	$2C_4$	$2\sigma_v$	$2\sigma_r'$
γ_1	1	1	1	1	1
γ_2	1	1	1	−1	−1
γ_3	1	1	−1	1	−1
γ_4	1	1	−1	−1	1
γ_5	2	−2	0	0	0

irreducible representations of $\mathfrak{C}_{4V}$ we obtain the compatibility table (Table III) which shows how the splitting of the bands occurs when we shift from a center of symmetry to an axis of fourfold mirror-rotation symmetry. A similar treatment can also be carried out for any other point.

10. CHANGES IN THE ENERGY SPECTRUM AND STATES WHEN SPIN IS INCLUDED. DOUBLE GROUPS.

The Schrödinger equation for an electron in a crystal, when we include the spin-orbit interaction, has the form

$$\left[-\Delta+V(\mathbf{r})-\frac{i}{2mc^2}(\nabla V(\mathbf{r})\times\nabla\sigma)\right]\psi(\mathbf{r})=E\psi(\mathbf{r}), \quad (10.1)$$

where $-\frac{i}{2mc^2}[\nabla V(\mathbf{r})\times\nabla\sigma]$ is the energy of the spin-orbit interaction and σ is the spin operator.

In this case the Hamiltonian

$$\hat{H}=\left[-\Delta+V(\mathbf{r})-\frac{i}{2mc^2}(\nabla V(\mathbf{r})\times\nabla\sigma)\right] \quad (10.2)$$

(and consequently also the wave functions) still has the translational and point symmetry of the lattice. We can therefore, as before, find the symmetry properties of the function ψ by considering the space group of the crystal. The only difference is that now we must deal with two-valued representations of the symmetry group.

So far, we have used coordinate functions as the basis of our representations. We must now generalize somewhat when we consider the electron spin and consequently have operators which act on spinors.

The transformation properties of spinors under the action of rotations are given by the matrices of the unimodular group which can be expressed in the form $\pm U(R_S)$, where

$$U_{11}(R_s)=U_{22}^*(R_s)=\exp\left[-\frac{i(\varphi+\psi)}{2}\right]\cos\frac{\theta}{2},$$
$$U_{12}(R_s)=U_{21}^*(R_s)=\exp\left[i\frac{\varphi+\psi}{2}\right]\sin\frac{\theta}{2}.$$

This representation of the rotation group is two-valued, since matrices with the reversed sign for all elements are also isomorphic to rotations. This is just the two-valuedness which is needed for the description of the spin. Thus, we are working with a group in which each rotation corresponds to two elements which are represented by the two matrices given above.

If we have a spinor whose two components are scalar functions of coordinates, the complete operator corresponding to this coordinate transformation is

$$\pm U(R_s)R_s. \quad (10.3)$$

Here R_S is an operator which acts on the two scalar components of the spinor which are functions of the coordinates.

If we are given a set of g spatial operators R_S which form a group, then the 2g operators of the form (10.3) corresponding to them form the double group. The operators of the double group corresponding to the inversion are

$$\pm U(I_s)I_s, \quad (10.4)$$

where $U(I_S)$ is the two-rowed unit matrix, and I_S is the operator for space inversion.

We shall denote the two operators of the double group corresponding to the operator R_S, which acts on scalar coordinate functions, by R and $\bar{R}$. Both correspond to an ordinary rotation in Cartesian space and thus their action on a vector is the same as the action of the corresponding spatial operator R_S. Consequently, the "doubled" point group will leave the lattice invariant if the simple point group corresponding to it leaves the lattice invariant. Therefore the possible double point groups are the double groups corresponding to the 32 simple point groups.

A similar situation prevails for the double space groups. The operators for the rotational part of the space group are now the operators of the 32 double point groups, R and $\bar{R}$.

TABLE III

$\mathfrak{O}_h$	Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_1'	Γ_2'	Γ_3'	Γ_4'	Γ_5'
$\mathfrak{C}_{4v}$	γ_1	γ_3	$\gamma_1+\gamma_3$	$\gamma_4+\gamma_5$	$\gamma_2+\gamma_5$	γ_2	γ_4	$\gamma_2+\gamma_4$	$\gamma_3+\gamma_5$	$\gamma_1+\gamma_5$

TABLE IV

$\mathfrak{O}_h$	E	$\bar{E}$	$8C_3$	$8\bar{C}_3$	$3C_2$ $3\bar{C}_2$	$6C_2$ $6\bar{C}_2$	$6C_4$	$6\bar{C}_4$	I	$\bar{I}$	$8C_3I$	$8\bar{C}_3I$	$3C_2I$ $3\bar{C}_2I$	$6C_2I$ $6\bar{C}_2I$	$6C_4I$	$6\bar{C}_4I$
Γ_6	2	−2	1	−1	0	0	$\sqrt{2}$	$-\sqrt{2}$	2	−2	1	−1	0	0	$\sqrt{2}$	$-\sqrt{2}$
Γ_7	2	−2	1	−1	0	0	$-\sqrt{2}$	$\sqrt{2}$	2	−2	1	−1	0	0	$-\sqrt{2}$	$\sqrt{2}$
Γ_8	4	−4	−1	1	0	0	0	0	4	−4	−1	1	0	0	0	0
Γ_6'	2	−2	1	−1	0	0	$\sqrt{2}$	$-\sqrt{2}$	−2	2	−1	1	0	0	$-\sqrt{2}$	$\sqrt{2}$
Γ_7'	2	−2	1	−1	0	0	$-\sqrt{2}$	$\sqrt{2}$	−2	2	−1	1	0	0	$\sqrt{2}$	$-\sqrt{2}$
Γ_8'	4	−4	−1	1	0	0	0	0	−4	4	1	−1	0	0	0	0

The two-valued representations of the space group will obviously be single-valued representations of the corresponding double space group, so that to find them we can apply standard procedures. We must find irreducible representations of the group of the wave vector $\mathbf{k}$ assuming that the point group $\mathfrak{G}_0(\mathbf{k})$, corresponding to the group of the wave vector, is now a double point group. From this consideration it follows that the knowledge of the irreducible representations of the 32 double point groups enables us to find the irreducible representations of all the double space groups for points within the Brillouin zone.

For an arbitrary point in the Brillouin zone, the inclusion of spin-orbit interaction leads to a change in the energy within the band, but this effect will be small in general since the spin-orbit interaction energy is small compared to the width of the band. At points of high symmetry within the Brillouin zone the introduction of the spin may result in a certain splitting. In order to clarify how this comes about, let us consider a state whose eigenfunctions transform according to the representation $\Gamma_i(\mathbf{k})$ of the group of the wave vector $\mathbf{k}$.

When we include the spin, the wave function is the product of a coordinate function with a spin function which transforms according to the representation $D_{1/2}$ of the rotation group. The total wave function will then transform as the direct product $\Gamma_i(\mathbf{k}) \times D_{1/2}$. This direct product can then be expanded into irreducible representations of the double group

$$\Gamma_i(\mathbf{k}) \times D_{1/2} = \sum a_{ij}\Gamma_j(\mathbf{k}). \tag{10.5}$$

(Only two-valued representations of the double group appear in the sum on the right.) If in the expansion of the direct product more than one representation of the double group appears, this will indicate a splitting of the band because of spin-orbit interaction.

As already mentioned, the group of the wave vector $\mathbf{k} = (000)$ in a cubic lattice is the full cubic group $\mathfrak{O}_h$, whose characters were given in Table I. The characters of the two-valued representations of the corresponding double group are given in Table IV. It is easy to see that, for example,

$$\Gamma_5 \times D_{1/2} = \Gamma_6 + \Gamma_8, \tag{10.6}$$

so that the three bands which touch when $\mathbf{k} = (000)$ are separated as the result of the spin-orbit interaction into two which touch and one which is non-degenerate.

One can also show how the splitting of the bands occurs when we shift from one point in the Brillouin zone to another, i.e., one can construct compatibility tables.*

11. TIME REVERSAL[11,7]

In addition to the degeneracy caused by spatial symmetries, there may also occur a degeneracy caused by the invariance of the Hamiltonian with respect to a reversal of the sign of the time. We know that when the Hamiltonian of a problem is real, as it is in the simplest Schrödinger theory, the presence of the time reversal operation has the effect that along with each function ψ for a given energy level we will also have the function ψ^*. The operator for time reversal for particles without spin is the operator of complex conjugation, i.e.,

$$\hat{K}\psi = \psi^*, \tag{11.1}$$

whereas the corresponding operator for particles with spin $\frac{1}{2}$ is the product of the Pauli matrix $i\sigma_y$ and the operator which transforms the spinor to its complex conjugate.

When we include spin-orbit interaction, the Hamiltonian will no longer be real, but it can be shown that

$$\hat{H}^* = \hat{\sigma}_y^{-1}\hat{H}\hat{\sigma}_y. \tag{11.2}$$

It then follows immediately that the effect of the time reversal operation on the eigenfunction gives

$$\hat{K}\psi = \hat{\sigma}_y\psi^*. \tag{11.3}$$

Now, with each function ψ there belongs to the same eigenvalue of the Hamiltonian the function $\hat{\sigma}_y\psi^*$.

Thus, the presence of the time reversal operation in general leads to a doubling of the degeneracy.

In both cases these operators transform all the coordinates into themselves, and all the momenta are reversed in sign.

If we apply the time reversal operator to a system of eigenstates whose degeneracy is due to the spatial symmetry of the Hamiltonian, we obtain a new system of eigenstates which transform according to a representation which is complex conjugate to the original

*Complete tables for the main symmetry groups can be found, for example, in the paper of Elliott.[8]

one. Obviously two possibilities can occur: the new system of states is a linear combination of the initial ones, or they are linearly independent. In the first case, there is no additional degeneracy, while it does occur in the second case.

Let us consider the case of spinless particles. The time reversal operator $\hat{\mathbf{K}}$, like any symmetry element, commutes with the Hamiltonian. It also commutes with all the spatial symmetry elements, so that the full symmetry group regarded as an abstract group is the direct product of the space symmetry group and the time reversal operator. Let us investigate the case when there is some spatial symmetry, and consider eigenfunctions $\psi_1, \psi_2, \ldots, \psi_l$ which with respect to this symmetry belong to an irreducible representation D:

$$O_R\psi_\varkappa = \sum_{\lambda=1}^{l} D(R)_{\lambda\varkappa}\psi_\lambda. \tag{11.4}$$

The wave functions ψ_K ($\kappa = 1, \ldots, l$) satisfy the Schrödinger equation for stationary states

$$\hat{H}\psi_\varkappa = E\psi_\varkappa. \tag{11.5}$$

Multiplying (11.5) on the left by $\hat{\mathbf{K}}$:

$$\hat{K}\hat{H}\psi_\varkappa = \hat{H}(\hat{K}\psi_\varkappa) = E(\hat{K}\psi_\varkappa), \tag{11.6}$$

we see that ψ_K and $\hat{\mathbf{K}}\psi_K$ are in general degenerate eigenfunctions. We note that, since the time reversal operator is not linear, we cannot apply the theory of representations in its usual form. We must therefore state in detail the assumptions made and consider three cases:

1. The representation D is real.
2. The representations D and D* are not equivalent.
3. The representations D and D* are equivalent, but they are not equivalent to a real representation.

Let us show that in the first case the time reversal operation leads to no additional degeneracy and only has the effect that all eigenfunctions belonging to the real form of the representation can be made real. Multiplying (11.4) by $\hat{\mathbf{K}}$, we have

$$\hat{K}O_R\psi_\varkappa = O_R\hat{K}\psi_\varkappa = \sum_{\lambda=1}^{l} D(R)_{\lambda\varkappa}\hat{K}\psi_\lambda. \tag{11.7}$$

From (11.4) and (11.7) it follows that ψ_K and $\hat{\mathbf{K}}\psi_K$ transform according to the same representation D. In place of the functions ψ_K and $\hat{\mathbf{K}}\psi_K$ we can consider their linear combinations

$$u_\varkappa = \psi_\varkappa + \hat{K}\psi_\varkappa \quad \text{and} \quad v_\varkappa = i(\psi_\varkappa - \hat{K}\psi_\varkappa) \quad (\varkappa = 1, 2, \ldots, l),$$

which also satisfy equation (11.7) and are real.

Both the functions u_K and the functions v_K transform into themselves under space transformations as well as time reversal, i.e., they form two separate sets of functions. This also means that the ψ_K and $\hat{\mathbf{K}}\psi_K$ form two separate sets.

According to (11.6), $\hat{\mathbf{K}}\psi_K$ and ψ_K are degenerate functions. However, since $\hat{\mathbf{K}}\psi_K$ and ψ_K transform according to the same representation D, the degeneracy of $\hat{\mathbf{K}}\psi_K$ and ψ_K is already included in the degeneracy caused by the spatial symmetry group. This is possible only in case 1 since $\hat{\mathbf{K}}\psi_K$ is linearly dependent on ψ_K, i.e., these sets coincide to within an equivalence transformation.

Let us proceed to consider case 2. If the eigenfunctions are all real, the corresponding representations are also real. This follows from (11.4):

$$(\psi_\mu, O_R\psi_\varkappa) = \int \psi_\mu \sum_{\lambda=1}^{l} D(R)_{\lambda\varkappa}\psi_\lambda\, d\tau = \sum_{\lambda=1}^{l} D(R)_{\lambda\varkappa} \int \psi_\mu\psi_\lambda\, d\tau$$

$$= \sum_{\lambda=1}^{l} D(R)_{\lambda\varkappa}\delta_{\mu\lambda} = D(R)_{\mu\varkappa};$$

because on the left both ψ_μ and $O_R\psi_K$ are real. If the representations cannot be made real, the corresponding functions also cannot be real.

Let us show that in this case the time reversal operation gives rise to an additional degeneracy not required by the spatial symmetry.

Even though the set of functions $\hat{\mathbf{K}}\psi_1, \ldots, \hat{\mathbf{K}}\psi_l$ belongs to the same eigenvalue as the set $\psi_1, \ldots, \psi_l$, the functions of the set $(\hat{\mathbf{K}}\psi)$ can no longer be expressed linearly in terms of the functions $\psi_1, \ldots, \psi_l$, as was the case for case 1. In fact, from (11.4) considering the non-linearity of the operator $\hat{\mathbf{K}}$, we obtain

$$\hat{K}O_R\psi_\varkappa = O_R\hat{K}\psi_\varkappa = \sum_{\lambda=1}^{l} D^*(R)_{\lambda\varkappa}\hat{K}\psi_\lambda. \tag{11.8}$$

On the basis of (11.4) and (11.8) we see that ψ_K belongs to D, while $\hat{\mathbf{K}}\psi_K$ belongs to D*. At the same time both sets (or both representations D and D*) belong to the same eigenvalue of the energy. Therefore the presence of the symmetry operation $\hat{\mathbf{K}}$ doubles the degeneracy.

Finally, in case 3, the new symmetry element — time reversal — also gives rise to additional degeneracy.

In fact, on the basis of the argument given above, since D and D* are equivalent but different, we can conclude that there is a coincidence of two energy eigenvalues with equivalent representations, i.e., there is an additional degeneracy not required by the spatial symmetry. In this case, the representation D is always double.

When we include the spin, the situation is different: The roles of case 1 and case 3 are interchanged. Degeneracy occurs in case 1, the number of states is doubled. The number of states is also doubled in case 2, but not in case 3 as in the preceding treatment.

The fact that the situation is different for spinless particles and particles with spin is not unexpected. For the case of a spinless particle, the square of the time reversal operator is the identity operator. We can always choose the eigenfunctions so that they are transformed into themselves under the action of the time reversal operation, so that they can always be chosen to be real. This cannot be done for particles

with spin ½. The square of the time reversal operator multiplies every spinor by −1. We can therefore not form eigenfunctions of the time reversal operator. Consequently, the spinor and the time reversed spinor are always linearly independent for particles with spin ½. Consequently, if the Hamiltonian is invariant under time reversal, the eigenfunctions of the problem are always at least doubly degenerate.

There is a simple criterion which enables us to determine to which of the three types a representation belongs. For the cases 1, 2 or 3 we have respectively

$$\sum_{Q}^{g} \chi(Q^2) = \begin{matrix} g \\ 0. \\ -g \end{matrix} \qquad (11.9)$$

Here Q is the matrix of the corresponding element of the group and g is the order of the group.

Thus for any group we can use the character table to decide whether there is additional degeneracy caused by invariance of the Hamiltonian with respect to time reversal.

Herring[7] studied in detail the application of this criterion to the case of space groups. He showed that relation (11.9) reduces to the following:

$$\sum_{Q_0} \chi(Q_0^2) = g.\ 0, -g, \qquad (11.10)$$

where Q_0 is the element of the space group which takes **k** into **−k**.

If the group of the wave vector does not contain the inversion I, then $Q_0 = I \times \mathfrak{K}$. If, however, I is contained in $\mathfrak{K}$, then Q_0 consists simply of the elements of the group $\mathfrak{K}$. The operation Q_0^2 is contained in the group $\mathfrak{K}$, so that the characters appearing above can be taken for the irreducible representation of the group of the wave vector **k** containing g elements. At an arbitrary point of the Brillouin zone, the factor group contains the group of translations and the identity, which now has corresponding to it two elements ϵ and $\bar{\epsilon}$. The only elements Q_0 which transform **k** into **− k** are the inversions $\{i|t\}$ and $\{\bar{1}|t\}$. Therefore, if the space group contains the inversion,

$$\sum_{Q_0} \chi(Q_0^2) = 2\chi(\varepsilon) = 2. \qquad (11.11)$$

This means that we have the first case and, at an arbitrary point of the Brillouin zone, there is always a twofold degeneracy because of time reversal. Thus, if the crystal has a center of symmetry, there must be twofold degeneracy at any point of the zone. Consequently, any degenerate representations which are found at points with higher symmetry must belong to cases 1 and 2.

In conclusion, we mention that the problem of time reversal has been treated most rigorously in the paper of Johnston[12] who investigated it on the basis of the relativistic Dirac equation.

12. CONCLUSION

The general principles for the investigation of quantum mechanical system by using group theory which have been presented in this paper show that the basic questions of the quantum physics of solids can be treated most profoundly and completely only by using the apparatus of the theory of space groups.

The symmetry properties of physical systems should be used more widely in the solution of specific problems and, especially, in the case of complex quantum mechanical systems, where exact quantitative computations cannot be done, so that it is important to obtain the greatest possible number of results by both methods. In addition, conclusions which are obtained by using group theory are the most rigorous ones because of the phenomenological character of the theory of symmetry.

The general statements presented in this survey can be applied to any crystal lattice. These general statements were made more detailed for the case of the band theory of solids, as treated from the point of view of group theory, and illustrated in particular by an investigation of the energy spectrum and classification of states for the linear chain and the properties of an electron in a field of cubic symmetry. In principle, they can be applied also to lattices with more complicated symmetry.

The theory of space groups is beginning to occupy a prominent place in investigations of magnetic symmetry. Here we should mention the extreme importance of investigation of the theory of magnetic space groups and its application to ferro- and antiferromagnetic materials.[13] The source of this latest work already appeared in the investigations by Landau and Lifshitz.[14]

[1] A. V. Sokolov and V. P. Shirokovskiĭ, Usp. Fiz. Nauk **60**, 617 (1956).

[2] H. A. Bethe, Ann. Physik **3**, 133 (1929).

[3] F. Seitz, Z. Krist. **88**, 433 (1934); **90**, 289 (1935); **91**, 336 (1935); **94**, 100 (1936); Ann. Math. **37**, 17 (1936).

[4] Bouckaert, Smoluchowski, and Wigner, Phys. Rev. **50**, 58 (1936).

[5] C. Herring, J. Franklin Inst. **233**, 525 (1942).

[6] W. Döring and V. Zehler, Ann. Physik 13, 214 (1953).

[7] C. Herring, Phys. Rev. **52**, 361 (1937).

[8] R. J. Elliott, Phys. Rev. **96**, 280 (1954).

[9] G. F. Koster, Solid State Phys. **5**, 173 (1957).

[10] V. P. Shirokovskiĭ, Физика металлов и металловедение (Phys. of Metals and Metallurgy) **6**, 3 (1958).

[11] E. P. Wigner, Göttinger Nachr., 546 (1932).

[12] D. F. Johnston, Proc. Roy. Soc. (London) **A243**, 546 (1958).

[13] I. E. Dzyaloshinskiĭ, JETP **32**, 1547 (1957); Soviet Phys. JETP **5**, 1259 (1957).

[14] L. D. Landau and E. M. Lifshitz, Статистическая физика (Statistical Physics), GITTL, M-L, 1951 [Transl. Pergamon Press, London, 1958]; E. M. Lifshitz, JETP **11**, 255 (1941).

Translated by M. Hamermesh
39