Chapter IV

EFFECTIVE MASS THEORY

§21. PERTURBATION THEORY. $\boldsymbol{kp}$-THEORY

To determine in detail the spectrum $E_n(\boldsymbol{k})$ throughout the Brillouin zone, one needs a numerical solution of the Schrödinger equation (17.1). Several approximate methods are now available to determine the wave functions and energy spectrum of electrons in the solid state, but a numerical solution of the Schrödinger equation is fraught with practical difficulties, involved mainly in determination of the self-consistent potential $V(\boldsymbol{x})$.

However, in most semiconductors a knowledge of the entire energy spectrum is not necessary, and a different approach has proved quite fruitful. In semiconductors, the number of current carriers — electrons in the conduction band and holes in the valence band — is usually small compared to the total number of atoms in the crystal; thus, if the allowed band is sufficiently wide, in thermodynamic equilibrium the electrons are concentrated near a minimum of the conduction band, and the holes are concentrated near a valence band maximum. Therefore, in most problems in the physics of semiconductors a knowledge of the energy spectrum near the band extrema is sufficient.

It is significant that the possible shapes of the spectrum near a given point can be determined by symmetry considerations alone, without numerical solution of the Schrödinger equation.

Since the shape of the spectrum near an extremum is the essential factor determining the experimentally observed properties of semiconductors, comparison of experimental data with theory indicates what shape the spectrum may have for a given semiconductor and enables one to find the constants appearing in the energy.

A necessary condition for a band extremum is that all the derivatives of the energy with respect to k_α vanish:

$$\frac{\partial E_n(\boldsymbol{k}_0)}{\partial k_\alpha} = 0, \tag{21.1}$$

i.e., the band has zero slope at the point $\boldsymbol{k}_0$.

Although in principle there may be a band extremum at any point of $\boldsymbol{k}$-space, in almost all semiconductors for which extremum positions are known they lie at points of the Brillouin zone where the zero slope of the band for all or some directions k_α is due to the symmetry of the point $\boldsymbol{k}_0$. Therefore, it is important to determine which points of the Brillouin zone may be points of zero slope and are consequently likely candidates for positions of the band extrema.

This problem can be solved by group-theoretic methods. Obviously, all the derivatives $\partial E_n/\partial k_\alpha$ may vanish simultaneously only at isolated points of $\boldsymbol{k}$-space. The "candidates" are points in $\boldsymbol{k}$-space at which all the velocity components vanish because of symmetry or time-reversal conditions, points on lines of symmetry at which two velocity components vanish identically, and also points of a plane on which one of the velocity components vanishes owing to symmetry or time reversal.

An extremum may of course be located at an arbitrary point of the Brillouin zone, possessing no symmetry, where all three velocity components vanish for some specific potential $V(\boldsymbol{x})$.

Qualitative investigation of the spectrum near a point $\boldsymbol{k}_0$ is based on "$\boldsymbol{kp}$-theory," which is a variant of perturbation theory essentially similar to the method used in §17 to investigate the analytical properties of wave functions and energies. In §17 we used the Bloch functions $\psi_{n\boldsymbol{k}}$, which are exact solutions of the Schrödinger equation (17.1), (17.2) for an ideal crystal. In semiconductors, where the significant regions in $\boldsymbol{k}$-space are those close to an extremum point $\boldsymbol{k}_0$, it is convenient to use instead the Luttinger-Kohn representation to investigate the energy spectrum of current carriers in an ideal crystal, especially in regard to the behavior of current carriers in slowly varying electric and magnetic fields. In this representation the wave vector $\boldsymbol{k}$ is measured from the extremum point $\boldsymbol{k}_0$, i.e., $\boldsymbol{k}=\boldsymbol{K}-\boldsymbol{k}_0$, where $\boldsymbol{K}$ is the wave vector of the point in $\boldsymbol{k}$-space (in §17 the difference $\boldsymbol{K}-\boldsymbol{k}_0$ was denoted by $\boldsymbol{\varkappa}$). Rather than the Bloch functions $\psi_{n\boldsymbol{K}}$ (17.5) the following functions $\varphi_{n\boldsymbol{k}}$ are taken as basis:

$$\varphi_{n\boldsymbol{k}}=\frac{1}{\sqrt{\mathscr{V}}}\psi_{n\boldsymbol{k}_0}e^{i\boldsymbol{k}\boldsymbol{x}}, \tag{21.2}$$

where $\psi_{n\boldsymbol{k}_0}$ is the Bloch function at the extremum point $\boldsymbol{k}_0$,

$$\psi_{n\boldsymbol{k}_0}=e^{i\boldsymbol{k}_0\boldsymbol{x}}u_{n\boldsymbol{k}_0},$$

which is an eigenfunction of the operator $\mathscr{H}_0=(p^2/2m)+V(\boldsymbol{x})$ with energy $E_n(\boldsymbol{k}_0)$.

The function $\varphi_{n\boldsymbol{k}}$ is normalized with respect to the crystal volume $\mathscr{V}=N\Omega_0$, where N is the number of primitive cells in the crystal and Ω_0 the volume of a primitive cell. The functions $u_{n\boldsymbol{k}_0}$ are normalized with respect to the volume of the primitive cell:

$$\frac{1}{\Omega_0}\int\limits_{\Omega_0}u^*_{n\boldsymbol{k}_0}u_{n\boldsymbol{k}_0}\,d\tau_0=1. \tag{21.3}$$

The function $\varphi_{n\boldsymbol{k}}$ transforms according to the same representation of the translation group (i.e., corresponds to the same $\boldsymbol{k}$) as the exact function $\psi_{n\boldsymbol{K}}$, but differs from it in that the function corresponding to $\varphi_{n\boldsymbol{k}}$ is the periodic function $u_{n\boldsymbol{k}_0}$ rather than $u_{n\boldsymbol{K}}$.

The functions $\varphi_{n\boldsymbol{k}}$ are orthogonal:

$$\langle\varphi_{n'\boldsymbol{k}'},\varphi_{n\boldsymbol{k}}\rangle=\delta_{nn'}\delta_{\boldsymbol{k}\boldsymbol{k}'}. \tag{21.4}$$

Here, for convenience, we assume that the crystal volume $\mathscr{V}$ is finite, and thus cyclic boundary conditions are imposed on the functions $\varphi_{n\boldsymbol{k}}$. Hence $\boldsymbol{k}$ may take only discrete values.

The orthogonality of the functions $\varphi_{n\boldsymbol{k}}$ with respect to $\boldsymbol{k}$, like that of the functions $\psi_{n\boldsymbol{K}}$, is due to the fact that under translation by a Bravais lattice vector the matrix element $\langle\varphi_{n'\boldsymbol{k}'}\varphi_{n\boldsymbol{k}}\rangle$ receives a factor $e^{i\boldsymbol{a}(\boldsymbol{k}-\boldsymbol{k}')}$, which is not equal to unity since the vectors $\boldsymbol{k}$ and $\boldsymbol{k}'$ are in the Brillouin zone. The orthogonality with respect to n is a consequence of the orthogonality of the periodic functions $u_{n\boldsymbol{k}_0}$.

The wave function ψ_{nK} can be expanded in the functions $\varphi_{n\boldsymbol{k}}$, which have the same period:

$$\psi_{nK}=\sum_{n'} c_{n'}\varphi_{n'\boldsymbol{k}}. \tag{21.5}$$

The coefficients $c_{n'}$ define the transformation from the Bloch representation to the Luttinger-Kohn representation.

We first calculate the spectrum $E(\boldsymbol{k})$ without spin-orbit coupling.

Substituting the expansion (21.5) into the Schrödinger equation $(\mathscr{H}_0-E)\psi=0$, multiplying on the left by $\varphi^*_{n\boldsymbol{k}}$ and integrating over $\boldsymbol{x}$, we obtain the system of equations

$$\sum_{n'}\left\{\left(E_n(\boldsymbol{k}_0)+\frac{\hbar^2k^2}{2m}-E\right)\delta_{nn'}+\frac{\hbar\boldsymbol{k}\boldsymbol{p}_{nn'}}{m}\right\}c_{n'}=0. \tag{21.6}$$

We have used the equation

$$\mathscr{H}_0(\boldsymbol{p})\,e^{i\boldsymbol{k}\boldsymbol{x}}=e^{i\boldsymbol{k}\boldsymbol{x}}\mathscr{H}_0(\boldsymbol{p}+\hbar\boldsymbol{k}). \tag{21.7}$$

In (21.6) $p_{nn'}$ are the momentum matrix elements between the Bloch functions $\psi_{n\boldsymbol{k}_0}$:

$$\boldsymbol{p}_{nn'}=\boldsymbol{p}_{n\boldsymbol{k}_0,\,n'\boldsymbol{k}_0}=\frac{1}{\mathscr{V}}\int\psi^*_{n\boldsymbol{k}_0}\boldsymbol{p}\psi_{n'\boldsymbol{k}_0}\,d\tau_0. \tag{21.8}$$

For small values of $\boldsymbol{k}$, the terms

$$\mathscr{H}_1=\frac{\hbar^2k^2}{2m},\qquad \mathscr{H}_2=\frac{\hbar\boldsymbol{k}\boldsymbol{p}}{m} \tag{21.9}$$

in equation (21.6) can be considered as a perturbation and, as in (15.35) and (15.45), only the operator $\mathscr{H}_2$ has interband matrix elements. To eliminate them, we apply the transformation (15.33); according to equation (15.47) this gives, to second order in $\boldsymbol{k}$,

$$\begin{aligned}\sum_{n'}\Bigg\{\left(E_n(\boldsymbol{k}_0)+\frac{\hbar^2k^2}{2m}-E\right)\delta_{nn'}+\frac{\hbar\boldsymbol{k}}{m}\,\boldsymbol{p}_{n\boldsymbol{k}_0,\,n'\boldsymbol{k}_0}+\\+\frac{\hbar^2}{m^2}\sum_{\alpha\beta}\sum_{n''\neq n}k_\alpha k_\beta\,\frac{p^\alpha_{n\boldsymbol{k}_0,\,n''\boldsymbol{k}_0}\,p^\beta_{n''\boldsymbol{k}_0,\,n'\boldsymbol{k}_0}}{E_n(\boldsymbol{k}_0)-E_{n''}(\boldsymbol{k}_0)}\Bigg\}c_{n'}=0.\end{aligned} \tag{21.10}$$

Nondegenerate bands. For a nondegenerate band, equation (21.10) implies that the energy $E(\boldsymbol{k}_0+\boldsymbol{k})$ is, up to $\boldsymbol{k}^2$ terms,

$$E_n(\boldsymbol{k}_0+\boldsymbol{k})=E_n(\boldsymbol{k}_0)+\frac{\hbar}{m}\sum_\alpha k_\alpha\,p^\alpha_{n\boldsymbol{k}_0,\,n\boldsymbol{k}_0}+\sum_{\alpha\beta}\frac{\hbar^2}{2m_{\alpha\beta}}\,k_\alpha k_\beta, \tag{21.11}$$

where, by (21.10),

$$\frac{1}{m_{\alpha\beta}}=\frac{1}{m^2}\sum_{n'\neq n}\frac{p^\alpha_{n\boldsymbol{k}_0,\,n'\boldsymbol{k}_0}\,p^\beta_{n'\boldsymbol{k}_0,\,n\boldsymbol{k}_0}+p^\beta_{n\boldsymbol{k}_0,\,n'\boldsymbol{k}_0}\,p^\alpha_{n'\boldsymbol{k}_0,\,n\boldsymbol{k}_0}}{E_n(\boldsymbol{k}_0)-E_{n'}(\boldsymbol{k}_0)}+\frac{1}{m}\delta_{\alpha\beta}. \tag{21.12}$$

It is significant that, starting only from symmetry considerations and using selection rules for the momentum matrix elements, we can determine the number of nonzero constants p^α and $1/m_{\alpha\beta}$ and establish the relations between them, thus obtaining a qualitative expression (21.11) for the spectrum $E_n(\boldsymbol{k}_0+\boldsymbol{k})$. Since the momentum operator transforms according to the representation $\mathscr{D}_1$ (which in the group $G_{\boldsymbol{k}_0}$ may be reducible), in the general case the number of nonzero $p^{\alpha}_{n\boldsymbol{k}_0,\, n\boldsymbol{k}_0}$ equals the number N_1 of identity representations occurring in the direct product

$$\mathscr{D}^{\boldsymbol{k}_0}_{\mu}\times\mathscr{D}_1\times\mathscr{D}^{*\boldsymbol{k}_0}_{\mu}=\mathscr{D}_1,$$

and is given by equation (19.43)

$$N_1=\frac{1}{h'}\sum_{g\in G'_{\boldsymbol{k}_0}}|\chi_{\mu}(g)|^2\chi_1(g)=\frac{1}{h'}\sum_{g\in G'_{\boldsymbol{k}_0}}\chi_1(g), \tag{21.13}$$

since for one-dimensional representations the product $\mathscr{D}^{*\boldsymbol{k}_0}_{\mu}\times\mathscr{D}^{\boldsymbol{k}_0}_{\mu}$ is always the identity representation.

It follows from (21.13) that the only nonzero constants $p^{\alpha}_{n\boldsymbol{k}_0,\, n\boldsymbol{k}_0}$ are those corresponding to components k_α which transform according to the identity representation of the group $G_{\boldsymbol{k}_0}$, i.e., are invariant under all transformations in $G_{\boldsymbol{k}_0}$.

Since $\partial E_n/\partial k_\alpha = p^{\alpha}_{n\boldsymbol{k}_0,\, n\boldsymbol{k}_0}$, equation (21.13) determines the points of zero slope, i.e., points $\boldsymbol{k}_0$ at which certain components of the velocity $\partial E/\partial k_\alpha$ vanish identically because of the symmetry conditions.

Thus, if the symmetry group of the point $\boldsymbol{k}_0$ contains inversion, all the momentum components vanish: $\partial E_n/\partial k_\alpha = 0$ for all directions of the vector $\boldsymbol{k}$. But if $\boldsymbol{k}_0$ lies on a twofold or fourfold symmetry axis, the only component of $\boldsymbol{k}_0$ invariant under all rotations will be that parallel to this axis, so that in this case the two velocity components normal to the symmetry axis vanish owing to symmetry. In case the point lies on a plane of symmetry, the component normal to the plane of symmetry vanishes, since it changes sign upon reflection in the plane.

For one-dimensional representations, it follows from (21.13) that these conditions are independent of the representation $\mathscr{D}^{\boldsymbol{k}_0}_{\mu}$ itself, depending only on the symmetry of the point $\boldsymbol{k}_0$.

Let us now see how the sum (21.12) transforms under the action of an element $g\in G_{\boldsymbol{k}_0}$.

If one of the bands n' over which the summation in (21.12) extends is m-fold degenerate, the sum (21.12) contains m terms, according to the number of degenerate functions $\psi_{s(n')\boldsymbol{k}_0}$ $(s=1, 2, \ldots, m)$ in the band n', with the same energy denominator $E_n(\boldsymbol{k}_0)-E_{n'}(\boldsymbol{k}_0)$. Under the action of an element $g\in G_{\boldsymbol{k}_0}$ the sum of these terms becomes

$$\sum_s\langle g\psi_{n\boldsymbol{k}_0}|gp^{\alpha}|g\psi_{s(n')\boldsymbol{k}_0}\rangle\langle g\psi_{s(n')\boldsymbol{k}_0}|gp^{\beta}|g\psi_{n\boldsymbol{k}_0}\rangle=$$

$$=\sum_{ss't\alpha'\beta'}\mathscr{D}^{*\boldsymbol{k}_0}_{\mu}(g)\,\mathscr{D}^{\alpha'\alpha}_1(g)\,\mathscr{D}^{\boldsymbol{k}_0}_{\mu's's}(g)\,\mathscr{D}^{*\boldsymbol{k}_0}_{\mu'ts}(g)\,\mathscr{D}^{\beta'\beta}_1(g)\,\mathscr{D}^{\boldsymbol{k}_0}_{\mu}(g)\times$$

$$\times\langle\psi_{n\boldsymbol{k}_0}|p^{\alpha'}|\psi_{s'(n')\boldsymbol{k}_0}\rangle\langle\psi_{t(n')\boldsymbol{k}_0}|p^{\beta'}|\psi_{n\boldsymbol{k}_0}\rangle,$$

where $\mathscr{D}_1(g)$ is the representation according to which the momentum $\boldsymbol{p}$ transforms and $\mathscr{D}^{\boldsymbol{k}_0}_{\mu'}$ the representation according to which the functions $\psi_{s(n')\boldsymbol{k}_0}$ transform. Using the unitarity relation (7.11)

$$\sum_s\mathscr{D}^{\boldsymbol{k}_0}_{\mu's's}(g)\,\mathscr{D}^{*\boldsymbol{k}_0}_{\mu'ts}(g)=\delta_{s't},$$

we obtain

$$\sum_{s}\langle g\psi_{n\boldsymbol{k}_0}|gp^{\alpha}|g\psi_{s(n')\boldsymbol{k}_0}\rangle\langle g\psi_{s(n')\boldsymbol{k}_0}|gp^{\beta}|g\psi_{n\boldsymbol{k}_0}\rangle=$$
$$=\sum_{s\alpha'\beta'}\langle\psi_{n\boldsymbol{k}_0}|p^{\alpha'}|\psi_{s(n')\boldsymbol{k}_0}\rangle\langle\psi_{s(n')\boldsymbol{k}_0}|p^{\beta'}|\psi_{n\boldsymbol{k}_0}\rangle\times$$
$$\times\mathscr{D}_1^{\alpha'\alpha}(g)\,\mathscr{D}_1^{\beta'\beta}(g)\,\mathscr{D}_{\mu}^{*\boldsymbol{k}_0}(g)\,\mathscr{D}_{\mu}^{\boldsymbol{k}_0}(g),$$

i.e., each term in (21.12) transforms like $\langle\psi_{n\boldsymbol{k}_0}|p^{\alpha}p^{\beta}+p^{\beta}p^{\alpha}|\psi_{n\boldsymbol{k}_0}\rangle$, and therefore the entire sum (21.12) (i.e., the constants $1/m_{\alpha\beta}$) transform according to the direct product

$$\mathscr{D}_{\mu}^{\boldsymbol{k}_0^*}[\mathscr{D}_1\times\mathscr{D}_1]\,\mathscr{D}_{\mu}^{\boldsymbol{k}_0}=[\mathscr{D}_1\times\mathscr{D}_1]. \tag{21.14}$$

The number N_2 of independent constants $1/m_{\alpha\beta}$ is, by (21.14) and (19.43),

$$N_2=\frac{1}{2h'}\sum_{g\in G'_{\boldsymbol{k}_0}}|\chi_{\mu}(g)|^2(\chi_1^2(g)+\chi_1(g^2))=\frac{1}{h'}\sum_{g\in G'_{\boldsymbol{k}_0}}[\chi_1^2(g)]. \tag{21.15}$$

Equations (21.14) and (21.15) imply that the quadratic part of the energy expansion (21.11) contains only combinations $k_{\alpha}k_{\beta}$ invariant under transformations from the group $G_{\boldsymbol{k}_0}$. Near an extremum point the energy is a quadratic function of k_{α}, and moreover the form $\frac{1}{m_{\alpha\beta}}k_{\alpha}k_{\beta}$ must be positive definite for an extremum to exist.

Near an extremum, the energy $E_n(\boldsymbol{k}_0+\boldsymbol{k})$ can be written as

$$E_n(\boldsymbol{k}_0+\boldsymbol{k})=E_n(\boldsymbol{k}_0)+\sum_{\alpha\beta}\frac{\hbar^2}{2m_{\alpha\beta}}k_{\alpha}k_{\beta}. \tag{21.16}$$

The constants $1/m_{\alpha\beta}$ are known as the inverse effective mass tensor.

Thus, near an extremum point, the energy is a parabolic function of k_{α} to lowest order in k_{α}, but in contrast to a free electron it is characterized by anisotropic effective mass.

The symmetric tensor $1/m_{\alpha\beta}$ may be reduced to principal axes, and then (21.16) becomes

$$E_n(\boldsymbol{k}_0+\boldsymbol{k})=E_n(\boldsymbol{k}_0)+\sum_{\alpha}\frac{\hbar^2}{2m_{\alpha}}k_{\alpha}^2. \tag{21.17}$$

Therefore, for a nondegenerate band the surfaces of constant energy near an extremum are always ellipsoidal, determined by six parameters: the six components of the symmetric tensor $1/m_{\alpha\beta}$, or the three principal values of $1/m_{\alpha}$ and the three directions of the principal axes of the energy ellipsoid.

The symmetry of the ellipsoid (21.16) depends only on the symmetry of the point $\boldsymbol{k}_0$. Thus, if the extremum point is on a symmetry axis, one of the axes of the energy ellipsoid coincides with this symmetry axis. If the symmetry axis is threefold, fourfold or sixfold, the energy ellipsoid is an ellipsoid of revolution and the other two axes of the ellipsoid can be chosen arbitrarily in a plane perpendicular to the symmetry axis. If $\boldsymbol{k}_0$ lies at the intersection of three mutually perpendicular twofold axes (has symmetry group D_2) or on a twofold axis in a reflection plane σ_v (symmetry group C_{2v}), the ellipsoid is characterized by three different effective masses. The principal axes of the ellipsoid are then directed along the symmetry elements,

TABLE 21.1

Symmetry group F_{k_0}	Vanishing components p^{α}	Energy near extremum point	Position of principal axes
e	—	$\frac{\hbar^2}{2m_1}k_1^2+\frac{\hbar^2}{2m_2}k_2^2+\frac{\hbar^2}{2m_3}k_3^2$	Arbitrary
C_i	$p^x,\ p^y,\ p^z$	$\frac{\hbar^2}{2m_1}k_1^2+\frac{\hbar^2}{2m_2}k_2^2+\frac{\hbar^2}{2m_3}k_3^2$	Arbitrary
C_s	p^z	$\frac{\hbar^2k_1^2}{2m_1}+\frac{\hbar^2k_2^2}{2m_2}+\frac{\hbar^2k_3^2}{2m_3}$	x- and y-axis in plane σ_h, z-axis normal to plane σ_h
C_2	$p^x,\ p^y$	$\frac{\hbar^2k_1^2}{2m_1}+\frac{\hbar^2k_2^2}{2m_2}+\frac{\hbar^2k_3^2}{2m_3}$	x- and y-axis in plane perpendicular to axis c_2; z-axis along axis c_2
C_{2h}	$p^x,\ p^y,\ p^z$	$\frac{\hbar^2k_1^2}{2m_1}+\frac{\hbar^2k_2^2}{2m_2}+\frac{\hbar^2k_3^2}{2m_3}$	z-axis along axis c_2, x- and y-axis in plane σ_h
C_{2v}	$p^x,\ p^y$	$\frac{\hbar^2k_1^2}{2m_1}+\frac{\hbar^2k_2^2}{2m_2}+\frac{\hbar^2k_3^2}{2m_3}$	z-axis along axis c_2, x- and y-axis along intersection of plane σ_v with planes perpendicular to c_2
D_2, D_{2h}	$p^x,\ p^y,\ p^z$	$\frac{\hbar^2k_1^2}{2m_1}+\frac{\hbar^2k_2^2}{2m_2}+\frac{\hbar^2k_3^2}{2m_3}$	x-, y- and z-axis along three twofold axes
S_4, D_{2d}	$p^x,\ p^y,\ p^z$	$\frac{\hbar^2\left(k_1^2+k_2^2\right)}{2m_1}+\frac{\hbar^2k_3^2}{2m_3}$	z-axis along axis s_4, x- and y-axis arbitrary in plane perpendicular to s_4
C_4, C_{4v}	$p^x,\ p^y$	$\frac{\hbar^2}{2m_1}\left(k_1^2+k_2^2\right)+\frac{\hbar^2k_3^2}{2m_3}$	z-axis along axis c_4, x- and y-axis arbitrary in plane perpendicular to c_4
D_{h}, C_{4h}, D_4	$p^x,\ p^y,\ p^z$	$\frac{\hbar^2}{2m_1}\left(k_1^2+k_2^2\right)+\frac{\hbar^2k_3^2}{2m_3}$	z-axis along axis c_4, x- and y-axis arbitrary in plane perpendicular to c_4
C_6, C_{6v}, C_3, C_{3v}	$p^x,\ p^y$	$\frac{\hbar^2}{2m_1}\left(k_1^2+k_2^2\right)+\frac{\hbar^2k_3^2}{2m_3}$	z-axis along axis c_3, x- and y-axis arbitrary in plane perpendicular to axis c_3
C_{3h}, D_{3h}, C_{6h}, D_3, S_6, D_{3d}, D_6, D_{6h}	$p^x,\ p^y,\ p^z$	$\frac{\hbar^2}{2m_1}\left(k_1^2+k_2^2\right)+\frac{\hbar^2k_3^2}{2m_3}$	As above
O, O_h, T, T_h, T_d	$p^x,p^y,\ p^z$	$\frac{\hbar^2}{2m_1}\left(k_1^2+k_2^2+k_3^2\right)$	Arbitrary

i.e., along the mutually perpendicular twofold axes or, in the case of C_{2v}, along the twofold axis and lines in which the reflection planes σ_v intersect the plane perpendicular to the twofold axis.

Table 21.1 lists the momentum components which vanish owing to the symmetry conditions (21.13) for the one-dimensional representations of all the point groups of $\boldsymbol{k}_0$, and describes the position of the axes of energy ellipsoids.

If the extremum lies at the center of the Brillouin zone, $\boldsymbol{k}_0 = 0$, the group $F_{\boldsymbol{k}_0}$ is the full point group F characterizing the crystal class, and there is one ellipsoid whose symmetry coincides with the macrosymmetry of the crystal.

A similar situation may occur when the extremum of a nondegenerate band is on the boundary of the Brillouin zone, if the symmetry group of $\boldsymbol{k}_0$ coincides with the space group.

If the band extremum is not at the center of the Brillouin zone, the symmetry of the point $\boldsymbol{k}_0$ is in general lower than that defined by the crystal class, and the symmetry of the energy $E(\boldsymbol{k})$ near the extremum may also be lower than that of the crystal. In this case, however, the star of $\boldsymbol{k}_0$ always contains other points $\boldsymbol{k}_0^i$, the energy at each point of the star is the same, $E(\boldsymbol{k}_0^i) = E(\boldsymbol{k}_0)$, and the form of $E(\boldsymbol{k}_0^i + \boldsymbol{k})$ near $\boldsymbol{k}_0^i$ is obtained from the ellipsoid (21.16) by applying those elements g_i of the group G which take the vector $\boldsymbol{k}_0$ into $\boldsymbol{k}_0^i$: $\boldsymbol{k}_0^i = g_i \boldsymbol{k}_0$. Thus the energy ellipsoid at $\boldsymbol{k}_0^i$ is obtained from the energy ellipsoid at $\boldsymbol{k}_0$ by the rotation of axes corresponding to the element g_i

By (17.10), the energy spectrum as a whole possesses the symmetry of the crystal.

This type of band is known as multi-ellipsoid or many-valley band structure. Here one has energy degeneracy at different points of $\boldsymbol{k}$-space, although the band at each extremum $\boldsymbol{k}_0^i$ is not degenerate.

Bands degenerate at $\boldsymbol{k}_0$. According to degenerate perturbation theory for equation (21.10), we must consider all states n' which have the same energy at the point $\boldsymbol{k}_0$. We shall denote these states by indices $i, j = 1, 2, \ldots, m$, where m is the degeneracy of the band at $\boldsymbol{k}_0$. The system of equations (21.10) for a degenerate band at $\boldsymbol{k}_0$ has the form

$$\sum_j \mathcal{H}_{ij}(\boldsymbol{k}_0, \boldsymbol{k})\, c_j = E c_i, \tag{21.18}$$

where, by (21.10),

$$\mathcal{H}_{ij}(\boldsymbol{k}_0, \boldsymbol{k}) = \sum_\alpha \frac{\hbar}{m} k_\alpha p^\alpha_{i\boldsymbol{k}_0,\, j\boldsymbol{k}_0} + \frac{\hbar^2}{2m}\delta_{ij} k^2 + \\ + \frac{\hbar^2}{2m^2} \sum_{\alpha\beta} k_\alpha k_\beta \sum_{n' \neq i,\, j} \frac{p^\alpha_{i\boldsymbol{k}_0,\, n'\boldsymbol{k}_0} p^\beta_{n'\boldsymbol{k}_0,\, j\boldsymbol{k}_0} + p^\beta_{i\boldsymbol{k}_0,\, n'\boldsymbol{k}_0} p^\alpha_{n'\boldsymbol{k}_0,\, j\boldsymbol{k}_0}}{E(\boldsymbol{k}_0) - E_{n'}(\boldsymbol{k}_0)}. \tag{21.19}$$

The condition for solvability of system (21.19) is the vanishing of the determinant

$$\mathrm{Det}\, \| \mathcal{H} - \boldsymbol{I}E \| = 0, \tag{21.20}$$

which yields an m-th order algebraic equation for the energy $E = E(\boldsymbol{k}_0 + \boldsymbol{k}) - E(\boldsymbol{k}_0)$ with m solutions E_j $(j = 1, 2, \ldots, m)$. Once the E_j have been found, equations (21.18) enable one to determine the coefficients c_{ji} corresponding to the energies E_j which determine the wave function $\varphi'_{j,\, \boldsymbol{k}_0+\boldsymbol{k}}$:

$$\varphi'_{j,\, \boldsymbol{k}_0+\boldsymbol{k}} = \sum_i c_{ji} \left\{ \psi_{i\boldsymbol{k}_0} + \frac{\hbar}{m} \sum_{\alpha,\, n' \neq i} k_\alpha \frac{p^\alpha_{n'\boldsymbol{k}_0,\, i\boldsymbol{k}_0}}{E(\boldsymbol{k}_0) - E_{n'}(\boldsymbol{k}_0)} \psi_{n'\boldsymbol{k}_0} \right\} e^{i\boldsymbol{k}\boldsymbol{x}} \tag{21.21}$$

The coefficients c_{ji} form a unitary matrix $\boldsymbol{c}$.

If equation (21.20) has exactly m different solutions, i.e., the degeneracy is completely removed, then all the minors of the determinant (21.20) differ from zero. In this case system (21.18) has a unique normalized solution, up to an arbitrary phase factor in each column. If there is only partial removal of degeneracy, i.e., equation (21.20) has less than m different solutions, equations (21.18) naturally yield solutions only up to an arbitrary unitary transformation of the wave functions belonging to states that remain degenerate. The energy $E(\boldsymbol{k}_0+\boldsymbol{k})$ and the wave functions $\varphi'_{l,\,\boldsymbol{k}_0+\boldsymbol{k}}$ near the point $\boldsymbol{k}_0$, obtained by solving equations (21.18) and (21.20), automatically satisfy the compatibility conditions for representations at $\boldsymbol{k}_0$ and at neighboring points $\boldsymbol{k}_0+\boldsymbol{k}$: in directions for which group theory tells us that degeneracy is not completely removed, equation (21.20) will have multiple solutions, and the wave functions $\psi_{l,\,\boldsymbol{k}_0+\boldsymbol{k}}=\sqrt{\mathscr{V}}\,\varphi'_{l,\,\boldsymbol{k}_0+\boldsymbol{k}}$ given by (21.21) transform according to the representations $\mathscr{D}_{\nu}^{\boldsymbol{k}_0+\boldsymbol{k}}$ into which the representation $\mathscr{D}_{\mu}^{\boldsymbol{k}_0}$ decomposes according to the compatibility conditions.

For sufficiently small $\boldsymbol{k}$, the k^2 terms in the matrix $\mathscr{H}$ are small compared with the linear $\boldsymbol{k}$ terms, and $\mathscr{H}$ is a linear function of k_α. Therefore, in the vicinity of the point $\boldsymbol{k}_0$ the energy correction $E_i(\boldsymbol{k}_0+\boldsymbol{k})-E_0(\boldsymbol{k}_0)$ is a homogeneous function of first degree in k_α. According to (21.18), the coefficient c_{ij}, as a quotient of two homogeneous functions of first degree, is homogeneous function of degree zero in k_α, so that it does not depend on the magnitude of the vector $\boldsymbol{k}$ but only on its direction.

Thus, in a sufficiently small neighborhood of a band degeneracy point $\boldsymbol{k}_0$ the wave functions $\psi_{l,\,\boldsymbol{k}_0+\boldsymbol{k}}$ depend on the direction of approach to $\boldsymbol{k}_0$. This was established previously in §17. Equation (21.21) enables one to determine just how $\lim\limits_{\boldsymbol{k}\to 0}\psi_{l,\,\boldsymbol{k}_0+\boldsymbol{k}}$ depends on the direction of the vector $\boldsymbol{k}$.

If there are no linear k_α terms, the correction to the energy $E_i(\boldsymbol{k})$ is a homogeneous quadratic function of k_α and $\psi_{l,\,\boldsymbol{k}_0+\boldsymbol{k}}$ depends as before only on the direction of the vector $\boldsymbol{k}$.

The explicit form of the matrix $\mathscr{H}_{lj}(\boldsymbol{k}_0,\,\boldsymbol{k})$, and therefore also of the matrix $\boldsymbol{c}(\boldsymbol{k})$, naturally depends on the specific choice of basis functions $\psi_{i\boldsymbol{k}_0}$, but the energy $E_i(\boldsymbol{k}_0+\boldsymbol{k})$, of course, does not depend on the choice of basis.

A necessary condition for an extremum, as in the case of nondegenerate bands, is the vanishing of all the momentum matrix elements:

$$p^{\alpha}_{i\boldsymbol{k}_0,\,j\boldsymbol{k}_0}=0.$$

All the previous considerations concerning "candidates" for the role of extremum points in the Brillouin zone remain valid for the case of degenerate bands.

If the wave functions at $\boldsymbol{k}_0$ transform according to the representation $\mathscr{D}_{\mu}^{\boldsymbol{k}_0}$, the number of nonzero components of the momentum operator $p^{\alpha}_{i\boldsymbol{k}_0,\,j\boldsymbol{k}_0}$ is equal to the number N_1 of identity representations occurring in the direct product

$$\mathscr{D}_{\mu}^{\boldsymbol{k}_0}\times\mathscr{D}_1\times\mathscr{D}_{\mu}^{*\boldsymbol{k}_0},$$

where $\mathscr{D}_1$ is the vector representation according to which the momentum components transform. The number N_1 can be determined in terms of the characters $\chi_{\mu}^{\boldsymbol{k}_0}$ and χ_1 using (19.43):

$$N_1=\frac{1}{h'}\sum_{g\in G'_{\boldsymbol{k}_0}}|\chi_{\mu}^{\boldsymbol{k}_0}(g)|^2\,\chi_1(g). \qquad (21.22)$$

Note that, in contrast to the case of a nondegenerate band, N_1 depends on the representation $\mathscr{D}_\mu^{\boldsymbol{k}_0}$ according to which the wave functions transform at $\boldsymbol{k}_0$.

Proceeding as in the derivation of equation (21.14) for a nondegenerate band, one can show that the matrix elements of $\mathscr{H}^{(2)}$ quadratic in k_α transform like the matrix element

$$\mathscr{H}_{ij}^{(2)} \sim \langle \psi_{i\boldsymbol{k}_0} | p^\alpha p^\beta + p^\beta p^\alpha | \psi_{j\boldsymbol{k}_0} \rangle. \tag{21.23}$$

Hence it follows that the number of independent constants that determine the matrix $\mathscr{H}^{(2)}$ and, accordingly, the spectrum near the extremum, equal the number N_2 of identity representations occurring in the expansion of the direct product

$$\mathscr{D}_\mu^{*\boldsymbol{k}_0}[\mathscr{D}_1 \times \mathscr{D}_1]\,\mathscr{D}_\mu^{\boldsymbol{k}_0},$$

i.e.,

$$N_2 = \frac{1}{h'} \sum_{g \in G'_{\boldsymbol{k}_0}} \left| \chi_\mu^{\boldsymbol{k}_0}(g) \right|^2 [\chi_1^2(g)]. \tag{21.24}$$

If the star $\{\boldsymbol{k}_0\}$ has several points $\boldsymbol{k}_0^l$, the band is m-fold degenerate at each of them, and the energy $E(\boldsymbol{k}_0^l + \boldsymbol{k})$ near $\boldsymbol{k}_0^l$ is derived, as in the nondegenerate case, from the energy $E(\boldsymbol{k}_0 + \boldsymbol{k})$ near $\boldsymbol{k}_0$ by a transformation g_l which takes $\boldsymbol{k}_0$ into $\boldsymbol{k}_0^l$: $E(\boldsymbol{k}_0^l + \boldsymbol{k}) = E(\boldsymbol{k}_0 + g_l^{-1}\boldsymbol{k})$.

Time reversal. Inclusion of time reversal symmetry in calculation of the energy spectrum has a twofold effect. First, the dimension of the representation according to which the wave functions at the point $\boldsymbol{k}_0$ belonging to the same energy $E(\boldsymbol{k}_0)$ transform must be doubled when time reversal dictates combination of representations at $\boldsymbol{k}_0$. Second, time reversal may impose additional requirements on the wave functions, sometimes causing vanishing of momentum matrix elements which did not vanish from symmetry considerations alone.

We first consider case 1, in which $\boldsymbol{k}_0$ and $-\boldsymbol{k}_0$ are equivalent. This is the case if $\boldsymbol{k}_0 = 0$, or if $\boldsymbol{k}_0$ is on the boundary of the Brillouin zone and $\boldsymbol{k}_0 \equiv -\boldsymbol{k}_0$. As shown in §18, this corresponds to three cases: (a_1), (b_1) and (c_1).

In case (a_1) the linear relation between the functions $K\psi_{s\boldsymbol{k}_0}$ and $\psi_{s\boldsymbol{k}_0}$ results in additional restrictions on the matrix elements of the operators p^α and $p^\alpha p^\beta$ entering $\mathscr{H}$ (21.19). Since p^α is odd with respect to time reversal, it follows from (19.45) that

$$N_1 = \frac{1}{h'} \sum_{g \in G'_{\boldsymbol{k}_0}} \chi_1(g) \{\chi_\mu^{\boldsymbol{k}_0\,2}(g)\}, \qquad N_2 = \frac{1}{h'} \sum_{g \in G'_{\boldsymbol{k}_0}} [\chi_1^2(g)] [\chi_\mu^{\boldsymbol{k}_0\,2}(g)], \tag{21.25}$$

where N_1 and N_2 are the number of linearly independent matrix elements of p^α and $p^\alpha p^\beta + p^\beta p^\alpha$ for $K^2 = 1$.

In cases (b_1) and (c_1) the wave functions $\psi_{s\boldsymbol{k}_0}^*$ and $\psi_{s\boldsymbol{k}_0}$ are linearly independent; in case (b_1) the representations $\mathscr{D}_\mu^{(\boldsymbol{k}_0)}$ and $\mathscr{D}_\mu^{(\boldsymbol{k}_0)*}$ of the full space group are complex and inequivalent, and in case (c_1) they are complex and equivalent. Therefore, in case (b_1) the representations $\mathscr{D}_\mu^{\boldsymbol{k}_0}$ and $\mathscr{D}_\mu^{*\boldsymbol{k}_0}$ of the little group are inequivalent and have complex conjugate characters, whereas in case (c_1) they are equivalent. In cases (b_1) and (c_1) the complete representations $\mathscr{D}_\mu^{(\boldsymbol{k}_0)}$ and $\mathscr{D}_\mu^{*(\boldsymbol{k}_0)}$ of the space group are combined, leading to combination of the representations $\mathscr{D}_\mu^{\boldsymbol{k}_0^l}$ and $\mathscr{D}_\mu^{*\boldsymbol{k}_0^l}$ of the little group at each point $\boldsymbol{k}_0^l$.

Thus, in cases (b_1) and (c_1) the band degeneracy at $\boldsymbol{k}_0$ is doubled, and the basis now consists of $2n_\mu$ functions $\psi_{i\boldsymbol{k}_0}$ and $\psi^*_{i\boldsymbol{k}_0}$, $i = 1, 2, \ldots, n_\mu$, where n_μ is the dimension of the representation $\mathscr{D}^{\boldsymbol{k}_0}_\mu$.

Under the group $G_{\boldsymbol{k}_0}$ these $2n_\mu$ functions transform according to the combined representation, which is reducible and has block-diagonal form:

$$\begin{Vmatrix} \mathscr{D}^{\boldsymbol{k}_0}_\mu & 0 \\ 0 & \mathscr{D}^{*\boldsymbol{k}_0}_\mu \end{Vmatrix}. \tag{21.26}$$

Calculation of the perturbation matrix $\mathscr{H}_{ij}$ (21.19) ($i, j = 1, 2, \ldots, 2n_\mu$) defining the spectrum and the wave functions near $\boldsymbol{k}_0$ involves both diagonal matrix elements of the operators p^α and $p^\alpha p^\beta$ between the functions $\psi_{i\boldsymbol{k}_0}$, $\psi_{j\boldsymbol{k}_0}$ or $\psi^*_{i\boldsymbol{k}_0}$, $\psi^*_{j\boldsymbol{k}_0}$ and off-diagonal matrix elements corresponding to states $\psi_{i\boldsymbol{k}_0}$, $\psi^*_{j\boldsymbol{k}_0}$. The number of independent diagonal matrix elements of the operators p^α and $p^\alpha p^\beta + p^\beta p^\alpha$ is determined according to (19.43) by equations (21.22) and (21.24), while the number of off-diagonal matrix elements N_1 and N_2 in cases (c_1) and (b_1) is determined by equations (21.25).

In case 2 $-\boldsymbol{k}_0$ is not equivalent to $\boldsymbol{k}_0$, but appears in the star $\{\boldsymbol{k}_0\}$. This will occur when there is an element R in G which takes $\boldsymbol{k}_0$ into $-\boldsymbol{k}_0$:

$$R\boldsymbol{k}_0 = -\boldsymbol{k}_0. \tag{21.27}$$

Since the star $\{\boldsymbol{k}_0\}$ is irreducible, this implies that the stars $\{\boldsymbol{k}_0\}$ and $-\{\boldsymbol{k}_0\}$ coincide. Relation (21.27) can be written for each point $\boldsymbol{k}^l_0$ of the star $\{\boldsymbol{k}_0\}$ with a specific $R^l = g_l R g_l^{-1}$, where g_l is an element of G which takes $\boldsymbol{k}_0$ into $\boldsymbol{k}^l_0$.

As shown in §18, in case 2 there are again three possibilities: (a_2), (b_2) and (c_2).

In case (a_2) there is a linear relation between the functions $K\psi_{s\boldsymbol{k}_0} = \psi^*_{s\boldsymbol{k}_0}$ and the functions $\psi_{s\boldsymbol{k}_0}$, and the complete representations $\mathscr{D}^{(\boldsymbol{k}_0)}_\mu$ and $\mathscr{D}^{*(\boldsymbol{k}_0)}_\mu$ of the space group are equivalent.

By (12.19), $\psi_{t,-\boldsymbol{k}_0} = R\psi_{t\boldsymbol{k}_0}$, and so in case ($a_2$), as shown in §18, there is a linear relation between functions $\psi_{t\boldsymbol{k}_0}$ and functions $KR\psi_{t\boldsymbol{k}_0}$ belonging to the same star. This results in additional requirements for the momentum matrix elements, and according to (19.46) N_1 and N_2 are given by

$$\begin{aligned} N_1 &= \frac{1}{2h'} \sum_{g \in G'_{\boldsymbol{k}_0}} \chi_1(g) \left|\chi^{\boldsymbol{k}_0}_\mu(g)\right|^2 - \chi_1(gR)\, \chi^{\boldsymbol{k}_0}_\mu((gR)^2), \\ N_2 &= \frac{1}{2h'} \sum_{g \in G'_{\boldsymbol{k}_0}} \left[\chi^2_1(g)\right] \left|\chi^{\boldsymbol{k}_0}_\mu(g)\right|^2 + \left[\chi^2_1(gR)\right] \chi^{\boldsymbol{k}_0}_\mu((gR)^2). \end{aligned} \tag{21.28}$$

In cases (b_2) and (c_2) the wave functions $\psi^*_{s\boldsymbol{k}_0}$ and $\psi_{t\boldsymbol{k}_0}$ are linearly independent. The complete representations $\mathscr{D}^{(\boldsymbol{k}_0)}$ and $\mathscr{D}^{*(\boldsymbol{k}_0)}$ of the space group must therefore be combined, leading to combination of the representations in the little group at $\boldsymbol{k}_0$.

Since the functions $\psi^*_{s\boldsymbol{k}_0}$ belong to the point $-\boldsymbol{k}_0$, it follows, as shown in §18, that at $\boldsymbol{k}_0$ the representations according to which functions $\psi_{s\boldsymbol{k}_0}$ and $KR\psi_{s\boldsymbol{k}_0}$ belonging to the same point $\boldsymbol{k}_0$ transform, i. e., the representations $\mathscr{D}^{\boldsymbol{k}_0}_\mu(g)$ and $\mathscr{D}^{*\boldsymbol{k}_0}_\mu(R^{-1}gR)$, whose characters are related by (18.35), must be combined. These representations are equivalent in case (c_2), inequivalent in case (b_2). Thus, in cases (b_2) and (c_2) the composite representation has the form

$$\begin{Vmatrix} \mathcal{D}_\mu^{\mathbf{k}_0}(g) & 0 \\ 0 & \mathcal{D}_\mu^{*\mathbf{k}_0}(R^{-1}gR) \end{Vmatrix}. \tag{21.29}$$

The number of independent diagonal elements of the matrix $\mathcal{H}$ for the functions $\psi_{s\mathbf{k}_0}$ and $KR\psi_{s\mathbf{k}_0}$ is determined by substituting the characters of the representations $\mathcal{D}_\mu^{\mathbf{k}_0}$ and $\mathcal{D}_\mu^{*\mathbf{k}_0}(R^{-1}gR)$ into (21.22) and (21.24), respectively. By (19.44), the number of independent off-diagonal elements of the matrix $\mathcal{H}$ in cases (b_2) and (c_2) is given by

$$\begin{aligned} N_1 &= \frac{1}{2h'} \sum_{g \in G'_{\mathbf{k}_0}} \{\chi_1(g)\,\chi_\mu^{\mathbf{k}_0}(g)\,\chi_\mu^{\mathbf{k}_0}(R^{-1}gR) - \chi_1(gR)\,\chi_\mu^{\mathbf{k}_0}((gR)^2)\}, \\ N_2 &= \frac{1}{2h'} \sum_{g \in G'_{\mathbf{k}_0}} [\chi_1^2(g)]\,\chi_\mu^{\mathbf{k}_0}(g)\,\chi_\mu^{\mathbf{k}_0}(R^{-1}gR) + [\chi_1^2(gR)]\,\chi_\mu^{\mathbf{k}_0}((gR)^2). \end{aligned} \tag{21.30}$$

In case 3 the vector $-\mathbf{k}_0$ is not in the star $\{\mathbf{k}_0\}$, i.e., there are no elements in the space group G which take $\mathbf{k}_0$ into $-\mathbf{k}_0$. When this happens, as noted in §18, only case (b_3) is possible. Because of time reversal, the complete representations of the space group $\mathcal{D}_\mu^{(\mathbf{k}_0)}$ and $\mathcal{D}_\mu^{*(\mathbf{k}_0)}$ must be combined, but since their stars $\{\mathbf{k}_0\}$ and $\{-\mathbf{k}_0\}$ do not coincide, the result is not an increase in the dimension of the representation at $\mathbf{k}_0$ and $-\mathbf{k}_0$ but coincidence of the energies at these points, which in case (b_3) is not in itself a consequence of the crystal symmetry.

In case (b_3) the numbers N_1 and N_2 of independent matrix elements of the operators p^α and $p^\alpha p^\beta + p^\beta p^\alpha$ are given by equations (21.22) and (21.24).

Spin-orbit coupling. Inclusion of spin-orbit coupling in band theory may be achieved in two ways.

In the first method, which is valid for spin-orbit coupling of any magnitude, the operator $\mathcal{H}_{so}$ (16.9) is incorporated from the start in the self-consistent potential, so that the new potential $V' = V + \mathcal{H}_{so}$ thereby becomes spin dependent. This approach classifies the wave functions according to the irreducible double-valued representations of the space group.

The method presented above to determine the spectrum is applicable in this case, except that the operator $\mathbf{p}$ must be replaced in the appropriate equations by $\boldsymbol{\pi}$(17.15) which, like $\mathbf{p}$, transforms according to the vector representation $\mathcal{D}_1$ and is odd with respect to time reversal:

$$K^{-1}\pi K = K_0\sigma_y\pi\sigma_y K_0 = -\pi.$$

In the case of double-valued representations, time reversal sometimes leads to combination of representations of the little group and also imposes additional conditions on the matrix elements.

Cases 1, 2 and 3, discussed in §18, are also applicable for double-valued representations, but when using Herring's test (18.32) one should bear in mind that for spinor representations $K^2 = -1$. As noted in §18, this changes the sign of the corresponding terms in the equations for N_1 and N_2.

Thus, in case (a_1), in which there is a linear dependence between the functions $K\psi_{s\mathbf{k}_0}$ and $\psi_{s\mathbf{k}_0}$,

$$N_1 = \frac{1}{h'} \sum_{g \in G'_{\mathbf{k}_0}} \chi_1(g)\,[\chi_\mu^{\mathbf{k}_0\,2}(g)], \tag{21.31}$$

$$N_2=\frac{1}{h'}\sum_{g\in G'_{\boldsymbol{k}_0}}[\chi_1^2(g)]\{\chi_\mu^{\boldsymbol{k}_0\,2}(g)\}. \tag{21.31}$$

In cases (b_1) and (c_1), the functions $K\psi_{s\boldsymbol{k}_0}$ and $\psi_{s\boldsymbol{k}_0}$ are linearly independent, and the representations $\mathscr{D}_\mu^{\boldsymbol{k}_0}$ and $\mathscr{D}_\mu^{*\boldsymbol{k}_0}$ are combined. The complete matrix of the representation of the little group has the form (21.26). The numbers of independent diagonal elements of the matrix $\mathscr{H}$ are determined by equations (21.22) and (21.24), the number of off-diagonal elements by equations (21.31).

In case (a_2), in which $\boldsymbol{k}_0$ and $-\boldsymbol{k}_0$ are in one star, the functions $\psi_{s\boldsymbol{k}_0}$ and $KR\psi_{i\boldsymbol{k}_0}$ are linearly dependent and the numbers N_1 and N_2 are given by

$$\begin{gathered}N_1=\frac{1}{2h'}\sum_{g\in G'_{\boldsymbol{k}_0}}\chi_1(g)\,|\chi_\mu^{\boldsymbol{k}_0}(g)|^2+\chi_1(gR)\,\chi_\mu^{\boldsymbol{k}_0}((gR)^2),\\ N_2=\frac{1}{2h'}\sum_{g\in G'_{\boldsymbol{k}_0}}[\chi_1^2(g)]\,|\chi_\mu^{\boldsymbol{k}_0}(g)|^2-[\chi_1^2(gR)]\,\chi_\mu^{\boldsymbol{k}_0}((gR)^2).\end{gathered} \tag{21.32}$$

In cases (b_2) and (c_2) the representations $\mathscr{D}_\mu^{\boldsymbol{k}_0}(g)$ and $\mathscr{D}_\mu^{*\boldsymbol{k}_0}(R^{-1}gR)$ are combined, and the representation matrix at $\boldsymbol{k}_0$ has the form (21.29). The numbers N_1 and N_2 for the diagonal elements of $\mathscr{H}$ are determined by equations (21.22) and (21.24) and for the off-diagonal elements by equation (19.44):

$$\begin{gathered}N_1=\frac{1}{2h'}\sum_{g\in G'_{\boldsymbol{k}_0}}\chi_1(g)\,\chi_\mu^{\boldsymbol{k}_0}(g)\,\chi_\mu^{\boldsymbol{k}_0}(R^{-1}gR)+\chi_1(gR)\,\chi_\mu^{\boldsymbol{k}_0}((gR)^2),\\ N_2=\frac{1}{2h'}\sum_{g\in G'_{\boldsymbol{k}_0}}\left[\chi_1^2(g)\right]\chi_\mu^{\boldsymbol{k}_0}(g)\,\chi_\mu^{\boldsymbol{k}_0}(R^{-1}gR)-\left[\chi_1^2(gR)\right]\chi_\mu^{\boldsymbol{k}_0}((gR)^2).\end{gathered} \tag{21.33}$$

In case (b_3) time reversal again makes the points $\boldsymbol{k}_0$ and $-\boldsymbol{k}_0$ equivalent, and the numbers N_1 and N_2 are given by equations (21.22) and (21.24).

We stress that when spin-orbit coupling is introduced, it is the characters of the double-valued representations of the space groups that appear in the corresponding equations.

However, spin-orbit coupling is usually of a considerably smaller order of magnitude than the atomic energy, and therefore it is frequently more convenient to use a second method, which treats spin-orbit coupling as a perturbation. This method brings out the properties of the band structure due specifically to spin-orbit coupling, and relativistically small terms can be isolated in the band parameters.

Since in many semiconductors the interband spacing is also much less than the atomic energy, various relations may hold between the spin-orbit splitting and the separation to the nearest band, which, in particular, may be the width of the band gap.

We first consider $\boldsymbol{kp}$-theory in the case that spin-orbit coupling is significantly less than the separation to the nearest band. In this case we can construct the spectrum with spin-orbit coupling only for the band under consideration.

Suppose that when spin is neglected the band is m-fold degenerate at $\boldsymbol{k}_0$ and the wave functions $\psi_{s\boldsymbol{k}_0}$ transform according to a certain representation $\mathscr{D}^{\boldsymbol{k}_0}$ of the little group (which may also be reducible owing to combination of irreducible representations). Then, when spin is included, we have $2m$ functions $\psi_{s\boldsymbol{k}_0}\alpha$ and $\psi_{s\boldsymbol{k}_0}\beta$, forming a basis for the representation $\mathscr{D}^{\boldsymbol{k}_0}\times\mathscr{D}_{1/2}$. This representation is in general reducible and may be expanded in terms of irreducible double-valued representations of the space group. This means that in general spin-orbit coupling lifts the m-fold band degeneracy at the point $\boldsymbol{k}_0$.

To calculate the spin-orbit splitting and to obtain the correct functions at the point $\boldsymbol{k}_0$, which in this case are linear combinations of the functions $\psi_{s\boldsymbol{k}_0}\alpha$ and $\psi_{s\boldsymbol{k}_0}\beta$, it is necessary to diagonalize the matrix $\mathscr{H}_{so}$ for the functions $\psi_{s\boldsymbol{k}_0}\alpha$ and $\psi_{s\boldsymbol{k}_0}\beta$. This reduces to solving the system of equations (21.20), with $\mathscr{H}_{so}$ in place of $\mathscr{H}$. Solving the characteristic equation of the system, we obtain $\mu\leqslant 2m$ solutions $E_i(\boldsymbol{k}_0)$, separated by energy intervals Δ_i which describe the spin-orbit splitting.

To determine the spectrum near $\boldsymbol{k}_0$ for each of the split-off bands $E_i(\boldsymbol{k}_0)$, we can use the above equations, replacing the operator $\boldsymbol{p}$ everywhere by $\boldsymbol{\pi}$. As opposed to the first method, here one can easily identify the band from which the new band has split off.

Separate treatment of each split-off band is possible only for quite small values of $\boldsymbol{k}$, when

$$E_i(\boldsymbol{k}_0+\boldsymbol{k})-E_i(\boldsymbol{k}_0)\ll\Delta_i.$$

If the energy measured from the bottom of the i-th split-off band becomes comparable with the separation to the nearest split-off bands, it is necessary to solve the $2m$-th order equation (21.20), including the spin-orbit coupling $\mathscr{H}_{so}$ along with the terms $k_\alpha p^\alpha$ and $k_\alpha k_\beta p^\alpha p^\beta$ in the matrix $\mathscr{H}$. If the spin-orbit splitting is of the same order as or even exceeds the separation to the nearest band, in particular, the width of the band gap, then in order to calculate the wave functions and the energy at the point $\boldsymbol{k}_0$ with allowance for $\mathscr{H}_{so}$ the basis must include the $2(m_1+m_2)$ functions belonging to both bands,

$$\psi_{s(n_1)\boldsymbol{k}_0}\alpha,\quad \psi_{s(n_1)\boldsymbol{k}_0}\beta,\quad \psi_{t(n_2)\boldsymbol{k}_0}\alpha,\quad \psi_{t(n_2)\boldsymbol{k}_0}\beta,$$
$$\text{where } s(n_1)=1,\ 2,\ \ldots,\ m_1;\quad t(n_2)=1,\ 2,\ \ldots,\ m_2.$$

To determine the correct wave functions at the point $\boldsymbol{k}_0$, which in this case may include all $2(m_1+m_2)$ functions of both bands, it is necessary to solve system (21.20), with the matrix $\mathscr{H}_{so}$, evaluated relative to this basis, playing the part of a perturbation, and the energy $E_i(\boldsymbol{k}_0)$ allowing for $\mathscr{H}_{so}$ is given by the roots of the corresponding characteristic equation (21.20).

The energy near $\boldsymbol{k}_0$ for each of the bands obtained in this way is determined by the method considered above. It may happen that when $\mathscr{H}_{so}$ is ignored the band has zero slope at $\boldsymbol{k}_0$, but because of the relativistic terms, the energy expansion near $\boldsymbol{k}_0$ may contain nonzero terms linear in k_α. In the linear spin-orbit coupling approximation, such terms arise only in second order perturbation theory for the operators $(\mathscr{H}_{so})_{\boldsymbol{k}_0}$ and $\hbar\boldsymbol{kp}/m$, as cross products of their matrix elements.

In first order perturbation theory, the contribution of the relativistic term to $\boldsymbol{\pi}$, i.e., of the operator $\frac{\hbar}{4mc^2}[\sigma\nabla V]$, is zero.

Indeed, this contribution is proportional to the matrix elements $\langle n\boldsymbol{k}_0|\frac{\partial V}{\partial x_\alpha}|n'\boldsymbol{k}_0\rangle$, where $n'=n$ for a nondegenerate band, while for a degenerate band the states

n and n' have the same energy. Calculating the matrix element of the commutator

$$\left\{ \frac{\partial}{\partial x_\alpha} \mathcal{H}_0 \right\} = \frac{\partial V}{\partial x_\alpha}, \tag{21.34}$$

we obtain

$$\langle n\boldsymbol{k}_0 | \frac{\partial V}{\partial x_\alpha} | n'\boldsymbol{k}_0 \rangle = i\left(E_n(\boldsymbol{k}_0) - E_{n'}(\boldsymbol{k}_0)\right) p^\alpha_{n\boldsymbol{k}_0,\, n'\boldsymbol{k}_0}, \tag{21.35}$$

which implies that for all the states n and n' with the same energy $E_n(\boldsymbol{k}_0) = E_{n'}(\boldsymbol{k}_0)$

$$\langle n\boldsymbol{k}_0 | \frac{\partial V}{\partial x_\alpha} | n'\boldsymbol{k}_0 \rangle = 0.$$

In cases when the linear k_α terms are relativistically small at $\boldsymbol{k}_0$, the point $\boldsymbol{k}_0$ is likely to be an extremum point, since small linear terms cause only a slight displacement from $\boldsymbol{k}_0$. The small linear k_α terms should then also be included in the matrix $\mathcal{H}$ along with the quadratic terms. The eigenvalues of the matrix $\mathcal{H}$ give the spectrum near $\boldsymbol{k}_0$ and the new positions $\boldsymbol{k}_0'$ of the band extrema.

The nature of the linear k_α terms can be determined from group-theoretic considerations. Use of single-valued representations of the space groups gives selection rules for the matrix elements of the operator $\boldsymbol{p}$, and the corresponding formulas for double-valued representations determine the number of nonzero components of the terms linear in $\boldsymbol{k}$ that appear in $\mathcal{H}$ when spin-orbit coupling is included. If the selection rules for double-valued representations dictate that $\pi^\alpha \neq 0$ for some component, but single-valued representations give $p^\alpha = 0$, then the matrix element in question is relativistic.

$\boldsymbol{kp}$-theory makes it possible to obtain the spectrum $E_i(\boldsymbol{k}_0 + \boldsymbol{k})$ near the point $\boldsymbol{k}_0$ in higher approximations as well.

In this context, we would like to define some dimensionless parameter of the energy expansion near the point $\boldsymbol{k}_0$, and to know whether the second-order expansion of the energy near the extremum is sufficient to describe the various physical phenomena in semiconductors.

As we see from the structure of the perturbation equations (21.10), a suitable dimensionless parameter is

$$\frac{\hbar}{m} \frac{\boldsymbol{k}\boldsymbol{p}_{n\boldsymbol{k}_0,\, n'\boldsymbol{k}_0}}{E_n(\boldsymbol{k}_0) - E_{n'}(\boldsymbol{k}_0)}, \tag{21.36}$$

where $\boldsymbol{p}_{n\boldsymbol{k}_0,\, n'\boldsymbol{k}_0}$ is the matrix element of the momentum (or of the operator $\boldsymbol{\pi}$ if spin-orbit coupling is included) between bands n and n', and $E_n(\boldsymbol{k}_0) - E_{n'}(\boldsymbol{k}_0)$ is the band separation at $\boldsymbol{k}_0$.

Using the expression (21.16) for the energy near an extremum, it is easily seen that the expansion parameter may also be written as $\sqrt{E_k/\overline{E}}$, where E_k is the carrier kinetic energy measured from the bottom of the band, and E is the separation of the band under consideration from the other bands n', averaged in some way. If the extrema of the valence band and the conduction band are located at the same point $\boldsymbol{k}_0$ and the matrix element of the momentum $\boldsymbol{p}_{v\boldsymbol{k}_0,\, c\boldsymbol{k}_0}$ between the valence band and the conduction band differs from zero, the major contribution to the energy in the sums (21.10) and (21.19) is the interaction of these two bands, and $\overline{E} \simeq E_g$. In this case the criterion for the applicability of the second-order expansion is

$$\sqrt{E_k/E_g} \ll 1. \tag{21.37}$$

Under conditions of thermodynamic equilibrium, $E_k \simeq kT$ or μ for a nondegenerate and degenerate electron gas, respectively, where k is Boltzmann's constant, T the absolute temperature and μ the Fermi level; thus it follows from (21.37) that in semiconductors with a sufficiently wide band gap, at relatively low temperatures and current carrier concentrations, the criterion (21.37) may be fulfilled with reasonable accuracy.

Two-band model. In semiconductors with a narrow band gap, however, marked deviations from the parabolic dependence of $E(\mathbf{k}_0+\mathbf{k})$ on $\mathbf{k}$ appear even at comparatively low temperatures and current carrier concentrations. To describe the spectrum in this case, the so-called two-band approximation is used. In the two-band approximation, the Hamiltonian $\mathcal{H}(\mathbf{k}_0+\mathbf{k})$ is constructed from the start for two close-lying bands, incorporating the $\mathbf{kp}$-interaction between these bands accurately, and the interaction with other, more distant bands is treated by perturbation theory.

In this model, the matrix elements $\mathcal{H}_{nn'}$ are determined by the general equation (15.46). In the second order in $\mathbf{k}$, the diagonal blocks of the matrix $\mathcal{H}$, i.e., the matrix elements between states within the conduction band (denoted by indices s, s') or states of the valence band (indices t, t'), incorporate the contributions from the more distant bands:

$$\begin{gathered}
\mathcal{H}_{s(c),\,s'(c)} = \frac{\hbar^2}{m^2} \sum_{n' \neq c,\,v} k_\alpha k_\beta \left(\frac{p^\alpha_{s(c)\mathbf{k}_0,\,n'\mathbf{k}_0} p^\beta_{n'\mathbf{k}_0,\,s'(c)\mathbf{k}_0}}{E_c(\mathbf{k}_0) - E_{n'}(\mathbf{k}_0)} + \right. \\
\left. + \frac{p^\beta_{s(c)\mathbf{k}_0,\,n'\mathbf{k}_0} p^\alpha_{n'\mathbf{k}_0,\,s'(c)\mathbf{k}_0}}{E_c(\mathbf{k}_0) - E_{n'}(\mathbf{k}_0)} \right) + \frac{\hbar^2 k^2}{2m} \delta_{ss'} + E_g \delta_{ss'}, \\
\mathcal{H}_{t(v),\,t'(v)} = \frac{\hbar^2}{m^2} \sum_{\alpha\beta} k_\alpha k_\beta \left\{ \sum_{n' \neq c,\,v} \frac{p^\alpha_{t(v)\mathbf{k}_0,\,n'\mathbf{k}_0} p^\beta_{n'\mathbf{k}_0,\,t'(v)\mathbf{k}_0}}{E_v(\mathbf{k}_0) - E_{n'}(\mathbf{k}_0)} + \right. \\
\left. + \frac{p^\beta_{t(v)\mathbf{k}_0,\,n'\mathbf{k}_0} p^\alpha_{n'\mathbf{k}_0,\,t'(v)\mathbf{k}_0}}{E_v(\mathbf{k}_0) - E_{n'}(\mathbf{k}_0)} \right\} + \frac{\hbar^2 k^2}{2m} \delta_{tt'}.
\end{gathered} \tag{21.38}$$

The off-diagonal blocks of $\mathcal{H}$, between states within the valence band and the conduction band, contain terms linear in k_α:

$$\mathcal{H}_{s(c),\,t(v)} = \frac{\hbar}{m} \sum_\alpha k_\alpha p^\alpha_{s(c)\mathbf{k}_0,\,t(v)\mathbf{k}_0}. \tag{21.39}$$

When necessary, interband terms of higher orders in $\mathbf{k}$ can be taken into account.

According to equation (15.34), a basis for this representation is formed by the functions

$$\begin{gathered}
\psi'_{s(c)\mathbf{k}_0} = \psi_{s(c)\mathbf{k}_0} + \frac{\hbar}{m} \sum_{n' \neq c,\,v,\,\alpha} \frac{k_\alpha p^\alpha_{n'\mathbf{k}_0,\,s(c)\mathbf{k}_0} \psi_{n'\mathbf{k}_0}}{E_c(\mathbf{k}_0) - E_{n'}(\mathbf{k}_0)}, \\
\psi'_{t(v)\mathbf{k}_0} = \psi_{t(v)\mathbf{k}_0} + \frac{\hbar}{m} \sum_{n' \neq c,\,v,\,\alpha} \frac{k_\alpha p^\alpha_{n'\mathbf{k}_0,\,t(v)\mathbf{k}_0} \psi_{n'\mathbf{k}_0}}{E_v(\mathbf{k}_0) - E_{n'}(\mathbf{k}_0)}.
\end{gathered} \tag{21.40}$$

For small $\mathbf{k}$, when the energy E_k measured from the bottom of the corresponding band E_c or E_v is small compared to $E_g = E_c(\mathbf{k}_0) - E_v(\mathbf{k}_0)$, the two-band model yields the same results as treatment of each band separately, but if $E_k \sim E_g$ the dependence of the energy on $\boldsymbol{k}$ in each band becomes nonparabolic.

The accuracy of the two-band approximation depends on the parameter $\sqrt{E_k/\bar{E}}$, where $\bar{E}$ is the separation to more distant bands; it is usually quite adequate for a satisfactory description of the spectrum in semiconductors with a narrow band gap.

§22. ELECTRON IN A CRYSTAL IN AN EXTERNAL FIELD

The equation of motion of an electron in a crystal, in the presence of external fields which perturb the periodic lattice field $V_0(\boldsymbol{x})$ is

$$\mathscr{H}\Psi = \left[\mathscr{H}_0\left(\boldsymbol{p} + \frac{e}{c}\mathscr{A}(\boldsymbol{x}, t)\right) + U(\boldsymbol{x}, t) + \frac{1}{2}g_0\mu_0(\sigma\boldsymbol{H})\right]\Psi = i\hbar\frac{\partial\Psi}{\partial t}, \qquad (22.1)$$

where

$$\mathscr{H}_0(\boldsymbol{p}) = \frac{p^2}{2m} + V_0(\boldsymbol{x}) + \frac{\hbar}{4m^2c^2}([\sigma\nabla V_0]\,\boldsymbol{p}) \qquad (22.1\text{a})$$

is the Schrödinger operator for an ideal lattice, $\mu_0 = e\hbar/2mc$ the Bohr magneton, and g_0 the g-factor for free electrons.

The external fields, electric and magnetic, are represented by a vector potential $\mathscr{A}(\boldsymbol{x}, t)$ and a scalar potential $U(\boldsymbol{x}, t)$:

$$\mathscr{E} = -\nabla U + \frac{1}{c}\frac{\partial\mathscr{A}}{\partial t}, \quad \boldsymbol{H} = \operatorname{rot}\mathscr{A}.$$

Although the gauge may always be chosen so that $U = 0$, we shall consider the general case, since in certain problems it is precisely the scalar potential that is more conveniently specified.

Exact solution of equation (22.1) is even more complicated than determination of the spectrum and wave functions in an ideal crystal. Let us assume, however, that the external fields are sufficiently "gentle"; in other words, they vary only slightly over distances of the order of the lattice constant, and the frequency of their time dependence is significantly smaller than $\Delta E/\hbar$, where ΔE is the separation of the nearest bands measured vertically, i.e., for the same $\boldsymbol{k}_0$. In that case, the motion of electrons near an extremum $\boldsymbol{k}_0$ in a perturbed periodic field is equivalent to the motion of a particle with a definite effective mass in a slowly varying external field. In the case of degenerate bands, the motion of this particle is described by a system of equations with the same operator $\mathscr{H}(\boldsymbol{k})$ (21.19) that determines the carrier spectrum near the extremum according to (21.20). Therefore, this approach is known as effective mass theory.

Write the wave function Ψ as the product of a slowly varying function $\mathscr{F}_n(\boldsymbol{x}, t)$ and the eigenfunctions $\psi_{n\boldsymbol{k}_0}(\boldsymbol{x})$ of the unperturbed Hamiltonian $\mathscr{H}_0$ at the point $\boldsymbol{k}_0$ (the "slow" functions are also known as envelope functions):

$$\Psi = \sum_n \mathscr{F}_n(\boldsymbol{x}, t)\,\psi_{n\boldsymbol{k}_0}. \qquad (22.2)$$

In the general case, the band is assumed to be degenerate, i.e., different functions $\psi_{n\boldsymbol{k}_0}$, including states with different spins, may belong to the same energy.

Expand the function $\mathscr{F}_n(\boldsymbol{x}, t)$ in Fourier series:

$$\mathscr{F}_n(\boldsymbol{x}, t) = \frac{1}{\sqrt{\mathscr{V}}} \sum_{\boldsymbol{k}} c_{n\boldsymbol{k}} e^{i\boldsymbol{k}\boldsymbol{x}}, \tag{22.3}$$

where

$$c_{n\boldsymbol{k}}(t) = \frac{1}{\sqrt{\mathscr{V}}} \int \mathscr{F}_n(\boldsymbol{x}, t) e^{-i\boldsymbol{k}\boldsymbol{x}} d\boldsymbol{x}. \tag{22.4}$$

Here, as in §21, the crystal volume $\mathscr{V}$ is assumed finite and so the values of $\boldsymbol{k}$ are discrete. If the functions $\mathscr{F}_n(\boldsymbol{x}, t)$ are sufficiently smooth, the main contribution to (22.3) comes from small values of $\boldsymbol{k}$. Substituting (22.3) into (22.2), we obtain an expansion of $\Psi(\boldsymbol{x}, t)$ in terms of the functions $\varphi_{n\boldsymbol{k}} = \frac{1}{\sqrt{\mathscr{V}}} e^{i\boldsymbol{k}\boldsymbol{x}} \psi_{n\boldsymbol{k}_0} = \frac{1}{\sqrt{\mathscr{V}}} e^{i(\boldsymbol{k}_0+\boldsymbol{k})\boldsymbol{x}} u_{n\boldsymbol{k}_0}$:

$$\Psi = \sum_n \sum_{\boldsymbol{k}} c_{n\boldsymbol{k}}(t) \varphi_{n\boldsymbol{k}}. \tag{22.5}$$

To obtain the $n\boldsymbol{k}$-representation, we substitute (22.5) into (22.1), multiply the right and left sides of the equation by $\varphi^*_{n'\boldsymbol{k}'} = \frac{1}{\sqrt{\mathscr{V}}} \psi^*_{n'\boldsymbol{k}_0} e^{-i\boldsymbol{k}'\boldsymbol{x}}$ and integrate over $\boldsymbol{x}$, noting that according to (21.4), (21.8) and (21.7)

$$\langle \varphi_{n'\boldsymbol{k}'} \varphi_{n\boldsymbol{k}} \rangle = \delta_{nn'} \delta_{\boldsymbol{k}\boldsymbol{k}'}, \quad \langle \varphi_{n'\boldsymbol{k}'} \pi \varphi_{n\boldsymbol{k}} \rangle = \pi_{n'n} \delta_{\boldsymbol{k}\boldsymbol{k}'},$$
$$\mathscr{H}(\boldsymbol{p}) e^{i\boldsymbol{k}\boldsymbol{x}} = e^{i\boldsymbol{k}\boldsymbol{x}} \mathscr{H}(\boldsymbol{p} + \hbar\boldsymbol{k}).$$

Therefore,

$$\begin{aligned} \langle \varphi_{n'\boldsymbol{k}'} | \mathscr{H} | \varphi_{n\boldsymbol{k}} \rangle &= \frac{1}{\mathscr{V}} \langle \psi_{n'\boldsymbol{k}_0} e^{i(\boldsymbol{k}-\boldsymbol{k}')\boldsymbol{x}} | \mathscr{H}_0(\boldsymbol{p}) | \psi_{n\boldsymbol{k}_0} \rangle + \\ &+ \frac{1}{2m} \frac{1}{\mathscr{V}} \langle \psi_{n'\boldsymbol{k}_0} e^{i(\boldsymbol{k}-\boldsymbol{k}')\boldsymbol{x}} | 2\hbar\boldsymbol{k}\pi + \frac{e}{c} (\mathcal{A}\pi + \pi\mathcal{A}) | \psi_{n\boldsymbol{k}_0} \rangle + \\ &+ \frac{1}{2m} \langle \varphi_{n'\boldsymbol{k}'} | \left(\hbar\boldsymbol{k} + \frac{e}{c}\mathcal{A}\right)^2 | \varphi_{n\boldsymbol{k}} \rangle + \langle \varphi_{n'\boldsymbol{k}'} | U(\boldsymbol{x}, t) | \varphi_{n\boldsymbol{k}} \rangle + \\ &+ \frac{1}{2} g_0 \mu_0 \langle \varphi_{n'\boldsymbol{k}'} | \sigma \boldsymbol{H} | \varphi_{n\boldsymbol{k}} \rangle. \end{aligned} \tag{22.6}$$

We next expand the functions $\mathcal{A}(\boldsymbol{x})$, $\mathcal{A}^2(\boldsymbol{x})$, $U(\boldsymbol{x})$, $\boldsymbol{H}(\boldsymbol{x})$ in Fourier series:

$$\begin{aligned} \mathcal{A}(\boldsymbol{x}) &= \sum_q \mathcal{A}_q e^{iq\boldsymbol{x}}, \qquad \mathcal{A}^2(\boldsymbol{x}) = \sum_q \mathcal{A}^{(2)}_q e^{iq\boldsymbol{x}}, \\ U(\boldsymbol{x}) &= \sum_q U_q e^{iq\boldsymbol{x}}, \qquad \boldsymbol{H}(\boldsymbol{x}) = \sum_q \boldsymbol{H}_q e^{iq\boldsymbol{x}}, \end{aligned} \tag{22.7}$$

where

$$\mathcal{A}_q = \frac{1}{\mathscr{V}} \int \mathcal{A}(\boldsymbol{x}) e^{-iq\boldsymbol{x}} d\boldsymbol{x}, \qquad \mathcal{A}^{(2)}_q = \frac{1}{\mathscr{V}} \int \mathcal{A}^2(\boldsymbol{x}) e^{-iq\boldsymbol{x}} d\boldsymbol{x} \quad \text{etc.};$$

when $\mathscr{V}$ is finite, the vectors $\boldsymbol{q}$, like $\boldsymbol{k}$, take only discrete values.

Equation (22.6) involves the matrix elements either of these functions themselves or of the products $\mathcal{A}\pi$ or $\pi\mathcal{A}$. Expanding the product $\psi^*_{n'\boldsymbol{k}_0}\psi_{n\boldsymbol{k}_0}$, which is a periodic function of the coordinates, in Fourier series with respect to the reciprocal lattice vectors $\boldsymbol{b}_M$,

$$\psi^*_{n'\boldsymbol{k}_0} \psi_{n\boldsymbol{k}_0} = \mathscr{V} \sum_M B^{n'n}_M e^{-i\boldsymbol{b}_M\boldsymbol{x}},$$

we find that for any function $f(\boldsymbol{x}) = \sum_{\boldsymbol{q}} f_{\boldsymbol{q}} e^{i\boldsymbol{q}\boldsymbol{x}}$

$$\langle \varphi_{n'\boldsymbol{k}'} | f(\boldsymbol{x}) | \varphi_{n\boldsymbol{k}} \rangle = \sum_{M\boldsymbol{q}} f_{\boldsymbol{q}} B_M^{n'n} \int e^{i(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{q}-\boldsymbol{b}_M)\boldsymbol{x}}\, d\boldsymbol{x} = \mathcal{V} \sum_M f_{\boldsymbol{k}'-\boldsymbol{k}+\boldsymbol{b}_M} B_M^{n'n}.$$

If $f(\boldsymbol{x})$ varies slowly over distances of the order of the lattice constant, all the components $f_{\boldsymbol{k}'+\boldsymbol{b}_M-\boldsymbol{k}}$ except the component with $M = 0$ may be neglected, and since by (21.4) we have $\mathcal{V} B_0^{n'n} = \frac{1}{\mathcal{V}} \langle \psi_{n'\boldsymbol{k}_0} \psi_{n\boldsymbol{k}_0} \rangle = \delta_{nn'}$, it follows that

$$\langle \varphi_{n'\boldsymbol{k}'} | f(\boldsymbol{x}) | \varphi_{n\boldsymbol{k}} \rangle = f_{\boldsymbol{k}-\boldsymbol{k}'} \delta_{nn'}. \tag{22.8}$$

In a similar manner we compute the Fourier series of the periodic function

$$\psi_{n'\boldsymbol{k}_0} \pi \psi_{n\boldsymbol{k}_0} = \mathcal{V} \sum_M \boldsymbol{R}_M^{n'n} e^{i\boldsymbol{b}_M\boldsymbol{x}}.$$

Then, after integration over $\boldsymbol{x}$ (assuming that $\mathcal{A}(\boldsymbol{x})$ is sufficiently smooth), the only term remaining from the sum

$$\frac{1}{\mathcal{V}} \langle \psi_{n'\boldsymbol{k}_0} e^{i(\boldsymbol{k}-\boldsymbol{k}')\boldsymbol{x}} | \mathcal{A}(\boldsymbol{x}) \pi | \psi_{n\boldsymbol{k}_0} \rangle = \mathcal{V} \sum_M \mathcal{A}_{\boldsymbol{k}'+\boldsymbol{b}_M-\boldsymbol{k}} \boldsymbol{R}_M^{n'n}$$

will again be that with $M = 0$. Since

$$\mathcal{V} \boldsymbol{R}_0^{n'n} = \frac{1}{\mathcal{V}} \langle \psi_{n'\boldsymbol{k}_0} \pi \psi_{n\boldsymbol{k}_0} \rangle = \pi_{n'n}, \text{ and } \pi e^{i\boldsymbol{q}\boldsymbol{x}} = e^{i\boldsymbol{q}\boldsymbol{x}} (\pi + \hbar \boldsymbol{q}),$$

it follows that

$$\begin{aligned} &\frac{1}{\mathcal{V}} \langle \psi_{n'\boldsymbol{k}_0} e^{i(\boldsymbol{k}-\boldsymbol{k}')\boldsymbol{x}} | \mathcal{A}(\boldsymbol{x}) \pi | \psi_{n\boldsymbol{k}_0} \rangle = \mathcal{A}_{\boldsymbol{k}'-\boldsymbol{k}} \pi_{n'n}, \\ &\frac{1}{\mathcal{V}} \langle \psi_{n'\boldsymbol{k}_0} e^{i(\boldsymbol{k}-\boldsymbol{k}')\boldsymbol{x}} | \pi \mathcal{A}(\boldsymbol{x}) | \psi_{n\boldsymbol{k}_0} \rangle = \mathcal{A}_{\boldsymbol{k}'-\boldsymbol{k}} (\pi_{n'n} + \hbar(\boldsymbol{k}' - \boldsymbol{k}) \delta_{nn'}). \end{aligned} \tag{22.9}$$

Using (22.7)–(22.9), we can represent the matrix $\mathcal{H}_{n'\boldsymbol{k}',n\boldsymbol{k}}$ as the sum of three matrices: a matrix $\mathcal{H}_0$,

$$(\mathcal{H}_0)_{n'\boldsymbol{k}',n\boldsymbol{k}} = E_n \delta_{n'n} \delta_{\boldsymbol{k}'\boldsymbol{k}}, \quad \text{where} \quad E_n = E_n(\boldsymbol{k}_0), \tag{22.10a}$$

a matrix $\mathcal{H}_1$ containing only intraband terms,

$$\begin{aligned} (\mathcal{H}_1)_{n'\boldsymbol{k}',n\boldsymbol{k}} = \frac{1}{2m} \Big\{ \hbar^2 k^2 \delta_{\boldsymbol{k}\boldsymbol{k}'} + \frac{e\hbar}{c} (\boldsymbol{k}' + \boldsymbol{k}) \mathcal{A}_{\boldsymbol{k}'-\boldsymbol{k}} + \frac{e^2}{c^2} \mathcal{A}^{(2)}_{\boldsymbol{k}'-\boldsymbol{k}} \Big\} \delta_{nn'} + \\ + \ U_{\boldsymbol{k}'-\boldsymbol{k}} \delta_{nn'} + \frac{1}{2} g_0 \mu_0 \sigma_{n'n} \boldsymbol{H}_{\boldsymbol{k}'-\boldsymbol{k}}, \end{aligned} \tag{22.10b}$$

and a matrix $\mathcal{H}_2$ containing the interband terms,

$$(\mathcal{H}_2)_{n'\boldsymbol{k}',n\boldsymbol{k}} = \frac{1}{m} \Big(\hbar \boldsymbol{k} \delta_{\boldsymbol{k}\boldsymbol{k}'} + \frac{e}{c} \mathcal{A}_{\boldsymbol{k}'-\boldsymbol{k}} \Big) \pi_{n'n}. \tag{22.10c}$$

As a result we obtain the system of equations

$$\sum_n \sum_{\boldsymbol{k}} \{ (\mathcal{H}_1 + \mathcal{H}_2)_{n'\boldsymbol{k}',n\boldsymbol{k}} + E_n \delta_{nn'} \delta_{\boldsymbol{k}\boldsymbol{k}'} \} c_{n\boldsymbol{k}} = i\hbar \frac{\partial c_{n'\boldsymbol{k}'}}{\partial t}. \tag{22.11}$$

The summation in system (22.11) extends over all the bands. To obtain an approximate equation involving only functions $c_{m\boldsymbol{k}}$ for one band, we must transform away the interband terms which appear in $\mathcal{H}_2$. If equation (22.1) included only the scalar potential, the operator $\mathcal{H}_2$ would be the same as $\mathcal{H}_2$ in (21.9), and we could eliminate the interband terms by the same procedure as was used there to derive (21.10). But if $\mathcal{H}_2$ also includes the term $\mathcal{A}\pi$ and the vector potential depends explicitly on t, it follows from (15.40) and (15.43) that the transformation matrix S (15.33) also depends explicitly on t.

Thus the transformation (15.33) converts the right side of equation (22.11) into

$$i\hbar e^{-\boldsymbol{S}}\frac{\partial}{\partial t}c_{n'\boldsymbol{k}'}=i\hbar\left(\frac{\partial\boldsymbol{S}}{\partial t}+\frac{\partial}{\partial t}\right)\bar{c}_{n'\boldsymbol{k}'}, \qquad (22.11\text{a})$$

where $\bar{\boldsymbol{c}}=e^{-\boldsymbol{S}}\boldsymbol{c}$ by (15.33a).

If the characteristic frequency of the variation of $\boldsymbol{\mathcal{A}}$, and consequently of $\boldsymbol{S}$, is ω, we see from (15.44) that the first term in (22.11a) is

$$\sum_{s}\frac{\hbar\omega}{E_s-E_{n'}}(\mathcal{H}_2)_{n'\boldsymbol{k}',\,s\boldsymbol{k}}\bar{c}_{n'\boldsymbol{k}'}.$$

Consequently, it is a factor of $\hbar\omega/(E_s-E_{n'})$ smaller than the principal term $(\mathcal{H}_2)_{n'\boldsymbol{k}',\,s\boldsymbol{k}}$ transformed away by (15.33). If $\hbar\omega$ is small compared to the separation to the nearest s-band for which the matrix element π_{ms} is nonzero, this term can be neglected. If this condition is not fulfilled, effective mass theory is not applicable, at any rate in the single-band model.

After dropping this term, we obtain a system of equations involving only intraband matrix elements:

$$\sum_{m\boldsymbol{k}}\mathcal{H}_{m'\boldsymbol{k}',\,m\boldsymbol{k}}\bar{\bar{c}}_{m\boldsymbol{k}}=i\hbar\frac{\partial}{\partial t}\bar{\bar{c}}_{m'\boldsymbol{k}'},\ \text{where}\quad \bar{\bar{c}}_{m\boldsymbol{k}}=\bar{c}_{m\boldsymbol{k}}\,e^{-iE_mt/\hbar}, \qquad (22.12)$$

and by (15.47), in second order perturbation theory,

$$\mathcal{H}_{m'\boldsymbol{k}',\,m\boldsymbol{k}}=(\mathcal{H}_1)_{m'\boldsymbol{k}',\,m\boldsymbol{k}}+(\mathcal{H}_2)_{m'\boldsymbol{k}',\,m\boldsymbol{k}}+\sum_{s\boldsymbol{k}''}\frac{(\mathcal{H}_2)_{m'\boldsymbol{k}',\,s\boldsymbol{k}''}(\mathcal{H}_2)_{s\boldsymbol{k}'',\,m\boldsymbol{k}}}{E_m-E_s}. \qquad (22.13)$$

Here, as against (15.47), each state is characterized by two indices n and $\boldsymbol{k}$, and therefore the summation over the intermediate states includes the summation over $\boldsymbol{k}''$, but, as is evident from (22.10a), the energies E_m and E_s do not depend on $\boldsymbol{k}''$. According to (22.3), in order to go back to the x-presentation in (22.12) we multiply the right and left sides by $(1/\sqrt{\mathcal{V}})e^{i\boldsymbol{k}'\boldsymbol{x}}$ and sum over $\boldsymbol{k}'$.

It is clear from (22.10) that here there are sums of four types. First we have sums of the type $\sum_{\boldsymbol{k}'\boldsymbol{k}}f_{\boldsymbol{k}'-\boldsymbol{k}}e^{i\boldsymbol{k}\boldsymbol{x}}c_{m\boldsymbol{k}}$, where the f_q are the Fourier components of U, H_α, $\mathcal{A}_\alpha$ or $\mathcal{A}^2$. According to (22.3) and (22.7), these sums may be transformed:

$$\frac{1}{\sqrt{\mathcal{V}}}\sum_{\boldsymbol{k}\boldsymbol{k}'}f_{\boldsymbol{k}'-\boldsymbol{k}}e^{i\boldsymbol{k}'\boldsymbol{x}}\bar{\bar{c}}_{m\boldsymbol{k}}=\frac{1}{\sqrt{\mathcal{V}}}\sum_{\boldsymbol{q}}f_{\boldsymbol{q}}e^{i\boldsymbol{q}\boldsymbol{x}}\sum_{\boldsymbol{k}}\bar{\bar{c}}_{m\boldsymbol{k}}e^{i\boldsymbol{k}\boldsymbol{x}}=f(\boldsymbol{x})F_m(\boldsymbol{x}), \qquad (22.14\text{a})$$

where

$$F_m(\boldsymbol{x})=\frac{1}{\sqrt{\mathcal{V}}}\sum_{\boldsymbol{k}}\bar{\bar{c}}_{m\boldsymbol{k}}e^{i\boldsymbol{k}\boldsymbol{x}},\qquad \boldsymbol{q}=\boldsymbol{k}'-\boldsymbol{k}.$$

As mentioned above, it is assumed here that $f(\boldsymbol{x})$ is a smooth function and the major contribution to (22.3) and (22.7) comes from the $\boldsymbol{q}$'s, which are small in comparison with the reciprocal lattice vectors. Under these conditions, the sum over $\boldsymbol{q}$ does not depend on the upper limit or, consequently, on $\boldsymbol{k}$, and the sum of products may be replaced by a product of sums.[*]

Under the same condition, we may proceed in the same way for sums of the type

$$\frac{1}{\sqrt{\mathscr{V}}}\sum_{k'k} k_\alpha^n e^{ik'x}\delta_{kk'}\bar{\bar{c}}_{mk} = \frac{1}{\sqrt{\mathscr{V}}}\left(-i\frac{\partial}{\partial x_\alpha}\right)^n \sum_{k} e^{ikx}\bar{\bar{c}}_{mk} = k_\alpha^n F_m(\boldsymbol{x}), \qquad (22.14\text{b})$$

where $\boldsymbol{k} = -i\nabla$. Similarly

$$\frac{1}{\sqrt{\mathscr{V}}}\sum_{kk'}(\boldsymbol{k}+\boldsymbol{k}')\mathscr{A}_{k'-k}e^{ikx}\bar{\bar{c}}_{mk} = \frac{1}{\sqrt{\mathscr{V}}}\sum_{qk}(2\boldsymbol{k}+\boldsymbol{q})\mathscr{A}_q e^{iqx}\bar{\bar{c}}_{mk}e^{ikx} = (\boldsymbol{k}\mathscr{A}+\mathscr{A}\boldsymbol{k})F_m(\boldsymbol{x}). \qquad (22.14\text{c})$$

The last term in (22.13), containing the interband terms, is transformed in the same manner to

$$\begin{aligned}\frac{1}{\sqrt{\mathscr{V}}}\sum_{k''k'k}\left(\hbar k_\alpha\delta_{k'k''}+\frac{e}{c}\mathscr{A}^\alpha_{k'-k''}\right)\left(\hbar k''_\beta\delta_{k''k}+\frac{e}{c}\mathscr{A}^\beta_{k''-k}\right)e^{ik'x}\bar{\bar{c}}_{mk} = \\ = \left(\hbar k_\alpha+\frac{e}{c}\mathscr{A}_\alpha\right)\left(\hbar k_\beta+\frac{e}{c}\mathscr{A}_\beta\right)F_m(\boldsymbol{x}).\end{aligned} \qquad (22.14\text{d})$$

Finally, the system of equations (22.11) in the $\boldsymbol{x}$-presentation becomes

$$\sum_m \mathscr{H}_{m'm}F_m(\boldsymbol{x},t) = i\hbar\frac{\partial}{\partial t}F_{m'}(\boldsymbol{x},t), \qquad (22.15)$$

where

$$\mathscr{H}_{m'm} = \frac{\hbar^2K^2}{2m}\delta_{mm'} + \frac{\hbar K}{m}\pi_{m'm} + \frac{\hbar^2}{m^2}\sum_s\frac{(K\pi_{m's})(K\pi_{sm})}{E_m-E_s} + U(\boldsymbol{x})\delta_{mm'} + \frac{1}{2}\mu_0 g_0(\sigma_{m'm}\boldsymbol{H}), \qquad (22.16)$$

$$\boldsymbol{K} = -i\nabla + \frac{e}{\hbar c}\mathscr{A}(\boldsymbol{x},t). \qquad (22.17)$$

In the presence of a magnetic field, when the components K_α do not commute, the product $K_\alpha K_\beta$ in (22.16) is conveniently represented in terms of the symmetrized and antisymmetrized products:

$$[K_\alpha K_\beta] = \frac{1}{2}(K_\alpha K_\beta + K_\beta K_\alpha), \qquad (22.18)$$

$$\{K_\alpha K_\beta\} = K_\alpha K_\beta - K_\beta K_\alpha = -i\frac{e}{\hbar c}\left(\frac{\partial A_\beta}{\partial x_\alpha}-\frac{\partial A_\alpha}{\partial x_\beta}\right) = -\frac{ie}{\hbar c}\sum_\gamma\delta_{\alpha\beta\gamma}H_\gamma, \qquad (22.19)$$

where $\delta_{\alpha\beta\gamma}$ is the unit antisymmetric tensor: $\delta_{\alpha\beta\gamma}=1$ if all three indices are different and in the natural order, i.e., a cyclic permutation of x, y, z; in the reverse order $\delta_{\alpha\beta\gamma}=-1$, and if there are equal indices $\delta_{\alpha\beta\gamma}=0$.

---

[*] The true potential $f(\boldsymbol{x})$ is thereby replaced by a smoothed potential

$$\bar{f}(\boldsymbol{x}) = \sum_{q<1/a_0} f_q e^{iqx},$$

which coincides "on the average" with $f(\boldsymbol{x})$ but may oscillate sharply over small distances, of the order of the lattice constant a_0. If f_q decreases slowly with q, as in the linear case $U(\boldsymbol{x}) = -e\mathscr{E}\boldsymbol{x}$ or $\mathscr{A}(\boldsymbol{x}) = {}^1/_2[\boldsymbol{Hx}]$ expressions (22.14) are valid only if $F(\boldsymbol{x})$ is sufficiently smooth. For example, in the latter case it is required that the magnetic length $L = (\hbar c/eH)^{1/2}$ be far greater than a_0: the error incurred by substituting $f(\boldsymbol{x})$ for $\bar{f}(\boldsymbol{x})$ is then of the order of $\exp\left(-L^2/a_0^2\right)$.

As a result, for points of zero slope, where $\boldsymbol{\pi}_{m'm}=0$, we may write (22.16) as

$$\mathcal{H}_{m'm}=\frac{\hbar^2}{2m}\sum_{\alpha\beta}[K_\alpha K_\beta]\left(\delta_{\alpha\beta}\delta_{m'm}+\frac{1}{m}\sum_{s}\frac{\pi^{\alpha}_{m's}\pi^{\beta}_{sm}+\pi^{\beta}_{m's}\pi^{\alpha}_{sm}}{E_m-E_s}\right)+$$
$$+U(\boldsymbol{x})\,\delta_{m'm}+\frac{1}{2}\mu_0\sum_{\gamma}H_\gamma\left(g_0\sigma^{\gamma}_{m'm}-\frac{i}{m}\sum_{s\alpha\beta}\frac{\pi^{\alpha}_{m's}\pi^{\beta}_{sm}}{E_m-E_s}\,\delta_{\alpha\beta\gamma}\right). \qquad (22.20)$$

In the absence of external fields, $F_m=\exp\{i[\boldsymbol{kx}-(Et/\hbar)]\}$, and equation (21.19), which determines the spectrum $E(\boldsymbol{k})$ near $\boldsymbol{k}_0$, follows from (22.15). If the band is degenerate owing to time reversal then, as indicated in §18, we can take as basis two linearly independent functions belonging to the energy E_m: $\psi_{1\boldsymbol{k}_0}$ and $\psi_{2\boldsymbol{k}_0}=KR\psi_{1\boldsymbol{k}_0}$, where K is the time reversal operation and R an operation which takes $\boldsymbol{k}_0$ into $-\boldsymbol{k}_0$. Correspondingly, the smooth functions $F_m(\boldsymbol{x})$ may be expressed as a single two-component function with components $F_1(\boldsymbol{x})$ and $F_2(\boldsymbol{x})$. If the degeneracy is not removed when $\boldsymbol{k}\neq 0$, it follows from (21.11) and (21.12) that the matrix $\mathcal{H}(\boldsymbol{k})$ is diagonal, and these functions are determined, according to (22.20), by the equation

$$\left(\sum_{\alpha\beta}\frac{\hbar^2}{2m_{\alpha\beta}}[K_\alpha K_\beta]+U(\boldsymbol{x},\,t)+\mu_0\left(\frac{1}{2}\,g_0\boldsymbol{\sigma}+\boldsymbol{L},\,\boldsymbol{H}\right)\right)F=i\hbar\frac{\partial F}{\partial t}. \qquad (22.21)$$

The tensor $m^{-1}_{\alpha\beta}$ is then given by (21.12), and by (22.20) the elements of the matrix $\boldsymbol{L}$ which describes the orbital contribution to the effective magnetic moment of the electron are

$$L^{\gamma}_{ij}=-\frac{i}{2m}\sum_{\alpha\beta s}\delta_{\alpha\beta\gamma}\frac{\pi^{\alpha}_{is}\pi^{\beta}_{sj}}{(E_m-E_s)} \qquad (i=1,\,2;\quad j=1,\,2). \qquad (22.22)$$

Thus equation (22.21) differs from the free-electron Schrödinger equation only in that the mass m is replaced by the effective mass, which is anitotropic in the general case, and $\boldsymbol{\sigma}$ by $\boldsymbol{\sigma}+\boldsymbol{L}$.

In the above equations, the spin-orbit coupling operator $\mathcal{H}_{\text{so}}$ is included everywhere in $\mathcal{H}_0$. If the spin-orbit splitting is small compared to the separation to the nearest band, the operator $\mathcal{H}_{\text{so}}$ may be regarded as a perturbation, as in §21, and we can take as basis the products of the coordinate functions and spin functions, $\psi_{m\mu\boldsymbol{k}_0}=\psi_{m\boldsymbol{k}_0}Y^{1/2}_{\mu}$ $(\mu=\pm 1/2)$. When this is done, the operator $\boldsymbol{\pi}$ should be replaced by $\boldsymbol{p}$, and $\mathcal{H}_{m'\boldsymbol{k},\,m\boldsymbol{k}}$ should incorporate $\mathcal{H}^0_{\text{so}}=$ $=\mathcal{H}_{\text{so}}(0)$, first order terms in $\mathcal{H}_{\text{so}}$ and cross terms of first order in $\mathcal{H}^0_{\text{so}}=$ $=\mathcal{H}_{\text{so}}(0)$ and $\mathcal{H}_2$ arising in second order perturbation theory. If it is necessary to introduce relativistic corrections to the effective mass and magnetic moment, we must also include terms from the third approximation: first order in $\mathcal{H}^0_{\text{so}}$ and second order in $\mathcal{H}_2$. According to (15.51) inclusion of these terms is equivalent to transforming (22.20) from the basis $\psi_{m\mu\boldsymbol{k}_0}$ to the basis

$$\bar{\psi}_{m\mu\boldsymbol{k}_0}=\psi_{m\mu\boldsymbol{k}_0}+\sum_{s\mu'}\frac{(\mathcal{H}_{\text{so}})_{s\mu',\,m\mu}\,\psi_{s\mu'\boldsymbol{k}_0}}{E_m-E_s}. \qquad (22.23)$$

The general equations given above are not specialized to any specific external fields. In practice, these fields are either created by impurities or defects in a semiconductor or applied from without. In connection with the latter, special interest attaches to the properties of a semiconductor in a constant magnetic field and crossed electric and magnetic fields.

We pause briefly to discuss the limits of applicability of effective mass theory. In the above derivation, we assumed that the fields $\boldsymbol{\mathscr{E}}(\boldsymbol{x})$ and $\boldsymbol{H}(\boldsymbol{x})$ are sufficiently smooth, so that we could ignore all the terms in (22.7) containing $\boldsymbol{q}$ values that "spill over" the boundary of the Brillouin zone; in other words, the significant values of $\boldsymbol{q}$ must be smaller than any of the reciprocal lattice vectors. In the case of an impurity center, this criterion is fulfilled only at sufficiently large distances from the impurity, while at distances of the order of the lattice constant the field of the center is no longer a Coulomb field and may vary rather sharply. The essential factor here is therefore the size of the region in which the electron is localized, which may be estimated from the binding energy. If this region occupies a sufficiently large volume compared with that of the primitive cell, effective mass theory is applicable. The ionization energy of such centers is small compared to the band gap, and they are therefore known as shallow centers. If the electron is localized within a primitive cell and the electron levels at the center lie near the middle of the band gap, effective mass theory is not applicable. We then speak of deep centers.

As mentioned previously, for an electron in an external magnetic field the smoothness of the field $\boldsymbol{H}$ is insufficient. The field should not be too strong, in order that the magnetic length, i.e., the localization region of the electron $L=(\hbar c/eH)^{1/2}$, substantially exceed the lattice constant. Similarly, the separation of the Landau levels $\hbar\omega_c = eH/m^*c$ must be small compared to the band gap E_g. In an electric field $\mathscr{E}$ the ratio $e\mathscr{E}a_0/E_g$ must be small.

In the case of crossed electric and magnetic fields, the condition imposed on the magnitude of the electric field is more rigid. An electron rotating about the z-axis, which is directed along the magnetic field, drifts in the x-direction (normal to $\boldsymbol{\mathscr{E}}$ and $\boldsymbol{H}$) with velocity $v_x = c\mathscr{E}/H$. The energy corresponding to this motion is $m^*v_x^2/2 = m^*c^2\mathscr{E}^2/2H^2$. A similar contribution comes from the average kinetic energy of the electron rotating while being accelerated in the electric field.

Effective mass theory is obviously applicable as long as this energy is small compared with the band gap, i.e., provided

$$\frac{\mathscr{E}}{H} \ll \left(\frac{E_g}{2m^*c^2}\right)^{1/2} = \frac{s}{c}, \qquad \text{where} \qquad s^2 = \frac{E_g}{2m^*}. \tag{22.24}$$

If the separation of two or more nearest bands is small compared to their separations to the other bands, we can substantially extend the limits of applicability of the theory, going over from the single-band to the two-band or multiband approximation. This approximation is applicable when the corresponding energies or frequencies of the external fields are small compared to the separation to the other, more distant bands, which may significantly exceed the band gap. Of course, the requirement that the external fields be smooth remains in force.

Multiband model. As mentioned in §21, in the multiband model, only interband terms corresponding to the interaction of the bands under discussion with the more distant ones are transformed out of equation (22.11) by (15.33), while the interaction of the nearest bands is accurately represented.

Here the energy is conveniently measured from an average $\bar{E}$ between the bands, and thus we must write $\bar{E}$ in the exponent in (22.12). The system of equations in the $\boldsymbol{k}$-representation will then differ from (22.12) only in that $\mathscr{H}_{m'\boldsymbol{k}',\,m\boldsymbol{k}}$ is replaced by $\mathscr{H}_{m'\boldsymbol{k}',\,m\boldsymbol{k}} + E_m\delta_{mm'}\delta_{\boldsymbol{k}\boldsymbol{k}'}$. According to the general equation

(15.46), in second order perturbation theory equation (22.13) is replaced by

$$\begin{aligned}\mathscr{H}_{m'\boldsymbol{k}',\,m\boldsymbol{k}} &= E_m\delta_{mm}\delta_{\boldsymbol{k}\boldsymbol{k}'} + (\mathscr{H}_1)_{m'\boldsymbol{k}',\,m\boldsymbol{k}} + (\mathscr{H}_2)_{m'\boldsymbol{k}',\,m\boldsymbol{k}} + \\ &+ \frac{1}{2}\sum_{s}\sum_{\boldsymbol{k}''}(\mathscr{H}_2)_{m'\boldsymbol{k}',\,s\boldsymbol{k}''}(\mathscr{H}_2)_{s\boldsymbol{k}'',\,m\boldsymbol{k}}\left(\frac{1}{E_m - E_s} + \frac{1}{E_{m'} - E_s}\right).\end{aligned} \tag{22.25}$$

Accordingly, in the $\boldsymbol{x}$-representation we have the following system of equations instead of (22.15) and (22.20):

$$\sum_{m}(\mathscr{H}_{m'm} + E_m\delta_{mm'})\,F_m = i\hbar\,\frac{\partial F_{m'}}{\partial t}, \tag{22.26}$$

where

$$\begin{aligned}\mathscr{H}_{m'm} = \frac{\hbar K}{m}\,\pi_{m'm} + \frac{\hbar^2}{2m}\sum_{\alpha\beta}[K_\alpha K_\beta]\times \\ \times\left\{\delta_{\alpha\beta}\delta_{m'm} + \frac{1}{2m}\sum_{s}(\pi^{\alpha}_{m's}\pi^{\beta}_{sm} + \pi^{\beta}_{m's}\pi^{\alpha}_{sm})\left(\frac{1}{E_m - E_s} + \frac{1}{E_{m'} - E_s}\right)\right\} + \\ + U(\boldsymbol{x})\,\delta_{mm'} + \frac{1}{2}\,\mu_0\sum_{\gamma}H_\gamma\left\{g_0\sigma^{\gamma}_{m'm} - \frac{i}{2m}\sum_{s\alpha\beta}\pi^{\alpha}_{m's}\pi^{\beta}_{sm}\delta_{\alpha\beta\gamma}\times\right. \\ \left.\times\left(\frac{1}{E_m - E_s} + \frac{1}{E_{m'} - E_s}\right)\right\}.\end{aligned} \tag{22.27}$$

As noted in §21, when determining the spectrum to the second order in $\boldsymbol{k}$ one may ignore the interband $\boldsymbol{k}^2$ terms in $\mathscr{H}_{m'm}$, since their contribution to the energy is of third order in $\boldsymbol{k}$. However, these terms may exert a significant influence on effects in external fields. Thus, these terms are the source of combined resonance, i.e., spin flip due to a varying electric field (see §33) rather than a magnetic field as in paramagnetic resonance. These terms result in the Pockels effect, i.e., birefringence in cubic crystals without inversion centers in a constant electric field, caused by the linear variation of the dielectric constant with the field.

Note that when the little group $G_{\boldsymbol{k}}$, contains inversion and its representations have a definite parity, the interband terms in the matrix $\mathscr{H}_{m'm}$ may be either linear in $\boldsymbol{k}$, if the representations belonging to bands m and m' have different parity, or only quadratic, if these representations have the same parity.

We briefly consider some specific problems solved by effective mass theory.

Shallow impurity centers. One of the most interesting cases in which the field is induced by impurities or defects is that of a shallow impurity center due to an impurity ion replacing one of the lattice atoms and possessing a positive or negative surplus charge Ze. Sufficiently far away from the ion, say more than one or two lattice constants, we may assume that the ion produces a Coulomb field:

$$U(\boldsymbol{x}) = -\frac{Ze^2}{\varkappa\,|\boldsymbol{x} - \boldsymbol{x}_0|}, \tag{22.28}$$

where $\boldsymbol{x}_0$ is the ion coordinate and $\varkappa$ the static dielectric constant. In anisotropic crystals, in which $\varkappa$ is a tensor, the quantity $\varkappa|\boldsymbol{x} - \boldsymbol{x}_0|$ must be replaced by

$$[\varkappa_{yy}\varkappa_{zz}(x - x_0)^2 + \varkappa_{zz}\varkappa_{xx}(y - y_0)^2 + \varkappa_{xx}\varkappa_{yy}(z - z_0)^2]^{1/2},$$

where the x-, y- and z-axes are directed along the principal crystal axes. The problem of a shallow impurity center in cubic crystals will be discussed in greater detail in §27.

Constant magnetic field. In this case $U(\boldsymbol{x})=0$; the gauge for the vector potential may be chosen in different ways; for example, the potential $\boldsymbol{\mathcal{A}}={}^1/_2[\boldsymbol{H}\boldsymbol{x}]$ satisfies rot $\boldsymbol{\mathcal{A}}=\boldsymbol{H}$. In practice, another gauge is often used: if the z-axis is in the direction of $\boldsymbol{H}$, it is convenient to take $\mathcal{A}_x=Hy$, $\mathcal{A}_y=\mathcal{A}_z=0$.

Constant electric field. In this case $U=-e\boldsymbol{\mathcal{E}}\boldsymbol{x}$ and $\boldsymbol{\mathcal{A}}=0$. In connection with various effects caused by a constant electric field $\boldsymbol{\mathcal{E}}$, such as electro-optical effects, tunnel effect, etc., it is convenient to transform (22.26) and (22.27) to the $\boldsymbol{k}$-representation. Then, since

$$\int x_\alpha e^{i\boldsymbol{k}\boldsymbol{x}}F(\boldsymbol{x})\,d\boldsymbol{x}=-i\frac{\partial}{\partial k_\alpha}\int e^{i\boldsymbol{k}\boldsymbol{x}}F(\boldsymbol{x})\,d\boldsymbol{x}=-i\frac{\partial}{\partial k_\alpha}F_m(\boldsymbol{k}),$$

we obtain a system of equations for the functions $F_m(\boldsymbol{k})$ which differs from (22.26), (22.27) in that the operators $\boldsymbol{k}$ are replaced by C-numbers $\boldsymbol{k}$ and the operator $U=-e\boldsymbol{\mathcal{E}}\boldsymbol{x}$ by

$$U=ie\boldsymbol{\mathcal{E}}\nabla_{\boldsymbol{k}}. \tag{22.29}$$

Thus, instead of a system of second order differential equations we obtain a system of first order equations, and the order is not increased even when the higher order $\boldsymbol{k}$ terms appearing in the following approximations are included.

Crossed electric and magnetic fields. To solve the motion of an electron in mutually perpendicular electric and magnetic fields, it is convenient to direct the y-axis along the field $\mathcal{E}$ and choose $\mathcal{A}_x=Hy$, $\mathcal{A}_y=\mathcal{A}_z=0$, $U=-e\mathcal{E}y$. Then the Hamiltonian (22.20) or (22.27) depends only on one variable, y. In this case the nature of the electron's motion depends essentially on the ratio of the electric and magnetic fields. In the simple two-band model (defined by equation (26.34) below), if $\mathcal{E}>(s/c)H$, where $s=(E_g/2m^*)^{1/2}$, the motion is infinite and as if in an electric field alone the electron moves in an open orbit. On the other hand, if $H>(c/s)\mathcal{E}$ the electron moves in a closed orbit, as if in a magnetic field alone. For more complicated bands, the value of the ratio of the fields corresponding to transition from one case to the other depends on the wave vector of the electron, i.e., for different electrons the transition from infinite to finite motion occurs in different fields. However, whenever the orbits in a magnetic field are closed they remain so in a sufficiently weak electric field. On the other hand, in a sufficiently strong electric field the orbits of all the electrons become open.

§23. REPRESENTATIONS OF SPACE GROUPS IN CUBIC AND HEXAGONAL CRYSTALS. DISTRIBUTION OF NORMAL MODES ACCORDING TO IRREDUCIBLE REPRESENTATIONS

In this section we shall use projective representations of the point groups to construct representations of the little group at different points of the

Brillouin zone for the space groups O_h^1, O_h^5, O_h^7, T_d^2, C_{6v}^4. In the following sections we shall determine the energy spectrum at a few representative points of the Brillouin zone.

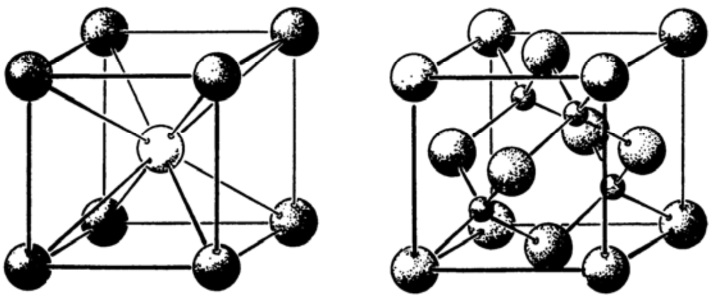
The group O_h^7 is the symmetry group of such "classical" semiconductors as Ge and Si, which have a diamond lattice: each primitive cell contains two like atoms (See Figure 24). The lattice is face-centered cubic Γ_c^f. Its basic vectors

$$\boldsymbol{a}_1 = \frac{a_0}{2}(110), \quad \boldsymbol{a}_2 = \frac{a_0}{2}(\bar{1}01), \quad \boldsymbol{a}_3 = \frac{a_0}{2}(011)$$

correspond to reciprocal lattice vectors

$$\boldsymbol{b}_1 = \frac{2\pi}{a_0}(11\bar{1}), \quad \boldsymbol{b}_2 = \frac{2\pi}{a_0}(1\bar{1}1), \quad \boldsymbol{b}_3 = \frac{2\pi}{a_0}(\bar{1}11), \tag{23.1}$$

where a_0 is the edge of the Bravais cube.



Halite Zinc blende

FIGURE 24. Halite and zinc blende lattices.

Place the zero point of the Bravais lattice at one of the atoms. Half of the elements of O_h^7, corresponding to the elements of the group T_d, do not have nonprimitive translations; the remaining elements, which are products of inversion i and the elements of T_d, occur together with a nonprimitive translation $\boldsymbol{\tau} = (a_0/4)(111)$. If we define the group O_h as in §3 by two generators c_4 and $s_6 = ic_3^2$, both occur in O_h^7 with the nonprimitive translation $\boldsymbol{\tau}$. The group O_h^7 also defines the symmetry of spinel crystals such as Fe_2MgO_4 and Fe_3O_4, so that the group-theoretic results obtained below are also applicable to spinels.

Semiconductors of the PbS, PbSe, PbTe type, whose structure and properties are now quite well known, have a halite lattice: face-centered Bravais lattice Γ_c^f, crystal class O_h. The primitive cell has one atom of each kind, one at a vertex of the Bravais cube and the other at the center (Figure 24). The space group therefore contains no nonprimitive translations. Such groups are called symmorphic. The symmorphic group for class O_h of the lattice Γ_c^f is O_h^5.

In this section we shall consider representations of the little group for the symmorphic group O_h^1 with the primitive cubic lattice Γ_c (the symmetry group of crystals of the CsCl type) and also representations of the little group for crystals with a zinc blende lattice (this is the type of most compounds of the A_3B_5 type, such as InSb, GaSb, GaAs). The crystal lattice is shown in Figure 24; it differs from the diamond lattice in that there are two unlike atoms in the primitive cell. Therefore, inversion is not an element of the symmetry group, which is the symmorphic group T_d^2. The Brillouin zones for the primitive Γ_c and face-centered lattices are shown in Figure 25.

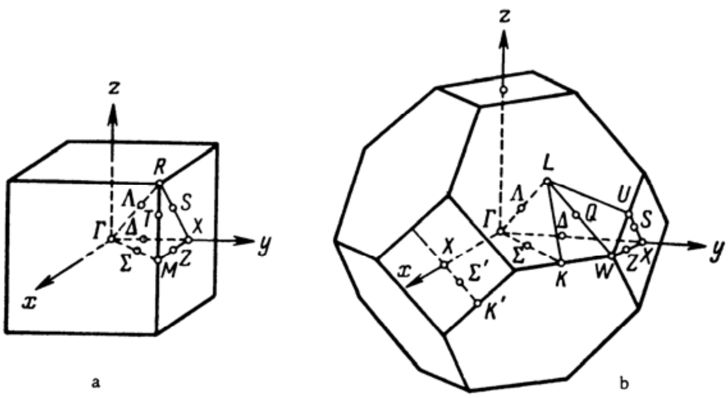


FIGURE 25. Brillouin zone for cubic lattices: a) primitive lattice (halite); b) face-centered (zinc blende, diamond).

Let us determine the symmetry of the points of Brillouin zone indicated in Figure 25. In so doing, we shall confine attention to the crystallographic point group F_k of each point k, bearing in mind, however, that the symmetry group G_k of k will contain different rotational elements with the same non-primitive translations as they have in the group G.

For crystals of class O_h, the symmetry group of the point Γ (the center of the Brillouin zone) is the crystallographic point group O_h.

The point Δ with coordinates $k_0\,(100)$ lies on a fourfold axis within the Brillouin zone; its symmetry group is C_{4v}. The Brillouin zone contains six points of this type: $\pm k_0\,(100)$, $\pm k_0\,(010)$, $\pm k_0\,(001)$, which go into one another upon inversion and rotation about the threefold axis.

The point Λ with coordinates $k_0\,(111)$ lies on a threefold axis; its symmetry group is C_{3v}. The Brillouin zone contains eight points of this type: $\pm k_0\,(111)$, $\pm k_0\,(11\bar{1})$, $\pm k_0\,(1\bar{1}1)$, $\pm k_0\,(\bar{1}11)$, which go into one another upon inversion and rotation about the fourfold axis.

The point Σ with coordinates $k_0\,(110)$ lies on the twofold axis joining the midpoints of opposite edges of the cube; its symmetry group is C_{2v}. There are twelve points of this type: $\pm k_0\,(110)$, $\pm k_0\,(101)$, $\pm k_0\,(011)$, $\pm k_0\,(1\bar{1}0)$, $\pm k_0\,(10\bar{1})$, $\pm k_0\,(01\bar{1})$.

In the lattice Γ_c, the point R with coordinates $k = (\pi/a_0)\,(111)$, where a_0 is the edge of the cube, lies at a vertex of the cube; it is the limiting position of the point Λ in Γ_c as it approaches the boundary of the Brillouin zone.

Since the vertices of the cube differ by integral multiples of reciprocal lattice vectors, all vectors $\boldsymbol{k}$ of the type R are equivalent. Since transformations in the cubic group bring these vertices into coincidence with one another, the symmetry group of R, like that of Γ, is the entire group O_h.

The point X with coordinates $(\pi/a_0)(100)$ lies at the center of a face of the cube; it is the limiting position of the point Δ on the surface of the Brillouin zone. Since points on opposite faces of the cube are equivalent, the symmetry group of X is obtained by adding inversion i to the elements of C_{4v}, and the result is the group $D_{4h} = D_{2d} \times C_i$. There are three inequivalent points of this type in the Brillouin zone.

The point M with coordinates $(\pi/a_0)(110)$ lies at the center of an edge of the cube. Since points at the midpoints of parallel edges of the cube differ by integral multiples of reciprocal lattice vectors, the corresponding vectors are equivalent. Thus the symmetry group of M must contain, in addition to the symmetry elements of Σ, a fourfold axis and inversion i, and so it is D_{4h}. The Brillouin zone contains three inequivalent points M.

In the lattice Γ_c^f, the interior points Γ, Δ, Σ, Λ of the Brillouin zone have the same symmetry as in Γ_c. The point X with coordinates $(2\pi/a_0)(100)$ in Γ_c^f, as in Γ_c, has symmetry group D_{4h}.

The point L with coordinates $(\pi/a_0)(111)$ lies at the center of a hexagonal face of the Brillouin zone, at its intersection with the threefold axis. Its symmetry group, in contrast to the point R in Γ_c, is $D_{3d} = C_{3v} \times C_i$. There are four inequivalent points of this type in the Brillouin zone.

The symmetry group of the point $K = (3\pi/2a_0)(110)$, like that of Σ, is C_{2v}, and therefore no additional symmetry elements appear in Γ_c^f when the point Σ crosses the boundary of the Brillouin zone.

The crystallographic point groups of the above points of the Brillouin zone, for lattices Γ_c and Γ_c^f of class O_h, are given in the following table.

Group	O_h	D_{4h}	C_{3v}	C_{2v}	C_{4v}	D_{3d}
Point	Γ, R	X, M	Λ	Σ, K	Δ	L
Number of points in star	1	3	8	12	6	4

For crystals of class T_d with lattice Γ_c^f, the symmetry group of the center of the Brillouin zone (the point Γ) is T_d.

The symmetry group of the point Δ is C_{2v}. As in the case of O_h, the Brillouin zone contains six points of this type, which go into one another under the transformations of T_d. The point Λ, as in O_h, has symmetry group C_{3v}, and the number of such points is four. In contrast to the group O_h, the points Λ and $-\Lambda$ are not in one star.

The symmetry group of the point Σ is C_s.

The point X has, in addition to the elements of the group C_{2v} of Δ, transformations s_4 and s_4^3 taking $\boldsymbol{k}$ into $-\boldsymbol{k}$; hence its symmetry group is D_{2d}; the number of points of this type is three.

The symmetry group of the point L, like that of Λ, is C_{3v}. The symmetry group of the point K, like that of Σ, is C_s.

The symmetry of the above points of the Brillouin zone for crystal class T_d and lattice Γ_c^f is given in the following table.

Group	T_d	D_{2d}	C_{3v}	C_{2v}	C_s
Point	Γ	X	L, Λ	Δ	Σ, K
Number of points in star	1	3	4	6	12

We now consider the representations of the little group and the possible band degeneracies at the above points in crystals of class O_h.

Class O_h

Point Γ. At this point, the Bloch functions coincide with the modulating Bloch factors $u(\mathbf{x})$; the representations $\mathscr{D}^{\mathbf{k}}(g)$ and $\mathscr{D}^{\mathbf{k}}(r)$ coincide and are representations of the point group O_h for space groups O_h^1, O_h^5, O_h^7.

The group $O_h = T_d \times C_i$ has ten single-valued representations $A_1^\pm$, $A_2^\pm$, $E^\pm$, $F_2^\pm$, $F_1^\pm$ and six double-valued representations $E_1'^\pm$, $E_2'^\pm$, $G^\pm$. The literature on band theory, following Boukaert, Smoluchowski, Wigner /16.1/, designates representations of little groups by the symbol for the corresponding point in the Brillouin zone.

The following table shows the connection between the representations of the point group O_h and the representations of the space group at Γ. All the representations at Γ belong to case (a_1), since the representations of the space group have real characters; they coincide with the characters of the representations of the point group O_h.

Representation of $O_h = T_d \times C_i$	A_1^+	A_2^+	E^+	F_1^+	F_2^+	A_1^-	A_2^-	E^-	F_1^-	F_2^-	$E_1'^-$	$E_2'^-$	$G'^\pm$	$E_1'^+$	$E_2'^+$
Representation at Γ	Γ_1	Γ_2	Γ_{12}	Γ_{15}'	Γ_{25}'	Γ_2'	Γ_1'	Γ_{12}'	Γ_{25}	Γ_{15}	Γ_6^-	Γ_7^-	$\Gamma_8^\pm$	Γ_7^+	Γ_6^+

At the point Γ we may have onefold, twofold and threefold band degeneracy if spin is not included, twofold and fourfold degeneracy counting spin.

In order to describe the band splitting at Γ due to spin-orbit coupling, we must expand the product $\Gamma_i \times \mathscr{D}_{1/2}$ in terms of double-valued irreducible representations of O_h, as shown in Table 16.1 (p. 148). The table shows that the representations Γ_1, Γ_1', Γ_2, Γ_2', Γ_{12}, Γ_{12}' are not split by spin-orbit coupling, and inclusion of spin for these representations only doubles the band degeneracy.

As noted in §11, the representation $\mathscr{D}_1$ of the rotation group with the basis p-functions x, y, z is also irreducible in the cubic group, becoming the representation Γ_{15}^-. The direct product $\mathscr{D}_1 \times \mathscr{D}_{1/2} = \mathscr{D}_{1/2} + \mathscr{D}_{3/2}$ contains the four- and two-dimensional representations $\mathscr{D}_{3/2}$ and $\mathscr{D}_{1/2}$. These representations are also irreducible in the cubic group, becoming respectively the representations Γ_8^- and Γ_6^-, in accordance with the decomposition $\Gamma_{15}^- \times \mathscr{D}_{1/2} = \Gamma_8^- + \Gamma_6^-$. We thus take $\psi_m^{3/2}$ $(m = \pm 1/2, \pm 3/2)$ as basis functions for the representation Γ_8^- in the Luttinger-Kohn representation /17.3/:

$$\psi_{3/2}^{3/2} = \frac{1}{\sqrt{2}}(X + iY)\,\alpha, \qquad \psi_{1/2}^{3/2} = \frac{i}{\sqrt{6}}[(X + iY)\,\beta - 2Z\alpha],$$
$$\psi_{-1/2}^{3/2} = \frac{1}{\sqrt{6}}[(X - iY)\,\alpha + 2Z\beta], \quad \psi_{-3/2}^{3/2} = \frac{i}{\sqrt{2}}(X - iY)\,\beta, \tag{23.2}$$

and $\psi_m^{1/2}$ $(m = \pm 1/2)$ as basis functions for Γ_6^-:

$$\psi_{1/2}^{1/2} = \frac{1}{\sqrt{3}}[(X + iY)\,\beta + Z\alpha], \quad \psi_{-1/2}^{1/2} = \frac{i}{\sqrt{3}}[-(X - iY)\,\alpha + Z\beta]. \tag{23.3}$$

Here X, Y, Z are the Bloch functions u_n for $k = 0$, which transform under operations of the cubic group as x, y, z, respectively.

All the above results for the point Γ are valid for all space groups of class O_h.

Point Δ. According to (12.26), the representations of the little group $\mathscr{D}_\nu^\Delta(g)$ at Δ are

$$\mathscr{D}_\nu^\Delta(g) = e^{-\boldsymbol{k}\beta}\mathscr{D}_\nu^\Delta(r) = e^{-ik_0\beta_x}\,\mathscr{D}_\nu^\Delta(r), \tag{23.4}$$

where $\mathscr{D}_\nu^\Delta(r)$ is a projective representation of the crystallographic point group of Δ, and $\boldsymbol{\beta} = \boldsymbol{a} + \boldsymbol{\tau}$, where $\boldsymbol{a}$ and $\boldsymbol{\tau}$ are respectively primitive and nonprimitive translations corresponding to the rotational element $r \in C_{4v}$, $g = (r|\beta)$. Since Δ is inside the Brillouin zone, the $\mathscr{D}_\nu^\Delta(r)$ are matrices of vector representations of the point group C_{4v}. In C_{4v} there are five single-valued representations A_1, A_2, B_1, B_2, E and two double-valued representations E_1' and E_2'. Of these, the representations A_1, A_2, B_1, B_2 are one-dimensional, while E and E_1', E_2' are two-dimensional. By (23.4), each of these representations of C_{4v} corresponds to a representation of the little group, of the same dimension; they are labeled by the letter Δ, as corresponding to the point Δ. Their relation to the representations of the group C_{4v} is given in the following table.

Representation of C_{4v}	A_1	A_2	B_1	B_2	E	E_1'	E_2'
Representation at Δ	Δ_1	Δ_1'	Δ_2	Δ_2'	Δ_5	Δ_6	Δ_7

Although the matrices $\mathscr{D}_\nu^\Delta(r)$ are the same for all space groups of class O_h, the representation matrices of the little group $\mathscr{D}_\nu^\Delta(g)$ are different for different space groups, in particular for the groups O_h^1, O_h^5, O_h^7.

Thus, in the groups O_h^1 and O_h^5, which do not contain nonprimitive translations, the matrices $\mathscr{D}_\nu^\Delta$ coincide (for elements which do not contain nonprimitive translations) with the representation matrices of C_{4v} and do not depend on the position of Δ on the fourfold axis.

In O_h^7 half the elements of the point group C_{4v}, namely c_4, c_4^3, σ_v, $c_4^2\sigma_v$, occur in combination with a nonprimitive translation $\boldsymbol{\tau}$, and therefore the matrices $\mathscr{D}_\nu^\Delta(g)$ and $\mathscr{D}_\nu^\Delta(r)$ for these elements differ by a factor $\lambda = e^{-ik_0\tau_x} = e^{-i(\pi/2)p}$, where $p = \frac{k_0}{2\pi/a_0}$ is the ratio of the length of $\boldsymbol{k}_0$ to the distance to the Brillouin zone boundary: $p=0$ at Γ and $p=1$ at X. Therefore, in the group O_h^7 the matrices $\mathcal{D}_\nu^\Delta(g)$ and the characters of the representations of the little group differ from the corresponding representations of C_{4v} and depend on the position of the point on the fourfold axis.

Equation (18.32) implies that both the single-valued and double-valued representations of a space group with star Δ belong to case (a_2), i.e., there is a linear relation between the functions ψ_{ik} and $RK\psi_{ik}$. As the element R which takes $\boldsymbol{k}$ into $-\boldsymbol{k}$ we may take inversion.

When Δ reaches the surface of the Brillouin zone, it becomes the point X, for which there is an essential difference between the representations of groups O_h^1, O_h^5 and O_h^7.

Point X. According to (23.4) the representations of the little group $\mathcal{D}_\nu^X(g)$ at X are

$$\begin{aligned}\mathcal{D}_\nu^X(g) &= e^{-i(\pi/a_0)\beta_x}\mathcal{D}_\nu^X(r) \quad (\text{for } O_h^1),\\ \mathcal{D}_\nu^X(g) &= e^{-i(2\pi/a_0)\beta_x}\mathcal{D}_\nu^X(r) \quad (\text{for } O_h^5,\ O_h^7),\end{aligned} \tag{23.5}$$

where $\mathcal{D}_\nu^X(r)$ runs through the projective representations of the group D_{4h}. Moreover, in the case of O_h^1 and O_h^5, which do not contain nonprimitive translations, it follows from (12.29) that the factor system is trivial and all projective representations belong to class K_0, so they are the usual representations of D_{4h}.

The group $D_{4h}=D_{2d}\times C_i$ has ten single-valued representations: five even and five odd (with respect to inversion), and four double-valued two-dimensional representations. By (23.5), each of these representations for O_h^1 and O_h^5 corresponds to a representation of the little group $\mathcal{D}^X$, denoted by the letter X_i according to the following table.

Representation of $D_{4h}=D_{2d}\times C_i$	$A_1^\pm$	$A_2^\pm$	$B_1^\pm$	$B_2^\pm$	$E^\pm$	$E_1'^\pm$	$E_2'^\pm$
Representation at X	$X_1^\pm$	$X_2^\pm$	$X_3^\pm$	$X_4^\pm$	$X_5^\pm$	$X_6^\pm$	$X_7^\pm$
Representation at M	$M_1^\pm$	$M_2^\pm$	$M_3^\pm$	$M_4^\pm$	$M_5^\pm$	$M_6^\pm$	$M_7^\pm$

Thus, the band is twofold degenerate counting spin. Since the little group contains inversion, all the representations belong to case 1, and since in the absence of nonprimitive translations Herring's test (18.32) reduces to Schur's test (18.27) for the group D_{4h}, all the representations at X for O_h^1 and O_h^5 belong to case (a_1).

We now consider the representations at X for the group O_h^7, which contains nonprimitive translations.

The factor system at X for O_h^7 is given in Table 23.1. As shown in §14, in order to determine to which class of the group D_{4h} this factor system belongs, we must find the quotient of its coefficients for pairs of commuting elements:

$$\alpha=\frac{\omega(u_2,c_4^2)}{\omega(c_4^2,u_2)}=1, \quad \beta=\frac{\omega(c_4,i)}{\omega(i,c_4)}=-1, \quad \gamma=\frac{\omega(u_2,i)}{\omega(i,u_2)}=1.$$

TABLE 23.1

r_1 \ r_2	$(e\|0)$	$(c_{4x}\|\tau)$	$(c_{2x}\|0)$	$(c_{4x}^3\|\tau)$	$(c_{2yz}\|\tau)$	$(c_{2z}\|0)$	$(c_{2y\bar{z}}\|\tau)$	$(c_{2y}\|0)$	$(i\|\tau)$	$(s_{4x}^3\|0)$	$(\sigma_x\|\tau)$	$(s_{4x}\|0)$	$(\sigma_{yz}\|0)$	$(\sigma_z\|\tau)$	$(\sigma_{y\bar{z}}\|0)$	$(\sigma_y\|\tau)$
$(e\|0)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$(c_{4x}\|\tau)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$(c_{2x}\|0)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$(c_{4x}^3\|\tau)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$(c_{2yz}\|\tau)$	1	−1	1	−1	−1	1	−1	1	−1	1	−1	1	1	−1	1	−1
$(c_{2z}\|0)$	1	−1	1	−1	−1	1	−1	1	−1	1	−1	1	1	−1	1	−1
$(c_{2y\bar{z}}\|\tau)$	1	−1	1	−1	−1	1	−1	1	−1	1	−1	1	1	−1	1	−1
$(c_{2y}\|0)$	1	−1	1	−1	−1	1	−1	1	−1	1	−1	1	1	−1	1	−1
$(i\|\tau)$	1	−1	1	−1	−1	1	−1	1	−1	1	−1	1	1	−1	1	−1
$(s_{4x}^3\|0)$	1	−1	1	−1	−1	1	−1	1	−1	1	−1	1	1	−1	1	−1
$(\sigma_x\|\tau)$	1	−1	1	−1	−1	1	−1	1	−1	1	−1	1	1	−1	1	−1
$(s_{4x}\|0)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$(\sigma_{yz}\|0)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$(\sigma_z\|\tau)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$(\sigma_{y\bar{z}}\|0)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$(\sigma_y\|\tau)$	1	−1	1	−1	−1	1	−1	1	−1	1	−1	1	1	−1	1	−1
u	1	i	−1	$-i$	i	−1	$-i$	1	i	−1	$-i$	1	1	i	−1	$-i$

Consequently, the factor system belongs to class K_3, in which there are four two-dimensional projective representations. Thus, in the group O_h^7 there is a twofold essential degeneracy.

In order to bring the factor system into the standard form (14.42) for its class, for which the matrices of the projective representations are given in §14, we must calculate the function u (14.43). For the factor system under discussion, we have

$$\omega(b, b) = \omega(u_2, u_2) = -1, \quad \omega(c, c) = \omega(i, i) = -1,$$
$$\omega_{a4} = \omega(c_4, c_4)\,\omega(c_4, c_4^2)\,\omega(c_4, c_4^3) = 1,$$

and the quantity α' (see (14.24)) is -1, i.e., $m = 2$ in equation (14.25). Since in the case at hand $\omega(a^k, b^p c^q) = 1$, we obtain

$$u(a^k b^p c^q) = \frac{i^{p+q+k}}{\omega(b^p, c^q)}.$$

The corresponding values of u are given in Table 23.1.

By (23.5) the representation matrices of the little group at the point X are

$$\mathscr{D}^X = e^{-i(2\pi/a_0)\beta_x}\,\mathscr{D}'(r)\,u(r), \tag{23.6}$$

where $\mathscr{D}'(r)$ runs through the matrices of the projective representations, which, if $r = a^k b^p c^q$, are the corresponding products of powers of matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$:

$$\mathscr{D}'(a^k b^p c^q) = \boldsymbol{A}^k \boldsymbol{B}^p \boldsymbol{C}^q. \tag{23.7}$$

The matrices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ are given in Table 14.2 (p. 113).

According to §16, the double-valued representations of a point group are projective representations of the point group. Thus the direct product of projective representations $\mathscr{D}(r)\times\mathscr{D}_{1/2}$ is a projective representation of the group D_{4h}. In §16 it was shown that the spinor representations of D_{4h} belong to class K_1. Therefore, the double-valued representations at X belong to class $K_1K_3=K_6$, in which there is one four-dimensional representation. Thus, with allowance for spin we have a fourfold degeneracy at X in O_h^7.

As mentioned in §16, the factor system corresponding to the usual spinor representations of D_{4h} differs from our standard factor system (14.42) for class K_1, but may be reduced to it by using equation (13.3) with the function u of (16.29). Thus formulas (23.6) and (23.7) also define a spinor representation of the little group, provided u is the product u_1u_2, where u_1 is the function defined above and u_2 is given in (16.29).

The characters of the representations of the little group at X are given in Table 23.2.

TABLE 23.2. Characters of representations of the little group at $\boldsymbol{X}$ for Ge-type lattice

Elements		$X_1\ (K_3)$	$X_2\ (K_3)$	$X_3\ (K_3)$	$X_4\ (K_3)$	$X_5\ (K_6)$
$(e\mid 0)$	$(e\mid 0)$	2	2	2	2	4
$(c_{4x}\mid\boldsymbol{\tau})$	$(c_4\mid\boldsymbol{\tau})$	0	0	0	0	0
$(c_{2x}\mid 0)$	$(c_4^2\mid 0)$	−2	2	−2	2	0
$(c_{4x}^3\mid\boldsymbol{\tau})$	$(c_4^3\mid\boldsymbol{\tau})$	0	0	0	0	0
$(c_{2yz}\mid\boldsymbol{\tau})$	$(u_2\mid\boldsymbol{\tau})$	2	0	−2	0	0
$(c_{2z}\mid 0)$	$(c_4u_2\mid 0)$	0	0	0	0	0
$(c_{2y\bar{z}}\mid\boldsymbol{\tau})$	$(c_4^2u_2\mid\boldsymbol{\tau})$	−2	0	2	0	0
$(c_{2y}\mid 0)$	$(c_4^3u_2\mid 0)$	0	0	0	0	0
$(i\mid\boldsymbol{\tau})$	$(i\mid\boldsymbol{\tau})$	0	0	0	0	0
$(s_{4x}^3\mid 0)$	$(c_4i\mid 0)$	0	0	0	0	0
$(\sigma_x\mid\boldsymbol{\tau})$	$(c_4^2i\mid\boldsymbol{\tau})$	0	0	0	0	0
$(s_{4x}\mid 0)$	$(c_4^3i\mid 0)$	0	0	0	0	0
$(\sigma_{yz}\mid 0)$	$(u_2i\mid 0)$	0	2	0	−2	0
$(\sigma_z\mid\boldsymbol{\tau})$	$(c_4u_2i\mid\boldsymbol{\tau})$	0	0	0	0	0
$(\sigma_{y\bar{z}}\mid 0)$	$(c_4^2u_2i\mid 0)$	0	2	0	−2	0
$(\sigma_y\mid\boldsymbol{\tau})$	$(c_4^3u_2i\mid\boldsymbol{\tau})$	0	0	0	0	0

Observe the essential difference between the representations X_i ($i=1,2,3,4$) and X_5. In both cases the matrices $\mathscr{D}'(r)$ form projective representations of the point group D_{4h}, but the representations X_i ($i=1,\ldots,4$), considered for all elements of the little group including primitive translations, are vector representations of the little group, whereas X_5 is a spinor representation (i.e., projective representation of class K_1).

Since $\boldsymbol{k}\equiv-\boldsymbol{k}$ at the point X, the summation in Herring's test (18.32) should extend over all the elements of D_{4h}, taking into account that the element g^2 may contain a primitive translation or (in case of the representation X_5) the element Q. All the representations X_i belong to case (a$_1$).

Point Λ. The representations of the little group at the point Λ are given by

$$\mathcal{D}^{\Lambda}(g) = e^{-ik_0(\beta_x+\beta_y+\beta_z)}\,\mathcal{D}(r), \tag{23.8}$$

where $\mathcal{D}(r)$ runs through the usual representations of the group C_{3v}. Since C_{3v} is a subgroup of T_d, its elements do not occur in combination with nonprimitive translations, and so for elements which do not contain primitive translations the representations $\mathcal{D}^{\Lambda}$ are identical for the groups O_h^1, O_h^5, and O_h^7, not depending on the position of the point Λ on the threefold axis and coinciding with single-valued or double-valued representations of the group C_{3v}.

In C_{3v} there are three single-valued representations: two one-dimensional A_1, A_2 and one two-dimensional E, and double-valued representations: two one-dimensional E_1', E_2' and one two-dimensional E_3'. The corresponding representations of the little group are denoted by Λ_1, Λ_2, Λ_3 and Λ_4, Λ_5, Λ_6, respectively. It follows from Herring's test (18.32) that the representations Λ_1, Λ_2, Λ_3 and Λ_6 belong to case (a_2), while the one-dimensional double-valued representations Λ_4 and Λ_5 with complex conjugate characters belong to case (b_2). In the latter case, combination of complete representations of the space group results in combination of the representations Λ_4 and Λ_5 of the little group.

Point L. The representations of the little group at the point L are

$$\mathcal{D}^{L}(g) = e^{-i(\pi/a_0)(\beta_x+\beta_y+\beta_z)}\,\mathcal{D}(r), \tag{23.9}$$

where $\mathcal{D}(r)$ runs through the projective representations of the group $D_{3d} = C_{3v} \times C_i$. For the group O_h^5, the representations $\mathcal{D}(r)$ coincide with the usual representations of D_{3d}. For each of the six single-valued representations $A_1^{\pm}$, $A_2^{\pm}$ (one-dimensional), $E^{\pm}$ (two-dimensional) and the six double-valued representations $A_1'^{\pm}$, $A_2'^{\pm}$ (one-dimensional), $E'^{\pm}$ (two-dimensional) of the group D_{3d}, we have by (23.9) a representation of the little group $\mathcal{D}^L$ as indicated in the following table.

Representation of $D_{3d} = C_{3v} \times C_i$	A_1^+	A_1^-	A_2^+	A_2^-	E^+	E^-	$A_1'^+$	$A_1'^-$	$A_2'^+$	$A_2'^-$	E'^+	E'^-
Representation at L	L_1	L_2'	L_2	L_1'	L_3	L_3'	L_4	L_5'	L_5	L_4'	L_6	L_6'

In the group O_h^7, the factor system for L is nonidentity. The group D_{3d} is isomorphic to D_6 and so, to determine to which class the factor system belongs, we must find the ratio $\omega(i, u_2)/\omega(u_2, i)$, which is in this case unity. Therefore, the factor system and all the representations of O_h^7 at L belong to class K_0 and are p-equivalent to vector representations of D_{3d}.

Using (12.29), we easily evaluate the factor system at L and, using (14.33), (14.34), find a function u which reduces the factor system to the unit factor system.

The representations of the little group at the point L are

$$\mathcal{D}^{L}(g) = e^{-(\pi/a_0)(\beta_x+\beta_y+\beta_z)}u(r)\,\mathcal{D}'(r),$$

where $\mathscr{D}'(r)$ runs through the matrices of the vector representations of the point group D_{3d}. It is easily proved directly that $\exp\{-i(\pi/a_0)(\tau_x+\tau_y+\tau_z)\}u(r)= =1$ for all elements of D_{3d}; therefore the representations $\mathscr{D}^L(g)$ of the little group at L for elements g not containing primitive translations simply coincide with the matrices of a representation of the point group D_{3d}, as is the case at points Λ within the Brillouin zone. This could have been expected, since in the absence of essential degeneracy at L the representation $\mathscr{D}^\Lambda$ goes continuously into $\mathscr{D}^L$ when Λ approaches L.

Herring's test shows that the representations $L_4^\pm$ and $L_5^\pm$ belong to case (b_1), and time reversal causes the representations L_4^+, L_5^+ and L_4^-, L_5^- to combine. The remaining representations belong to case (a_1).

Point Σ. The representations of the little group $\mathscr{D}^\Sigma$ at Σ are

$$\mathscr{D}^\Sigma(g)=e^{-ik_0(\beta_x+\beta_y)}\,\mathscr{D}(r), \tag{23.10}$$

where $\mathscr{D}(r)$ runs through the representations of the group C_{2v}.

For O_h^1 and O_h^5, the representations $\mathscr{D}^\Sigma$ for elements not containing primitive translations do not depend on the position of Σ on the twofold axis and coincide with the representations of the group C_{2v}. For the elements c_2 and $c_2\sigma_v$ of the group O_h^7, we have a factor $e^{-i(3\pi/4)p}$ in (23.10), where $p=k_0\Big/\frac{3}{2}\frac{\pi}{a_0}$: $p=0$ for the point Γ and $p=1$ for the point K. In the group O_h^7, therefore, the representations $\mathscr{D}^\Sigma$ depend on the position of Σ on the twofold axis.

The group C_{2v} has four one-dimensional single-valued representations A_1, A_2, B_1, B_2 and one two-dimensional double-valued representation E', which lead via (23.10) to representations of the little group of the same respective dimensions, Σ_1, Σ_2, Σ_3, Σ_4 and Σ_5.

According to Herring's test, both the single-valued and the double-valued representations at Σ belong to case (a_2); this indicates the presence of a linear relation between functions ψ and $(i|\tau)K\psi$ belonging to the same point of the star of Σ.

As noted above, as Σ approaches the boundary of the Brillouin zone in the lattice Γ_c^f, its symmetry is not increased, and so all the results for Σ in the groups O_h^7 and O_h^5 remain valid for the point $K=\frac{3}{2}\frac{\pi}{a_0}(110)$.

Point M. The representations $\mathscr{D}^M$ at M are

$$\mathscr{D}^M(g)=e^{-i(\pi/a_0)(\beta_x+\beta_y)}\,\mathscr{D}(r). \tag{23.11}$$

Since the group O_h^1 contains no nonprimitive translations, $\mathscr{D}^M(r)$ coincides with $\mathscr{D}(r)$, which (at the point M) is a representation of the group D_{4h}. Thus all results obtained above for the point X in the group O_h^5 remain valid for the point M.

A similar procedure yields the representations of the little group at any other point of the Brillouin zone.

Note that, in accordance with the remark made in §18, the presence of inversion in the group O_h at any point of the Brillouin zone implies (when allowance is made for spin) at least twofold degeneracy of the energy spectrum.

Group T_d^2

We now consider the representations of the little group at a few points in $\boldsymbol{k}$-space for the group T_d^2. Since T_d^2 contains no nonprimitive translations, all the representations of the little group coincide (for elements not containing primitive translations) with the corresponding representations of the crystallographic point group.

At the point Γ, the representations of the space group coincide with the representations of the little group, which are representations of the group T_d. The group T_d has five single-valued representations A_1, A_2, E, F_2, F_1, denoted by Γ_1, Γ_2, Γ_{12}, Γ_{15}, Γ_{25} respectively, and also three spinor representations E_1', E_2', G', denoted by Γ_6, Γ_7, Γ_8.

All the representations at Γ belong to case (a_1).

At the point Δ the representations of the little group coincide with the representations of the group C_{2v}. This group is abelian, and so it has only one-dimensional single-valued representations A_1, A_2, B_1, B_2, denoted by Δ_1, Δ_2, Δ_3, Δ_4, respectively.

When spin is included, the band is always twofold degenerate at Δ, since C_{2v} has a two-dimensional double-valued representation E', denoted by Δ_5, such that $\Delta_i \times \mathscr{D}_{1/2} = \Delta_5$ $(i = 1, 2, 3, 4)$. All the representations at Δ belong to case (b_3): they are complex and inequivalent. In this case time reversal brings about coincidence of energy at the points $\boldsymbol{k}$ and $-\boldsymbol{k}$, i.e., $E(\boldsymbol{k}) = E(-\boldsymbol{k})$.

At the point X, the representations X_i of the little group, for elements not containing primitive translations, coincide with the representations of D_{2d}. The latter has five single-valued representations A_1, A_2, B_1, B_2, E and two two-dimensional spinor representations E_1', E_2'. The single-valued representations at X are denoted by X_i $(i = 1, 2, 3, 4, 5)$, and the spinor representations by X_6, X_7.

Representation of D_{2d}	A_1	A_2	B_1	B_2	E	E_1'	E_2'
Representation at X	X_1	X_2	X_3	X_4	X_5	X_6	X_7

For the point X, we have $\boldsymbol{k} \equiv -\boldsymbol{k}$, and all representations belong to case (a_1).

Now for the points L and Λ. Since these points have the same symmetry in the group T_d^2, the representations of the corresponding little groups are the same, coinciding with the representations of the little group at Λ in the groups O_h^5 and O_h^7 (see above). The difference between the points Λ and L is revealed only by time reversal.

The representations at Λ belong to case (b_3), since the little group does not contain elements which transform $\boldsymbol{k}$ into $-\boldsymbol{k}$. At the point L we have $\boldsymbol{k} \equiv -\boldsymbol{k}$. The representations L_1, L_2, L_3, L_6 with real characters belong to case (a_1); the two double-valued representations L_4 and L_5 (as at the point Λ in the group O_h) with complex characters belong to case (b_2) so that these two representations of the little group must be combined.

Thus, in the group T_d^2 the band is always twofold degenerate at L when spin is included, whereas at Λ the band may split completely even when spin is introduced.

In the group T_d^2, the points Σ and K have symmetry group C_s. This group has two one-dimensional single-valued representations A^+ and A^- and two one-dimensional double-valued representations A'^+ and A'^-, to which correspond four representations of the little group, $\Sigma_1, \Sigma_2, \Sigma_3$, Σ_4 (or K_1, K_2, K_3, K_4 for the point K). At these points, therefore, there may be only onefold de-degeneracy. All the representations for the points Σ and K belong to case (b_3).

Compatibility conditions

We now consider compatibility conditions for representations in crystals of the cubic group; these conditions express the relationships between representations at different points of the Brillouin zone. The compatibility conditions for representations at two interior points of the Brillouin zone are the same for all space groups in one crystal class. If one of the points is on the boundary of the Brillouin zone, the compatibility conditions may differ for different space groups.

The compatibility conditions for representations at the points Γ and Δ for groups of class O_h are

$$\begin{array}{llll}
\Gamma_1 \to \Delta_1, & \Gamma_2 \to \Delta_2, & \Gamma_1' \to \Delta_1', & \Gamma_2' \to \Delta_2', \quad \Gamma_{12} \to \Delta_1 + \Delta_2 \\
\Gamma_{12}' \to \Delta_1' + \Delta_2', & \Gamma_{15} \to \Delta_1 + \Delta_5, & \Gamma_{15}' \to \Delta_1' + \Delta_5, & \Gamma_{25} \to \Delta_2 + \Delta_5, \\
\Gamma_{25}' \to \Delta_2' + \Delta_5, & \Gamma_6^{\pm} \to \Delta_6, & \Gamma_7^{\pm} \to \Delta & \Gamma_8^{\pm} \to \Delta_6 + \Delta_7^{\pm}.
\end{array}$$

The compatibility conditions for representations at the points X and Δ are different for the groups O_h^1, O_h^5 on the one hand, and O_h^7 on the other. For O_h^1, O_h^5:

$$\begin{gathered}
X_1^+,\ X_4^- \to \Delta_1, \quad X_1^-,\ X_4^+ \to \Delta_2', \quad X_2^+,\ X_3^- \to \Delta_1', \\
X_2^-,\ X_3^+ \to \Delta_2, \quad X_5^{\pm} \to \Delta_5, \quad X_6^{\pm} \to \Delta_6, \quad X_7^{\pm} \to \Delta_7,
\end{gathered}$$

and for O_h^7:

$$X_{1,3} \to \Delta_5, \quad X_2 \to \Delta_1 + \Delta_2', \quad X_4 \to \Delta_1' + \Delta_2, \quad X_5 \to \Delta_6 + \Delta_7.$$

Compatibility conditions for points moving along the threefold axis in the groups O_h^5 and O_h^7:

$$\begin{array}{lll}
\Gamma_1,\ \Gamma_2' \to \Lambda_1, & \Gamma_2,\ \Gamma_1' \to \Lambda_2, & \Gamma_{12},\ \Gamma_{12}' \to \Lambda_3, \Gamma_{25}', \Gamma_{15} \to \Lambda_1 + \Lambda_3, \\
\Gamma_{15}',\ \Gamma_{25} \to \Lambda_3 + \Lambda_2, & \Gamma_7^{\pm},\ \Gamma_6^{\pm} \to \Lambda_6, & \Gamma_8^{\pm} \to \Lambda_4 + \Lambda_5 + \Lambda_6; \\
L_1,\ L_2' \to \Lambda_1, & L_2,\ L_1' \to \Lambda_2, & L_3,\ L_3' \to \Lambda_3, \\
L_4,\ L_5' \to \Lambda_4, & L_5,\ L_4' \to \Lambda_5, & L_6,\ L_6' \to \Lambda_6.
\end{array}$$

For the group O_h^1, the compatibility conditions at Γ and Λ are the same as in O_h^5 and O_h^7; at Λ and R they are the same as for Λ and Γ.

Compatibility conditions at Γ and Σ:

$$\begin{array}{lll}
\Gamma_1 \to \Sigma_1, \quad \Gamma_2 \to \Sigma_4, & \Gamma_{12} \to \Sigma_1 + \Sigma_4, & \Gamma_{15}' \to \Sigma_2 + \Sigma_3 + \Sigma_4, \\
\Gamma_{25}' \to \Sigma_1 + \Sigma_2 + \Sigma_3, & \Gamma_1' \to \Sigma_2,\ \Gamma_2' \to \Sigma_3, & \Gamma_{12}' \to \Sigma_2 + \Sigma_3, \\
\Gamma_{15} \to \Sigma_1 + \Sigma_4 + \Sigma_3, & \Gamma_{25} \to \Sigma_1 + \Sigma_2 + \Sigma_4, & \Gamma_6^{\pm} \to \Sigma_5,\ \Gamma_7^{\pm} \to \Sigma_5, \\
\Gamma_8^{\pm} \to 2\Sigma_5. & &
\end{array}$$

In the groups O_h^5 and O_h^7, the representations at Σ and K transform continuously into one another; the compatibility conditions for these groups at Σ and K are therefore trivial.

Compatibility conditions at M and Σ in the group O_h^1:

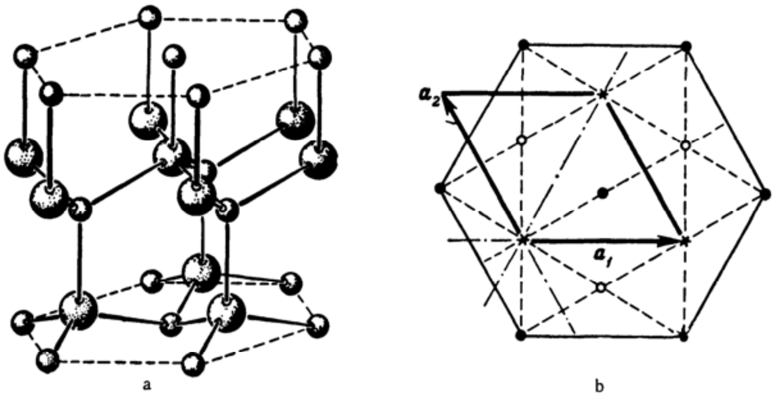
$$M_1^{\pm},\ M_4^{\pm} \to \Sigma_1,\quad M_2^{\pm},\ M_3^{\pm} \to \Sigma_2,\quad M_5^{\pm} \to \Sigma_2 + \Sigma_3,\quad M_6^{\pm},\ M_7^{\pm} \to \Sigma_5.$$

Compatibility conditions for the group T_d^2:

$$\begin{gathered}
\Gamma_1 \to \Delta_1,\quad \Gamma_2 \to \Delta_2,\quad \Gamma_{12} \to \Delta_1 + \Delta_2,\quad \Gamma_{15} \to \Delta_1 + \Delta_3 + \Delta_4,\\
\Gamma_{25} \to \Delta_2 + \Delta_3 + \Delta_4,\quad \Gamma_6 \to \Delta_5,\quad \Gamma_7 \to \Delta_5,\quad \Gamma_8 \to 2\Delta_5,\\
X_1 \to \Delta_1,\qquad X_2 \to \Delta_2,\quad X_3 \to \Delta_2,\quad X_4 \to \Delta_1,\quad X_5 \to \Delta_3 + \Delta_4,\\
X_6 \to \Delta_5,\quad X_7 \to \Delta_5,\\
\Gamma_1 \to \Lambda_1,\quad \Gamma_2 \to \Lambda_2,\quad \Gamma_{12} \to \Lambda_3,\quad \Gamma_{15} \to \Lambda_1 + \Lambda_3,\quad \Gamma_{25} \to \Lambda_2 + \Lambda_3,\\
\Gamma_6 \to \Lambda_6,\quad \Gamma_7 \to \Lambda_7,\quad \Gamma_8 \to \Lambda_4 + \Lambda_5 + \Lambda_6,\\
\Gamma_1 \to \Sigma_1,\quad \Gamma_2 \to \Sigma_2,\quad \Gamma_{12} \to \Sigma_1 + \Sigma_2,\quad \Gamma_{25} \to \Sigma_1 + 2\Sigma_2,\\
\Gamma_{15} \to \Sigma_2 + 2\Sigma_1,\qquad \Gamma_6 \to \Sigma_3 + \Sigma_4,\quad \Gamma_7 \to \Sigma_3 + \Sigma_4,\quad \Gamma_8 \to 2\Sigma_3 + 2\Sigma_4.
\end{gathered}$$

Group C_{6v}^4

Let us consider the representations of the little group at a few points of $\boldsymbol{k}$-space in crystals with a wurtzite lattice. This lattice is characteristic for a number of semiconductors such as hexagonal CdS and CdSe.



a

b

FIGURE 26. Wurtzite lattice: general view (a) and projection (b). The atoms of the first layer are indicated by solid dots, the atoms of the second layer by white dots, the position of the sixfold axes by asterisks. The thick lines indicate the base of the primitive cell.

The wurtzite lattice (the hexagonal form of ZnS) is shown in Figure 26. The points of the Bravais lattice are at the sites of the sulfur atoms, depicted in Figure 26 by small spheres. In the space group C_{6v}^4 of the wurtzite lattice the sixfold axis is a screw axis and half of the reflection planes are glide planes. The twelve rotational elements of the group C_{6v}^4, together with nonprimitive translations, are given in the following table.

$(e\,\vert\,0)$	$(c_6\,\vert\,\tau)$	$(c_6^2\,\vert\,0)$	$(c_6^3\,\vert\,\tau)$	$(c_6^4\,\vert\,0)$	$(c_6^5\,\vert\,\tau)$
$(\sigma_1\,\vert\,\tau)$	$(c_6\sigma_1\,\vert\,0)$	$(c_6^2\sigma_1\,\vert\,\tau)$	$(c_6^3\sigma_1\,\vert\,0)$	$(c_6^4\sigma_1\,\vert\,\tau)$	$(c_6^5\sigma_1\,\vert\,0)$

Here $\tau = c/2$ is a nonprimitive translation vector directed along the axis c_6, where c is the edge of the primitive cell along the sixfold axis.

The Brillouin zone for a hexagonal lattice is a six-faced prism (see Figure 27). The point Γ is at the center of the Brillouin zone. The points Δ and A lie on the sixfold axis, Δ within and A on the boundary of the Brillouin zone. The points H, P and K are on a lateral edge of the prism, H at a vertex of the prism, K at the midpoint of the edge, and P at an arbitrary position on the edge.

The point group F_k at these points and the number of points in the star are given in the following table.

Group F_k	C_{6v}	C_{3v}
Points	Γ, Δ, A	H, P, K
Number of points in star	1	2

The representations of the little group at the point $\Gamma(\boldsymbol{k}=0)$ coincide with the representations of the point group C_{6v}. This group has four one-dimensional single-valued representations A_1, A_2, B_2, B_1, and two two-dimensional ones E_1, E_2, denoted by $\Gamma_1, \ldots, \Gamma_6$; the double-valued representations E_1', E_2', E_3', denoted by $\Gamma_7, \Gamma_8, \Gamma_9$, are two-dimensional. All the representations at Γ belong to case (a$_1$). At the point Δ, $\boldsymbol{k} = k_0\,(001)$, the representations $\Delta_i\,(i = 1, \ldots, 9)$ of the little group for elements not containing primitive translations are obtained from the representations of C_{6v} by multiplying by $e^{-ik_0\tau_r}$, where τ_r is the nonprimitive translation corresponding to the rotational element r. The representations Δ_i depend on the position of the point Δ on the symmetry axis. Since no additional symmetry elements appear as Δ approaches the point A on the boundary of the Brillouin zone, the representations at A are obtained from those at Δ by setting $k_0 = \pi/c$.

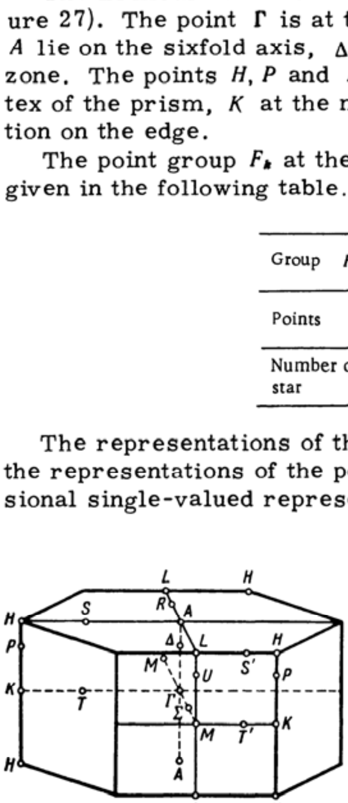


FIGURE 27. Brillouin zone for wurtzite lattice.

The difference between the points Δ and A is revealed when time reversal is introduced. Since for the point Δ the group C_{6v} contains no elements which transform $\boldsymbol{k}$ into $-\boldsymbol{k}$, all the representations at Δ belong to case (b$_3$).

At the point A we have $\boldsymbol{k} \equiv -\boldsymbol{k}$; Herring's test indicates that the representations $A_1, \ldots, A_8$ belong to case (b$_1$), but A_9 belongs to case (c$_1$). According to the results of § 18, time reversal makes the representations combine in pairs, $A_1 - A_4$, $A_2 - A_3$, $A_5 - A_6$, $A_7 - A_8$, and the representation A_9 must be doubled.

We now consider the representations of the little group for points on an edge of the prism. The symmetry group of such points is C_{3v}, and so, by the results of § 14, all the projective representations are equivalent to vector representations. All the coefficients of the factor system for these points equal unity, since the little group contains a nonprimitive translation $\boldsymbol{\tau}$ directed along the sixfold axis, and all the vectors $\boldsymbol{b}$ in (12.29) lie in a plane perpendicular to this axis. Thus the representations of the little group at P, K and H (for elements not containing primitive translations) are single- or double-valued representations of C_{3v} multiplied by $e^{-ik_z\tau_r}$, where $k_z = 0$ for K and $k_z = \pi/c$ for H.

The group C_{3v} has three single-valued representations A_1, A_2, E (one- and two-dimensional) and three double-valued representations E_1', E_2', E_3' (two-dimensional), giving representations $K_1, \ldots, K_6$, $P_1, \ldots, P_6$ and $H_1, \ldots, H_6$ of corresponding dimensions. The compatibility conditions of these representations for motion along the edge of the prism are trivial. At the point K, $\boldsymbol{k}$ and $-\boldsymbol{k}$ occur in the same star, and the representations $K_1, \ldots, K_6$ belong to case (a$_2$). At the point P, all the representations belong to case (b$_3$). At the point H, the representations H_1, H_2, H_4 and H_5 belong to case (b$_2$), the representations H_3 and H_6 to case (a$_2$). In this case the representations H_1, H_2 and H_4, H_5 are combined because of time reversal.

Distribution of normal modes according to irreducible representations

Using equations (15.28), we shall determine the irreducible representations according to which the normal modes transform at symmetric points of $\boldsymbol{q}$-space in diamond, zinc blende, halite and wurtzite lattices.

We first consider the diamond lattice. Table 23.3 presents the characters of the representations $\mathscr{D}_u$ according to which the components $u^{\nu}_{\varkappa l}(\boldsymbol{q})$ transform in this lattice. As noted above, the primitive cell of this lattice contains two atoms, at sites (000) and $(a_0/4)(111)$.

Since all 24 elements containing the nonprimitive translation $\boldsymbol{\tau} = (a_0/4)(111)$ take both atoms to inequivalent positions, we have $\chi^u_q(g) = 0$ for these elements. The other 24 elements leave the first atom in place and displace the second by one of the lattice vectors $\boldsymbol{a}_i$. Table 23.3 shows these displacements and gives the values of $\chi^u_q(g)$ for arbitrary $\boldsymbol{q}$, and also for the points $\Gamma(000)$, $X\left(\frac{2\pi}{a_0}(100)\right)$ and $L\left(\frac{\pi}{a_0}(111)\right)$ for all elements in the little group at these points. By (15.28a),

$$\chi^u_q(r\,|\,\boldsymbol{\tau}) = \chi_1(r) \sum_{\varkappa} \exp\left\{i\boldsymbol{q}\left[r^{-1}(\boldsymbol{X}_{\varkappa} - \boldsymbol{\tau}) - \boldsymbol{X}_{\varkappa}\right]\right\} \delta_{\varkappa,\, g\varkappa},$$

where $\chi_1(r)$ is the character of a vector representation. At the point Γ, for all 24 elements not containing the nonprimitive translation $\boldsymbol{\tau}$ (i.e., all elements of group T_d), $\chi_u(g) = 2\chi_{F_2}(g)$; for the other 24 elements $\chi_u(g) = 0$.

TABLE 23.3. Characters of representation $\mathscr{D}_u$ for diamond lattice

Group element	$\chi_1(r)$	Displacement of atom	Characters χ_q^u			
			arbitrary q	at $\Gamma\,(000)$	at $X\left(\frac{2\pi}{a_0}\,(100)\right)$	at $L\left(\frac{\pi}{a_0}\,(111)\right)$
$(e\,\vert\,0)$	3	0	6	6	6	6
$(c_{2x}\,\vert\,0)$	-1	a_3	$-1-e^{iqa_3}$	-2	-2	—
$(c_{2y}\,\vert\,0)$	-1	a_2	$-1-e^{iqa_2}$	-2	0	—
$(c_{2z}\,\vert\,0)$	-1	a_1	$-1-e^{iqa_1}$	-2	0	—
$(s_{4x}\,\vert\,0)$	-1	a_1	$-1-e^{iqa_1}$	-2	0	—
$(s_{4y}\,\vert\,0)$	-1	a_3	$-1-e^{iqa_3}$	-2	—	—
$(s_{4z}\,\vert\,0)$	-1	a_2	$-1-e^{iqa_2}$	-2	—	—
$(s_{4x}^3\,\vert\,0)$	-1	a_2	$-1-e^{iqa_2}$	-2	0	—
$(s_{4y}^3\,\vert\,0)$	-1	a_1	$-1-e^{iqa_1}$	-2	—	—
$(s_{4z}^3\,\vert\,0)$	-1	a_3	$-1-e^{iqa_3}$	-2	—	—
$(\sigma_{xy}\,\vert\,0)$	1	0	2	2	—	2
$(\sigma_{x\bar{y}}\,\vert\,0)$	1	a_1	$1+e^{iqa_1}$	2	—	—
$(\sigma_{xz}\,\vert\,0)$	1	0	2	2	—	2
$(\sigma_{x\bar{z}}\,\vert\,0)$	1	a_3	$1+e^{iqa_3}$	2	—	—
$(\sigma_{yz}\,\vert\,0)$	1	0	2	2	2	2
$(\sigma_{y\bar{z}}\,\vert\,0)$	1	a_3	$1+e^{iqa_3}$	2	2	—
$8\,(c_3\,\vert\,0),\,(c_3^2\,\vert\,0)$	0	—	0	0	0	0
$24(gi\,\vert\,\tau)$	—	Goes to inequivalent position	0	0	0	0

Direct lattice vectors:

$$a_1=(a_0/2)\,(110),\quad a_2=(a_0/2)\,(101),\quad a_3=(a_0/2)\,(011).$$

Reciprocal lattice vectors:

$$b_1=(2\pi/a_0)\,(11\bar{1}),\quad b_2=(2\pi/a_0)\,(\bar{1}11),\quad b_3=(2\pi/a_0)\,(\bar{1}11).$$

Consequently, $\mathscr{D}_u = F_2^+ + F_2^-$. For acoustical modes, describing the displacement of the cell as a whole, the components $a_{q\nu}$ transform like the components of a polar vector; hence acoustical modes belong to the representation $F_2^- = \Gamma_{15}$, while optical modes belong to $F_2^+ = \Gamma_{25}'$. Table 23.4 gives the decomposition of $\mathscr{D}_u$ into its irreducible constituents for the other points.

In a similar manner one easily determines the representations according to which the normal modes transform in the other lattices. The results are summarized in Table 23.4.

§24. ELECTRON SPECTRUM IN CUBIC CRYSTALS WITH SPACE GROUP O_h^1, O_h^5, O_h^7

In the present section we shall determine the possible shape of the spectrum in crystals with space groups O_h^1, O_h^5 and O_h^7 at the points of the Brillouin zone discussed in §23.

TABLE 23.4. Distribution of vibrations according to irreducible representations of $\boldsymbol{O_h}$

Type of crystal	Γ		X	Λ	L	Δ	Σ, K	H, P, K
	ac.	opt.						
Ge, Si	Γ_{15}	Γ'_{25}	X_1, X_2, X_3	$2\Lambda_1, 2\Lambda_3$	$L_1, L'_2,$ L_3, L'_3	$2\Delta_1, 2\Delta_5$	$2\Sigma_1, 2\Sigma_3,$ Σ_1, Σ_5	—
InSb	Γ_{25}	Γ_{25}	$X_1, X_1,$ $2X_5$	$2\Lambda_1, 2\Lambda_3$	$2L, 2L_3$	$2\Delta_1, 2\Delta_3,$ $2\Delta_4$	$4\Sigma_1, 2\Sigma_2$	—
PbS	Γ_{15}	Γ'_{25}	$2X_4^-, 2X_5^-$	$2\Lambda_1, 2\Lambda_3$	$L_1, L'_2,$ L_3, L'_3	$2\Delta_1, 2\Delta_5$	$2\Sigma_1, 2\Sigma_3,$ Σ_4, Σ_5	—
ZnS (wurtzite)	Γ_1, Γ_5	$\Gamma_1,$ $2\Gamma_4,$ $\Gamma_5, 2\Gamma_6$	—	—	—	$2\Delta_1, 2\Delta_4,$ $2\Delta_5, 2\Delta_6$	—	$2H_1,$ $2H_2,$ $4H_3$

The point Γ is a point of zero slope for all representations, since the group O_h contains inversion.

For the one-dimensional representations Γ_1, Γ_2, Γ'_1, Γ'_2, Table 21.1 (p.192) shows that the spectrum near $\boldsymbol{k}=0$ is described by one effective mass m^*:

$$E(\boldsymbol{k})=\frac{\hbar^2k^2}{2m^*}. \tag{24.1}$$

Since the two-dimensional representations Γ_{12} and Γ'_{12} differ only in their behavior under inversion, the spectrum for both these representations is the same.

Remembering that the momentum transforms according to the vector representation Γ'_{15}, we see from (21.25) that $N_2 = 2$, i.e., the matrix (21.19) and the spectrum for the representations Γ_{12} and Γ'_{12} are determined by two independent constants.

To construct the matrix $\mathcal{H}$, we take as basis functions of the representation Γ_{12}

$$\psi_1=z^2-\frac{1}{2}(x^2+y^2), \qquad \psi_2=\frac{\sqrt{3}}{2}(x^2-y^2).$$

Then, by (21.23),

$$\mathcal{H}=\begin{Vmatrix} Ak^2+\frac{B}{2}(k^2-3k_z^2) & \frac{\sqrt{3}}{2}B(k_x^2-k_y^2) \\ \frac{\sqrt{3}}{2}B(k_x^2-k_y^2) & Ak^2-\frac{B}{2}(k^2-3k_z^2) \end{Vmatrix}, \tag{24.2}$$

where

$$\begin{aligned} A&=\frac{\hbar^2}{2m}+\frac{\hbar^2}{2m^2}\sum_n\frac{|\langle 1|p_x|n\rangle|^2+|\langle 2|p_x|n\rangle|^2}{E_0-E_n}, \\ B&=\frac{2}{\sqrt{3}}\frac{\hbar^2}{m^2}\sum_n\frac{\langle 1|p_x|n\rangle\langle n|p_x|2\rangle}{E_0-E_n}. \end{aligned} \tag{24.3}$$

The eigenvalues of $\mathscr{H}$, which yield the two branches of the spectrum $E_{1,2}$, are

$$E_{1,2}(\boldsymbol{k}) = Ak^2 \pm |B| \sqrt{k^4 - 3(k_x^2 k_y^2 + k_x^2 k_z^2 + k_y^2 k_z^2)}. \tag{24.4}$$

In the [111] direction, when $k_x = k_y = k_z$, the band degeneracy is not removed: $E_1(\boldsymbol{k}) = E_2(\boldsymbol{k})$, in accordance with the compatibility conditions, which indicate that in the [111] direction the representations Γ_{12} and Γ_{12}' merge with the two-dimensional representation Λ_3.

Since each of the three-dimensional representations Γ_{15}, $\Gamma_1'^5$, Γ_{25} and Γ_{25}' is obtained by multiplying one of them by any one-dimensional representation of the cubic group, they are all projectively equivalent and, as will be shown in §25, the corresponding spectra near $\boldsymbol{k} = 0$ are identical. To construct the matrix $\mathscr{H}$ for the representation Γ_{15}', we choose basis functions X, Y, Z.

Using (21.19), we see that relative to this basis the matrix $\mathscr{H}$ is

$$\mathscr{H} = \begin{Vmatrix} Lk_x^2 + M(k_y^2 + k_z^2) & Nk_xk_y & Nk_xk_z \\ Nk_xk_y & Lk_y^2 + M(k_x^2 + k_z^2) & Nk_yk_z \\ Nk_xk_z & Nk_yk_z & Lk_z^2 + M(k_y^2 + k_x^2) \end{Vmatrix}, \tag{24.5}$$

where L, M, N are independent constants:

$$\begin{aligned} L &= \frac{\hbar^2}{2m} + \frac{\hbar^2}{m^2} \sum_n \frac{|\langle X | p^x | n \rangle|^2}{E_0 - E_n}, \\ M &= \frac{\hbar^2}{2m} + \frac{\hbar^2}{m^2} \sum_n \frac{|\langle X | p^y | n \rangle|^2}{E_0 - E_n}, \\ N &= \frac{\hbar^2}{m^2} \sum_n \frac{\langle X | p^x | n \rangle \langle n | p^y | Y \rangle + \langle X | p^y | n \rangle \langle n | p^x | Y \rangle}{E_0 - E_n}. \end{aligned} \tag{24.6}$$

The problem of determining the spectrum reduces to solving the third order characteristic equation $|\mathscr{H} - \boldsymbol{I}E| = 0$ (see (21.20)). An explicit solution may be obtained for the three symmetric directions [100], [111], [110]:

$$\begin{gathered} E_1 = Lk^2, \quad E_2 = E_3 = Mk^2, \quad k_x = k, \quad k_y = k_z = 0; \\ E_1 = \left(\frac{L + 2M}{3} + \frac{2}{3} N\right) k^2, \quad E_2 = E_3 = \frac{L + 2M - N}{3} k^2, \\ k_x = k_y = k_z = \frac{k}{\sqrt{3}}; \\ E_1 = \frac{L + M + N}{2} k^2, \quad E_2 = \frac{L + M - N}{2} k^2, \quad E_3 = Mk^2, \\ k_x = k_y = \frac{k}{\sqrt{2}}, \quad k_z = 0. \end{gathered} \tag{24.7}$$

It is evident from (24.7) that degeneracy is not completely removed along the [100] and [111] directions, but a twofold band degeneracy remains. This degeneracy is maintained along the threefold and fourfold axes, up to the boundary of the Brillouin zone, because of the compatibility conditions for these directions: $\Gamma_{15}' \to \Delta_1 + \Delta_5$, $\Gamma_{15}' \to \Lambda_1 + \Lambda_3$, although, of course, at a sufficient distance from Γ the spectrum will differ from (24.7).

If spin-orbit coupling is introduced, we have twofold representations $\Gamma_6^{\pm}$ and $\Gamma_7^{\pm}$ and fourfold representations $\Gamma_8^{\pm}$ at the point Γ; the spectra for the representations $\Gamma_6^{\pm}$, $\Gamma_7^{\pm}$ are identical, as they are for Γ_8^{+} and Γ_8^{-}.

Let us consider the spectrum for twofold degenerate bands.

By (21.23), the elements $\mathscr{H}_{ij}$ of the matrix $\mathscr{H}$ transform like

$$k_\alpha k_\beta \int \varphi_i^* (\pi^\alpha \pi^\beta + \pi^\beta \pi^\alpha) \varphi_j \, d\tau,$$

where φ_i, φ_j ($i, j = 1, 2$; $\alpha, \beta = x, y, z$) are basis functions of the representation $\Gamma_6^\pm$ or $\Gamma_7^\pm$. Since the representations for the point Γ belong to case (a_1), it follows from (21.31) that the selection rules for these matrix elements depend on the representation that belongs to the antisymmetrized product of the functions $\varphi_i\varphi_j$:

$$\{\Gamma_6^\pm \times \Gamma_6^\pm\} = \{\Gamma_7^\pm \times \Gamma_7^\pm\} = \Gamma_1.$$

Thus the matrix element is nonzero only for invariant combinations $\pi_\alpha\pi_\beta$, i.e., for $\pi^2 = \pi_x^2 + \pi_y^2 + \pi_z^2$. Moreover, as the representations are orthogonal,

$$k^2 \int \varphi_i^* \pi^2 \varphi_j \, d\tau \sim k^2 \delta_{ij},$$

so that $\mathscr{H}_{ij} \sim E(\boldsymbol{k})\delta_{ij}$ and the spectrum $E(\boldsymbol{k})$ is twofold degenerate, coinciding with the spectrum (24.1) for a nondegenerate band.

Note that this result was derived independently of any specific choice of basis functions; it is therefore valid for any spin-orbit coupling.

We now consider the spectrum near Γ for the representation $\Gamma_8^\pm$. When electron spin is introduced, the threefold degeneracy at Γ, corresponding to the representation Γ_{15}', is doubled and becomes sixfold, and the basis functions

$$X\alpha, \quad Y\alpha, \quad Z\alpha, \quad X\beta, \quad Y\beta, \quad Z\beta \tag{24.8}$$

transform according to the representation $\Gamma_{15}' \times \mathscr{D}_{1/2}$. In (24.8) α and β are spin functions corresponding to spin projection ± 1/2. If spin-orbit coupling is disregarded, the matrix $\mathscr{H}$ relative to the basis (24.8) is

$$\mathscr{H} = \begin{Vmatrix} \mathscr{H}_1 & 0 \\ 0 & \mathscr{H}_1 \end{Vmatrix}, \tag{24.9}$$

where $\mathscr{H}_1$ is the matrix (24.5).

Spin-orbit coupling causes the sixfold degenerate band at Γ to split into a fourfold degenerate band (representation Γ_8^-) and a twofold degenerate band Γ_6^-. If the spin-orbit splitting Δ of the bands Γ_8^- and Γ_6^- at $\boldsymbol{k} = 0$ is significantly less than the separation to other bands, the functions (23.2) and (23.3) may be taken as basis functions for the representations Γ_8^- and Γ_6^-.

In the new representation defined by functions (23.2) and (23.3) relative to which the operator $\mathscr{H}_{so}$ is diagonal, the matrix $\mathscr{H}$ of (24.9) becomes

$$\mathscr{H}' = S^{-1}\mathscr{H}S, \tag{24.10}$$

where S is a unitary matrix transforming the functions (24.8) into (23.2) and (23.3). Direct computation of $\mathscr{H}'_{j'm',\, jm}$ yields the matrix shown in Table 24.1.

Table 24.1 uses the following abbreviations:

$$\begin{aligned} &F = Ak^2 + \frac{B}{2}(k^2 - 3k_z^2), \qquad G = Ak^2 - \frac{B}{2}(k^2 - 3k_z^2), \\ &H = -Dk_z(\underset{\mathsf{X}}{k_z} + ik_y), \qquad I = \frac{\sqrt{3}}{2}B(k_x^2 - k_y^2) - iDk_x k_y, \end{aligned} \tag{24.11}$$

where

$$A = \frac{L + 2M}{3}, \quad B = \frac{L - M}{3}, \quad C^2 = D^2 - 3B^2, \quad D = \frac{N}{\sqrt{3}}. \tag{24.12}$$

To determine the spectrum with allowance for spin-orbit splitting in the general case, we must compute the eigenvalues of the matrix in Table 24.1.

TABLE 24.1

j', m' \ j, m	$\frac{3}{2}, \frac{3}{2}$	$\frac{3}{2}, \frac{1}{2}$	$\frac{3}{2}, -\frac{1}{2}$	$\frac{3}{2}, -\frac{3}{2}$	$\frac{1}{2}, \frac{1}{2}$	$\frac{1}{2}, -\frac{1}{2}$
$\frac{3}{2}, \frac{3}{2}$	F	H	I	0	$\frac{iH}{\sqrt{2}}$	$-i\sqrt{2}I$
$\frac{3}{2}, \frac{1}{2}$	H^*	G	0	I	$\frac{(G-F)i}{\sqrt{2}}$	$\frac{i\sqrt{3}H}{\sqrt{2}}$
$\frac{3}{2}, -\frac{1}{2}$	I^*	0	G	$-H$	$-i\sqrt{\frac{3}{2}}H^*$	$\frac{i(G-F)}{\sqrt{2}}$
$\frac{3}{2}, -\frac{3}{2}$	0	I^*	$-H^*$	F	$-i\sqrt{2}I^*$	$-\frac{iH^*}{\sqrt{2}}$
$\frac{1}{2}, \frac{1}{2}$	$-\frac{iH^*}{\sqrt{2}}$	$-\frac{(G-F)i}{\sqrt{2}}$	$i\sqrt{\frac{3}{2}}H$	$i\sqrt{2}I$	$\frac{(F+G)}{2}-\Delta$	0
$\frac{1}{2}, -\frac{1}{2}$	$i\sqrt{2}I^*$	$-i\sqrt{\frac{3}{2}}H^*$	$-\frac{i(G-F)}{\sqrt{2}}$	$\frac{iH}{\sqrt{2}}$	0	$\frac{(F+G)}{2}-\Delta$

If Δ is large compared to the current carrier kinetic energy, measured from the bottom of each of the split-off bands, the problem is simplified, for then it is sufficient to determine the eigenvalues of the fourth-order matrix

$$\begin{Vmatrix} F & H & I & 0 \\ H^* & G & 0 & I \\ I^* & 0 & G & -H \\ 0 & I^* & -H^* & F \end{Vmatrix} \tag{24.13}$$

and of a second-order diagonal matrix; these matrices give the spectrum of each of the split-off bands separately. Solving the corresponding characteristic equations, we find the spectrum in the band Γ_8:

$$E_{1,2} = \frac{F+G}{2} \pm \sqrt{\left(\frac{F-G}{2}\right)^2 + |H|^2 + |I|^2}, \tag{24.14}$$

whence

$$E_{1,2} = Ak^2 \pm \sqrt{B^2k^4 + C^2(k_x^2k_y^2 + k_x^2k_z^2 + k_y^2k_z^2)}, \tag{24.14a}$$

and in the band Γ_6:

$$E_3 = \frac{F+G}{2} - \Delta = Ak^2 - \Delta. \tag{24.14b}$$

Each eigenvalue is twofold degenerate, owing to time reversal.

For the degenerate band Γ_8, it follows from the results of §17 that the second derivative of the energy with respect to the components of the vector $\boldsymbol{k}$ has a discontinuity at $\boldsymbol{k} = 0$. The matrix (24.13) and the spectrum (24.14a) for the representation Γ_8 correspond to the weak spin-orbit coupling approximation. It is easily shown that for any spin-orbit coupling the matrix $\mathcal{H}$ and spectrum $E(\boldsymbol{k})$ have the same form as (24.13) and (24.14); in particular, they are determined by the three independent constants A, B, D, as may be verified from (21.31).

However, the relations (24.12) between the constants A, B, D for the level Γ_8 and L, M, N for the level Γ_{15}, like the relationship of the effective mass in the band Γ_6 to these constants, are valid only for weak spin-orbit coupling (see (26.15a)).

The surfaces of constant energy (24.14) are warped spheres.

By (21.21), the wave functions of the holes, to lowest order in k, are

$$\psi_{l\boldsymbol{k}} = e^{i\boldsymbol{k}\boldsymbol{x}} \sum_{m'=\pm 3/2,\ \pm 1/2} c^l_{m'}(\boldsymbol{k})\,\psi^{3/2}_{m'}, \tag{24.15}$$

where the functions $\psi^{3/2}_{m'}$ are defined by (23.2).

It is convenient to write the wave functions (24.15) as a column matrix of expansion coefficients $c^l_{m'}(\boldsymbol{k})$ of $\psi_{l\boldsymbol{k}}$ in terms of the functions $\psi^{3/2}_{m'}$ of the band bottom:

$$\psi_{l\boldsymbol{k}} = e^{i\boldsymbol{k}\boldsymbol{x}} \begin{Vmatrix} c^l_{3/2}(\boldsymbol{k}) \\ c^l_{1/2}(\boldsymbol{k}) \\ c^l_{-1/2}(\boldsymbol{k}) \\ c^l_{-3/2}(\boldsymbol{k}) \end{Vmatrix} \equiv e^{i\boldsymbol{k}\boldsymbol{x}} \mathcal{F}_{l\boldsymbol{k}}. \tag{24.16}$$

The matrix $\boldsymbol{c}$ is determined by solving the system of equations (21.18) with the matrix $\mathcal{H}$ (24.13). Since the eigenvalues of this matrix are doubly degenerate owing to time reversal, the coefficients of the matrix $\boldsymbol{c}$ are determined not uniquely but only up to a unitary transformation of the degenerate functions. Since the representation Γ_8 belongs to case (a$_1$), there is a linear relation (18.7) between the basis functions $\psi_m^{3/2}$ and the functions $K\psi_m^{3/2}=\sigma_y K_0 \psi_m^{3/2}$ It is easy to see that for the functions (23.2) this gives

$$K\psi_{3/2}^{3/2}=\psi_{-3/2}^{3/2},\quad K\psi_{1/2}^{3/2}=-\psi_{-1/2}^{3/2},\quad K\psi_{-1/2}^{3/2}=\psi_{1/2}^{3/2},\quad K\psi_{-3/2}^{3/2}=-\psi_{3/2}^{3/2}. \tag{24.17}$$

These relations determine the matrix $\boldsymbol{T}$ (18.7).

Relative to the basis $K\psi$, the matrix $\mathcal{H}$ becomes*

$$\boldsymbol{T}^{-1}\mathcal{H}\boldsymbol{T}=\mathcal{H}^{*},$$

and this implies that the eigenvalues of the matrix (24.13) are doubly degenerate. The coefficients c_i^l and $c_i^{l'}$ of the two degenerate functions $\psi_{l\boldsymbol{k}}$ and $\psi_{l'\boldsymbol{k}}$ may be subjected to a condition similar to (24.17):

$$c_{3/2}^{l'}=-c_{-3/2}^{*l},\quad c_{1/2}^{l'}=c_{-1/2}^{*l},\quad c_{-1/2}^{l'}=-c_{1/2}^{*l},\quad c_{-3/2}^{l'}=c_{3/2}^{*l}. \tag{24.18}$$

The eigenfunctions $\mathcal{F}_{\boldsymbol{k}}$ of the matrix $\mathcal{H}$ which satisfy condition (24.18) are

$$\begin{aligned}
\mathcal{F}_{\boldsymbol{k}1}&=\frac{1}{\sqrt{(E_1-F)(E_1-E_2)}}\left\|\begin{matrix} H \\ E_1-F \\ 0 \\ I^{*} \end{matrix}\right\|;\\
\mathcal{F}'_{\boldsymbol{k}1}&=\frac{1}{\sqrt{(E_1-F)(E_1-E_2)}}\left\|\begin{matrix} -I \\ 0 \\ -(E_1-F) \\ H^{*} \end{matrix}\right\|\\
\mathcal{F}_{\boldsymbol{k}2}&=\frac{1}{\sqrt{(E_2-F)(E_2-E_1)}}\left\|\begin{matrix} H \\ E_2-F \\ 0 \\ I^{*} \end{matrix}\right\|;\\
\mathcal{F}'_{\boldsymbol{k}2}&=\frac{1}{\sqrt{(E_2-F)(E_2-E_1)}}\left\|\begin{matrix} -I \\ 0 \\ -(E_2-F) \\ H^{*} \end{matrix}\right\|,
\end{aligned} \tag{24.19}$$

where $E_{1,2}$, I, F and H are defined by (24.11) and (24.14).

In agreement with the general results of §17 and §21, the functions $\mathcal{F}_{\boldsymbol{k}l}$ and $\psi_{\boldsymbol{k}l}$ depend only on the direction of the vector $\boldsymbol{k}$, and as functions of $\boldsymbol{k}$ have a discontinuity at $\boldsymbol{k}=0$.

In cases where the energy of the holes is comparable with the spin-orbit splitting of the bands Γ_8 and Γ_6, we must find the eigenvalues of the matrix of Table 24.1. This reduces to solution of a cubic equation, since each eigenvalue is doubly degenerate owing to time reversal.

Using the equations (15.49) and (15.49a) of perturbation theory, we easily find the corrections to the energy $E_{1,2}(\boldsymbol{k})$ due to the interaction of the bands Γ_8

---

* Note that the invariance of $\mathcal{H}$ under time reversal and the form of the matrix $\boldsymbol{T}$ as determined by (24.17) uniquely determine the form (24.13) of $\mathcal{H}$ for the representation Γ_8.

and Γ_6. This means introducing suitable corrections to the matrix (24.13); for the matrix $\mathcal{H}$ of Table 24.1, these are

$$\begin{aligned} F' &= F + \frac{1}{2\Delta}(|H|^2 + 4|I|^2), \\ G' &= G + \frac{1}{2\Delta}((F-G)^2 + 3|H|^2), \\ H' &= H + \frac{1}{2\Delta}(H(G-F) - 2\sqrt{3}H^{*}I), \\ I' &= I - \frac{1}{2\Delta}(\sqrt{3}H^2 + 2I(G-F)). \end{aligned} \tag{24.20}$$

Hence we obtain an energy correction $\delta E_{1,2}(k) \sim k^4/\Delta$, which shows that the interaction of the bands Γ_8 and Γ_6 distorts the parabolicity of the spectrum:

$$\delta E_{1,2} = \frac{1}{\Delta}\{B^2k^4 + C^2\Theta \pm \frac{1}{2}\frac{2B^3k^6 + 9B\left(\frac{D^2}{3} - B^2\right)k^2\Theta + 27\left(\frac{2D}{\sqrt{3}} + B\right)k_x^2k_y^2k_z^2}{\sqrt{B^2k^4 + C^2\Theta}}, \tag{24.21}$$

where $\Theta = k_x^2k_y^2 + k_x^2k_z^2 + k_y^2k_z^2$.

The equations obtained above determine the possible types of spectrum at Γ for all the space groups of the class O_h.

Point Δ. The representations at Δ belong to case (a_2); thus, to determine the number of nonzero momentum matrix elements we must use (21.28) and (21.32), with inversion playing the part of the element R. As noted in §18, at interior points of the Brillouin zone equations (21.28) and (21.32) involve the characters of the representations of the crystallographic point group, and so the results obtained below for Δ are valid for all the groups of class O_h. In view of the fact that the operator $\boldsymbol{\pi}$ (or $\boldsymbol{p}$) is odd with respect to both space inversion and time reversal, equations (21.28) and (21.32) have the form

$$N = \sum_{r \in F_{\boldsymbol{k}}} \chi(r)(|\chi_v(r)|^2 + K^2\chi_v(r^2)). \tag{24.22}$$

For all one-dimensional single-valued representations, $[\Delta_i^* \times \Delta_i] = A_1$; for double-valued representations, $\{\Delta_i^* \times \Delta_i\} = A_1$. For two-dimensional single-valued representations, $[\Delta_5^* \times \Delta_5] = A_1 + A_2 + B_2$.

The representation according to which the components of the vectors $\boldsymbol{\pi}$ or $\boldsymbol{p}$ transform is reducible; its irreducible constituents are the identity representation A_1, according to which π_x transforms, and the two-dimensional representation E, according to which π_y and π_z transform. It follows from (24.22) that, for all the representations, the only nonzero component of the momentum is $\pi_x(p_x)$.

According to Table 21.1, the surface of constant energy for the one-dimensional representations at Δ near the extremum is an ellipsoid of revolution:

$$E(\boldsymbol{k}) = E(\boldsymbol{k}_0) + \frac{\hbar^2k_x^2}{2m_\parallel} + \frac{\hbar^2(k_y^2 + k_z^2)}{2m_\perp}. \tag{24.23}$$

For the two-dimensional representation Δ_5 with basis functions Y, Z, the matrix $\mathcal{H}$ at the extremum point has the form

$$\mathcal{H} = \begin{Vmatrix} Ak_x^2 + Bk_y^2 + Ck_z^2 & Dk_yk_z \\ Dk_yk_z & Ak_x^2 + Ck_y^2 + Bk_z^2 \end{Vmatrix}, \tag{24.24}$$

where

$$A=\frac{\hbar^2}{2m}+\frac{\hbar^2}{m^2}\sum_n\frac{|\langle Y|p^x|n\rangle|^2}{E_0-E_n},$$
$$D=\frac{\hbar^2}{m^2}\sum_n\frac{\langle Y|p^z|n\rangle\langle n|p^y|Z\rangle+\langle Z|p^z|n\rangle\langle n|p^y|Y\rangle}{E-E_n};\qquad(24.25)$$

B and C are given by formulas like that for A, except that p^x is replaced by p^y and p^z, respectively.

The spectrum $E_{1,2}$ for the representations Δ_5 is

$$E_{1,2}(\boldsymbol{k})=E(\boldsymbol{k}_0)+Ak_x^2+\frac{B+C}{2}(k_y^2+k_z^2)\pm\sqrt{\left(\frac{C-B}{2}\right)^2(k_y^2-k_z^2)^2+D^2k_y^2k_z^2},\qquad(24.26)$$

depending on the four parameters A, B, C, D.

When spin and spin-orbit coupling are introduced, the one-dimensional representations become two-dimensional representations Δ_6 and Δ_7, while the four-dimensional representation $\Delta_5\times\mathscr{D}_{1/2}$ is split into two two-dimensional representations.

Since $\{\Delta_7^*\times\Delta_7\}=\{\Delta_6^*\times\Delta_6\}=A_1$, the number of independent constants defining the spectrum at a point of zero slope is equal to the number of identity representations occurring in the direct products $[\Delta_1\times\Delta_1]=A_1$ and $[\Delta_5\times\Delta_5]=$ $=A_1+B_2+A_2$, i.e., two. These invariant combinations $\pi^\alpha\pi^\beta+\pi^\beta\pi^\alpha$ are π_x^2 and $\pi_y^2+\pi_z^2$, and therefore, for double-valued representations, $\mathscr{H}_{ij}=$ $=(Ak_x^2+B(k_y^2+k_z^2))\delta_{ij}$; in other words, for double-valued representations the spectrum at Δ coincides with the spectrum (24.23) in a nondegenerate band.

Point X. Since the point group D_{4h} at X contains inversion, which commutes with all elements of the space groups O_h^1 and O_h^5, the wave functions at X possess definite parity and therefore X is a point of zero slope for all representations of O_h^1 and O_h^5.

Since $D_{4h}=C_{4v}\times C_i$, the spectrum at the point X for O_h^1 and O_h^5 must coincide with the spectrum at a point of zero slope Δ and is defined by (24.23) and (24.26). The spectrum at the point M for the group O_h^1 must conform to the various band structures at the point X for this group. The situation is otherwise at the point X for the group O_h^7, where there is an essential degeneracy.

In the group O_h^7 inversion i occurs together with a nonprimitive translation $\boldsymbol{\tau}$; thus it does not commute with all the elements of the space group and the basis functions of the representations of the little group do not possess definite parity. Therefore, in the group O_h^7 the mere presence of inversion in the point group does not automatically imply the vanishing of momentum matrix elements. Since the representations at X belong to case (a_1), the number of linearly independent components of the momentum operator must be determined via (21.25) and (21.31).

Since $D_{4h}=D_{2d}\times C_i$, we shall expand the product $X_i\times X_i$ $(i=1,2,\ldots,5)$ in terms of the representations of the group $D_{2d}\times C_i$.

Using the character tables we readily verify that

$$\{X_1^*\times X_1\}=\{X_3^*\times X_3\}=A_2^-,\ [X_1^*\times X_1]=[X_3^*\times X_3]=A_1^++B_1^-+B_2^+,$$
$$\{X_2^*\times X_2\}=\{X_4^*\times X_4\}=B_2^-,\ [X_2^*\times X_2]=[X_4^*\times X_4]=A_1^++A_1^-+B_2^+,$$
$$[X_5^*\times X_5]=A_2^++A_2^-+B_1^++B_2^++2E^++E^-,$$
$$\{X_5^*\times X_5\}=A_1^++A_1^-+B_1^-+B_2^++E^-.\qquad(24.27)$$

Since the momentum components at X transform according to the representations B_2^- (π_x) and E^- (π_y, π_z), it follows from (24.27) that X is a point of zero slope for the representations X_1 and X_3, while for the representations X_2 and X_4 the matrix element p_x is nonzero (and hence so is $\partial E/\partial k_x$). For the spinor representation X_5, all three components π_x, π_y, π_z are nonzero. The matrix elements π_y, π_z are relativistically small, but π_x is relativistically small only for the representations X_5 which originate from representations X_1, X_3.

Let us examine the spectrum at X for the representations X_1, X_3. According to (21.25), the matrix elements of the quadratic components of the momentum are determined by the symmetrized products.

We expand the symmetrized product $[\mathscr{D}_1^2]$ in terms of irreducible representations of $D_{2d} \times C_i$:

$$[\mathscr{D}_1^2] = 2A_1^+ + B_1^+ + B_2^+ + E^+.$$

Thus, by (24.27), the matrix $\mathscr{H}$ and the spectrum $E(\boldsymbol{k})$ are defined by combinations of components: $k_x^2, k_y^2 + k_z^2$, which transform according to the representation A_1, and $k_y k_z$, which transforms according to B_2. The form of the matrix $\mathscr{H}$ depends on the choice of basis functions. The construction of the matrix $\mathscr{H}$ in (24.2), (24.5) and (24.24) utilized the explicit form of the basis functions. However, we can construct $\mathscr{H}$ without knowing the explicit form of the basis functions, provided the representation matrices are at our disposal.

For the representations X_1 and X_3, we construct $\mathscr{H}$ relative to the basis functions ψ_1 and ψ_2 defined by the transformation matrices given in Table 14.2 (p. 114) for the generators of the group D_{2d}.

For the invariant combinations p_x^2 and $p_y^2 + p_z^2$, the matrix elements (21.23) of $\mathscr{H}_{ij}$ are

$$\langle\psi_i|p_x^2|\psi_j\rangle = \langle\psi_1|p_x^2|\psi_1\rangle\,\delta_{ij}, \quad \langle\psi_i|p_y^2 + p_z^2|\psi_j\rangle = \langle\psi_1|p_y^2 + p_z^2|\psi_1\rangle\,\delta_{ij}.$$

For the diagonal matrix elements, $\langle\psi_1|p_y p_z|\psi_1\rangle = \langle\psi_2|p_y p_z|\psi_2\rangle = 0$, since Table 14.2 indicates that they change sign under the transformations s_4. The only off-diagonal element of $\mathscr{H}$ relative to this basis is $\langle\psi_1|p_y p_z|\psi_2\rangle$; therefore, by (21.23), the matrix $\mathscr{H}$ for the representations X_1, X_3 is

$$\mathscr{H} = \begin{Vmatrix} Ak_x^2 + B(k_y^2 + k_z^2) & Ck_y k_z \\ C^* k_y k_z & Ak_x^2 + B(k_y^2 + k_z^2) \end{Vmatrix}, \tag{24.28}$$

where A, B, C are constants.

The energy spectrum $E_{1,2}(\boldsymbol{k})$ for the representations X_1 and X_3 is

$$E_{1,2}(\boldsymbol{k}) = E(\boldsymbol{k}_0) + Ak_x^2 + B(k_y^2 + k_z^2) \pm |C|\, k_y k_z, \tag{24.29}$$

and the surfaces of constant energy (24.29) are two ellipsoids with principal axes in the [100], [011] and $[01\bar{1}]$ directions.

Although none of these is invariant under all transformations in the group D_{4h}, the spectrum (24.29) as a whole is invariant under D_{4h}.

Point Λ. At the point Λ, for the one-dimensional single-valued representations Λ_1 and Λ_2, the momentum has one nonzero component, parallel to the threefold axis. At a point of zero slope the surfaces of constant energy are ellipsoids of revolution with principal axis (z) directed along the threefold axis. Relative to these axes,

$$E(\boldsymbol{k}) = E(\boldsymbol{k}_0) + \frac{\hbar^2 k_z^2}{2m_\parallel} + \frac{\hbar^2 (k_y^2 + k_x^2)}{2m_\perp}. \tag{24.30}$$

For the two-dimensional single-valued representation Λ_3, which belongs to case (a_2), we must use equations (21.28) and (21.32), with inversion taken as the element R, to determine N_1. It follows from (21.28) that the nonzero momentum components for the representation Λ_3 are those which transform according to $[E \times E] = A_1 + E$, i. e , all three components. Thus Λ is not a point of zero slope for the representation Λ_3.

The one-dimensional double-valued representations Λ_4 and Λ_5 belong to case (b_2) and must be combined. According to (21.22) there is one nonzero linearly independent constant, corresponding to the diagonal momentum matrix element π_z; the off-diagonal momentum matrix elements vanish.

By (21.23), at a point of zero slope, where the longitudinal momentum component vanishes, the spectrum is twofold degenerate and is defined by equation (24.30). The two-dimensional double-valued representation Λ_6 belongs to case (a_2), and by (21.32) the numbers N_1 and N_2 depend upon the antisymmetrized product $\{\Lambda_6 \times \Lambda_6\} = A_1$. At the point Λ, therefore, only the longitudinal component π_z is nonzero for the representation Λ_6, and at a point of zero slope the spectrum is defined by equation (24.29).

Thus, at the point Λ near the extremum the surfaces of constant energy for all the representations are ellipsoids of revolution (24.30) for all the groups O_h^1, O_h^5 and O_h^7.

Point L. Since the little group at L contains inversion, which commutes with all the group elements, L is a point of zero slope for all the representations.

For the single-valued one-dimensional representations $L_1^\pm$ and $L_2^\pm$, the spectrum $E(\boldsymbol{k})$ at L is defined by equation (24.30). The spectrum is the same for the two-dimensional single-valued representations $L_3^\pm$. For L_3^-, the basis functions may be taken as

$$\psi_1 = X + iY, \quad \psi_2 = X - iY,$$

where the x- and y-axes lie in a plane perpendicular to the threefold z-axis. Relative to this basis, the matrix $\mathcal{H}$ has the form

$$\mathcal{H} = \begin{Vmatrix} Ak_z^2 + B(k_x^2 + k_y^2) & Ck_zk_- \\ C^*k_zk_+ & Ak_z^2 + B(k_x^2 + k_y^2) \end{Vmatrix} \tag{24.31}$$

giving the spectrum

$$E_{1,2}(\boldsymbol{k}) = E(\boldsymbol{k}_0) + Ak_z^2 + B(k_y^2 + k_x^2) \pm \tfrac{1}{2}|C|k_zk_\perp; \tag{24.32}$$

in these equations $k_\pm = (k_x \pm ik_y)/2$, $k_\perp = \sqrt{k_x^2 + k_y^2}$, and A, B, C are constants.

For the two-dimensional spinor representations $L_6^\pm$, we see from (21.31) that $\{L_6^\pm \times L_6^\pm\} = A_1^+$; thus the appropriate spectrum is defined by equation (24.30) and is twofold degenerate.

For the combined representations L_4^+, L_5^+ and L_4^-, L_5^-, we must use (21.22) and (21.24). It follows from these formulas that for these representations the matrix $\mathcal{H}$ contains only the combinations k_z^2, $k_x^2 + k_y^2$, which are invariant under the group D_{3d}. The matrix is therefore a multiple of the identity; its eigenvalues are twofold degenerate and are defined by (24.30).

Point Σ. Table 21.1 tells us that for the one-dimensional representations Σ_1, Σ_2, Σ_3 and Σ_4 only one matrix element of the momentum component p_z does not vanish, where the z-axis lies along the twofold axis.

At a point of zero slope the surfaces of constant energy for the one-dimensional representations are ellipsoids:

$$E(\boldsymbol{k}) = E(\boldsymbol{k}_0) + \frac{\hbar^2 k_x^2}{2m_1} + \frac{\hbar^2 k_y^2}{2m_2} + \frac{\hbar^2 k_z^2}{2m_3}, \qquad (24.33)$$

where the x- and y-axes are directed along the intersection of a plane perpendicular to the twofold axis and the reflection planes σ_v.

For the two-dimensional spinor representation Σ_5, it follows from (21.28) that only the matrix element of the momentum component π_z is nonzero. At a point of zero slope the spectrum is twofold degenerate and is determined by equation (24.33).

All the results obtained for the point Σ are valid for the corresponding representations at the point K.

The Kane model

In certain semiconductors with the zinc blende lattice, such as InSb, the conduction and valence band extrema lie at the same point Γ and the band gap is very narrow. For these crystals, therefore, the $\boldsymbol{k}^2$ approximation for $E(\boldsymbol{k})$ often proves inadequate, and $\mathscr{H}(\boldsymbol{k})$ must incorporate both conduction and valence bands from the start. If we ignore the relativistic linear $\boldsymbol{k}$ terms, which will be considered in §26, then $\mathscr{H}(\boldsymbol{k})$ for these crystals at Γ does not differ from the corresponding Hamiltonian for the Ge and Si lattices as is the case for other cubic crystals. In accordance with (23.2), we take the odd functions ψ_m^{j-} with $j = 3/2$ and $j = 1/2$ as basis functions for the valence band, and the even functions $\psi_m^{1/2+}$ as basis functions for the conduction band. By (21.38), the matrix $\mathscr{H}(\boldsymbol{k})$ then has the form shown in Table 24.2, where

$$p = \frac{\hbar}{m}\langle S|\, p_x\, |X\rangle, \quad k_\pm = \frac{1}{\sqrt{2}}(k_x \pm ik_y). \qquad (24.34)$$

The constants L', M', N' appearing in F', G' H', I' differ from the constants L, M, N in F, G, H, I in Table 24.1 since they incorporate the interaction of the valence band with all bands other than the conduction band, whereas L, M, N also include the interaction with the conduction band. In order to go over from Table 24.2 to Table 24.1, we must eliminate the interband terms representing this interaction. Similarly, the sum

$$\frac{1}{2m_c'} = \frac{1}{m^2}\sum_n \frac{|\langle S|\, p_x\, |n\rangle|^2}{E_0 - E_n} + \frac{1}{2m}$$

includes the interaction with all bands n except the valence band.

TABLE 24.2. Matrix $\mathscr{H}(\boldsymbol{k})$ for two-band model

	1/2, 1/2	1/2, −1/2	3/2, 3/2	3/2, 1/2	3/2, −1/2	3/2, −3/2	1/2, 1/2	1/2, −1/2
1/2, 1/2	$E_g + \frac{\hbar^2 k^2}{2m_c'}$	0	pk_+	$-i\sqrt{\frac{2}{3}}\,pk_z$	$\frac{1}{\sqrt{3}}\,pk_-$	0	$\frac{1}{\sqrt{3}}\,pk_z$	$-i\sqrt{\frac{2}{3}}\,pk_-$
1/2, −1/2	0	$E_g + \frac{\hbar^2 k^2}{2m_c'}$	0	$\frac{1}{\sqrt{3}}\,pk_+$	$\sqrt{\frac{2}{3}}\,pk_z$	ipk_-	$\sqrt{\frac{2}{3}}\,pk_+$	$\frac{i}{\sqrt{3}}\,pk_z$
3/2, 3/2	pk_-	0	F'	H'	I'	0	$\frac{iH'}{\sqrt{2}}$	$-i\sqrt{2}\,I'$
3/2, 1/2	$i\sqrt{\frac{2}{3}}\,pk_z$	$-\frac{i}{\sqrt{3}}\,pk_-$	$H^{*\prime}$	G'	0	I'	$\frac{i}{\sqrt{2}}(G'-F')$	$i\sqrt{\frac{3}{2}}\,H'$
3/2, −1/2	$\frac{1}{\sqrt{3}}\,pk_+$	$\sqrt{\frac{2}{3}}\,pk_z$	$I^{*\prime}$	0	G'	$-H'$	$-i\sqrt{\frac{3}{2}}\,H^{*\prime}$	$\frac{i}{\sqrt{2}}(G'-F')$
3/2, −3/2	0	$-ipk_+$	0	$I^{*\prime}$	$-H^{*\prime}$	F'	$-i\sqrt{2}\,I^{*\prime}$	$-\frac{i}{\sqrt{2}}\,H^{*\prime}$
1/2, 1/2	$\frac{1}{\sqrt{3}}\,pk_z$	$\sqrt{\frac{2}{3}}\,pk_-$	$-\frac{i}{\sqrt{2}}\,H^{*\prime}$	$-\frac{i}{\sqrt{2}}(G'-F')$	$i\sqrt{\frac{3}{2}}\,H'$	$i\sqrt{2}\,I'$	$-\Delta + \frac{F'+G'}{2}$	0
1/2, −1/2	$i\sqrt{\frac{2}{3}}\,pk_+$	$-\frac{i}{\sqrt{3}}\,pk_z$	$i\sqrt{2}\,I^{*\prime}$	$-i\sqrt{\frac{3}{2}}\,H^{*\prime}$	$-\frac{i}{\sqrt{2}}(G'-F')$	$\frac{i}{\sqrt{2}}\,H'$	0	$-\Delta + \frac{F'+G'}{2}$

§25. THEORY OF INVARIANTS

As demonstrated in the preceding sections, the Hamiltonian of the electron in the effective mass approximation is an operator $\mathcal{H}(\boldsymbol{k})$ in the form of an $n_s \times n_s$ matrix, where n_s is the dimension of the corresponding representation $\mathcal{D}^{k_s}$ of the little group G_{k_s}. For free carriers the eigenvalues of the matrix $\mathcal{H}(\boldsymbol{k})$ determine the spectrum $E(\boldsymbol{k})$. The construction of $\mathcal{H}(\boldsymbol{k})$ made use of perturbation theory, so that we were able to establish the general form of $\mathcal{H}(\boldsymbol{k})$ and to develop explicit expressions for the constants entering $\mathcal{H}$ in terms of matrix elements. However, it is usually sufficient to know only the general form of $\mathcal{H}$, since the various constants are determined as a rule from experimental data. To determine the general form of $\mathcal{H}(\boldsymbol{k})$ there is no need for the cumbersome apparatus of perturbation theory: all we need is the law of transformation of the basis functions $\psi_{k_s m}$ and of the components k_i under the operations in the little group.

Conditions imposed on $\mathcal{H}(\boldsymbol{k})$ by symmetry

Consider the Hamiltonian $\mathcal{H}$ relative to basis functions ψ_i $(i = 1, 2, \ldots, n_s)$ which transform according to an n_s-dimensional representation $\mathcal{D}(g)$ of a group G, $g \in G$. We shall assume that $\mathcal{H}$ depends not only on the wave vector $\boldsymbol{k}$, but also on external fields: a magnetic field $\boldsymbol{H}$ and an electric field $\mathcal{E}$, allowing $\mathcal{H}$ in the latter case to depend explicitly on the coordinate $\boldsymbol{x}$. In addition, as we shall show in detail in the next chapter, $\mathcal{H}$ may include terms proportional to the components of the strain tensor. All these quantities will be denoted by $\mathscr{K}$; it will be assumed that the components $\mathscr{K}_i$ transform under an operation g as dictated by the transformation laws of the corresponding vectors or the tensors. An element g of the group G takes the functions $\psi_j(\boldsymbol{x})$ into

$$\psi'_j(\boldsymbol{x}) = \psi_j(g^{-1}\boldsymbol{x}) = \sum_i \mathcal{D}_{ij}(g)\,\psi_i(\boldsymbol{x})$$

and it follows from (7.12) that $\mathcal{H}(\mathscr{K})$ goes into $\mathcal{D}^{-1}(g)\,\mathcal{H}(\mathscr{K})\,\mathcal{D}(g)$.

On the other hand, a transformation g of the coordinate system is equivalent to the inverse transformation g^{-1} of the vectors or tensors $\mathscr{K}$, so that $\mathcal{H}(\mathscr{K})$ goes into $\mathcal{H}(g^{-1}\mathscr{K})$. Consequently, $\mathcal{H}(\mathscr{K})$ must satisfy the conditions

$$\mathcal{D}^{-1}(g)\,\mathcal{H}(\mathscr{K})\,\mathcal{D}(g) = \mathcal{H}(g^{-1}\mathscr{K}), \qquad g \in G. \tag{25.1a}$$

These relations are the usual invariance conditions of the Hamiltonian under the transformation g:

$$\mathcal{H}'(\mathscr{K}') = \mathcal{D}(g)\,\mathcal{H}(g^{-1}\mathscr{K})\,\mathcal{D}^{-1}(g) = \mathcal{H}(\mathscr{K}). \tag{25.1b}$$

Each matrix equation (25.1) represents n_s^2 equations for the elements $\mathcal{H}_{ij}(\mathscr{K})$ of $\mathcal{H}(\mathscr{K})$.

Equations (25.1) must hold for all operations $g \in G$. However, it will suffice if they hold for all the generators of the group G, since any element g may be written as a product of powers of generators.

If we disregard the additional requirements imposed on $\mathcal{H}(\mathscr{K})$ by the time reversal symmetry of the Hamiltonian, which will be discussed below, the system of equations (25.1) for all the generators completely determines the form of $\mathcal{H}(\mathscr{K})$. If the representation matrices are known, we can construct $\mathcal{H}(\mathscr{K})$ directly from these equations. Though it may seem cumbersome, in practice this method turns out to be quite simple. However, as we demonstrate below, in most cases we can construct $\mathcal{H}(\mathscr{K})$ knowing only the characters of the representations. Of course, in this case too $\mathcal{H}(\mathscr{K})$ is constructed relative to a definite basis, but the basis is not specified.

The $n_s \times n_s$ matrix $\mathcal{H}(\mathscr{K})$ may be expressed in terms of products of the components $\mathscr{K}_i$ and n_s^2 linearly independent matrices $X_i (i = 1, 2, \ldots, n_s^2)$ which do not depend on $\mathscr{K}$. Let us consider the properties of the basis matrices X_i in greater detail. By (10.9), the effect of applying an operation g to any of these matrices is

$$g^{-1} X_i = X_i' = \mathscr{D}(g) X_i \mathscr{D}^{-1}(g), \tag{25.2}$$

and the transformed matrix X_i' may be expanded in the complete set of matrices X_i:

$$X_i' = \sum_j \mathscr{D}^X_{ji}(g) X_j. \tag{25.3}$$

The matrices $\mathscr{D}^X(g)$ form a representation of group G (this is easily proved by the arguments used in §7 to derive equation (7.5)). The dimension of the representation $\mathscr{D}^X$ is n_s^4. To determine its characters, we give the matrices X_i a double index, X^{lk} $(l = 1, 2, \ldots, n_s,\ k = 1, 2, \ldots, n_s)$, and choose them so that each matrix X^{lk} has only one nonzero element $X^{lk}_{lk} = 1$:

$$X^{lk}_{l'k'} = \delta_{ll'} \delta_{kk'}.$$

By (25.2) and (25.3), in this representation,

$$\mathscr{D}^X_{iji'j'}(g) = \mathscr{D}_{i'i}(g) \mathscr{D}^{-1}_{jj'}(g) = \mathscr{D}_{ii'}(g) \mathscr{D}^{*}_{j'j}(g),$$

i.e., each matrix of $\mathscr{D}^X$ is a direct product $\mathscr{D}(g)\mathscr{D}^{*}(g)$. Hence the character of the representation $\mathscr{D}^X$, which does not depend on the specific choice of the matrices X_i, is

$$\chi_X(g) = \sum_{ij} \mathscr{D}^X_{ijij}(g) = \chi(g)\chi^{*}(g) = |\chi(g)|^2, \tag{25.4}$$

where $\chi(g)$ is the character of the representation $\mathscr{D}(g)$.

The representation $\mathscr{D}^X$ is in general reducible and may be expanded in terms of its irreducible constituents. This means that from the n_s^2 linearly independent matrices X_i we can form linear combinations $X_l^{\varkappa}$ which, by (25.2) and (25.3), transform according to the irreducible representations $\mathscr{D}_{\varkappa}$ of the group G occurring in $\mathscr{D} \times \mathscr{D}^{*}$, and so satisfy the relations

$$g^{-1} X_l^{\varkappa} = \mathscr{D}(g) X_l^{\varkappa} \mathscr{D}^{-1}(g) = \sum_{l'} \mathscr{D}^{\varkappa}_{l'l}(g) X^{\varkappa}_{l'}. \tag{25.5}$$

In the same way, the vector or tensor components also yield combinations $\mathscr{K}_l^{\varkappa}$ which transform according to the representations $\mathscr{D}_{\varkappa}$ of G. Then, by the rules of §9 for constructing invariants (see (9.27)), the matrices $\mathcal{H}(\mathscr{K})$ may be expressed in terms of products of the matrices $X_l^{\varkappa}$ and the components $\mathscr{K}_l^{\varkappa *}$, which transform according to complex conjugate representations:

$$\mathscr{H}(\mathcal{K}) = \sum_{\varkappa} a_{\varkappa} \sum_{l} \mathbf{X}_l^{\varkappa} \mathcal{K}_l^{*\varkappa}, \tag{25.6}$$

where $a_{\varkappa}$ are arbitrary constants, which may be either complex, real or pure imaginary, so chosen that $\mathscr{H}(\mathcal{K})$ is hermitian. If the components $\mathcal{K}_l^{\varkappa}$ are real and the matrices $\mathbf{X}_l^{\varkappa}$ hermitian, these constants must be real. The expansion (25.6) may involve the same components $\mathcal{K}_l^{\varkappa}$ multiplied by several different sets of matrices $\mathbf{X}_l^{\varkappa}$ which transform according to one representation, and consequently with different constants $a_{\varkappa}$.

Using (25.5), we easily show that the operator $\mathscr{H}(\mathcal{K})$ defined by (25.6) indeed satisfies condition (25.1). In so doing, we must use the fact that by (7.3) the components $\mathcal{K}_l^{\varkappa *}$ transform according to

$$\mathcal{K}_l'^{\varkappa *} = g^{-1}\mathcal{K}_l^{\varkappa *} = \sum_{l'} \mathscr{D}_{l'l}^{\varkappa *}(g)\,\mathcal{K}_{l'}^{\varkappa *} = \sum_{l'} \mathscr{D}_{ll'}^{\varkappa^{-1}}(g)\,\mathcal{K}_{l'}^{\varkappa *}. \tag{25.7}$$

Consequently, relations (25.4)–(25.6) imply that $\mathscr{H}(\mathcal{K})$ may involve only components $\mathcal{K}_l^{\varkappa}$ that transform according to the constituents $\mathscr{D}_{\varkappa}$ of the direct product $\mathscr{D} \times \mathscr{D}^{*}$; these constituents are of course the representations for which formula (19.43) yields a nonzero $N_{\varkappa}$.

The last equation may also be derived by other means. Substituting (25.6) into (25.1b) and comparing coefficients of like components $\mathcal{K}_l^{\varkappa}$, we obtain

$$\mathbf{X}_l^{\varkappa} = \sum_{l'} \mathscr{D}_{l'l}^{\varkappa}(g)\,\mathscr{D}^{-1}(g)\,\mathbf{X}_{l'}^{\varkappa}\mathscr{D}(g)$$

or, in expanded form,

$$X_{l,\,ij}^{\varkappa} = \sum_{l'i'j'} \mathscr{D}_{l'l}^{\varkappa}(g)\,\mathscr{D}_{i'i}^{*}(g)\,\mathscr{D}_{j'j}(g)\,X_{l',\,i'j'}^{\varkappa}. \tag{25.8}$$

This sum is precisely the expression (19.4) defining the transformation law of the matrix elements V_{ij}^{l} under an operation g. Therefore, the number of sets of linearly independent matrices $\mathbf{X}_l^{\varkappa}$ which transform according to a given irreducible representation $\mathscr{D}_{\varkappa}$ of the point group F is just the number of times $\mathscr{D}_{\varkappa}$ occurs in the product $\mathscr{D}(g) \times \mathscr{D}^{*}(g)$, determined by (19.6).

At no point in the preceding arguments did we assume that the representation $\mathscr{D}(g)$ is irreducible; this was because it is frequently necessary to construct $\mathscr{H}(\mathcal{K})$ for several representations simultaneously. Let us consider the general case, in which the representation $\mathscr{D}(g)$ is reducible and decomposes, say, into two representations $\mathscr{D}_{\text{I}}$ and $\mathscr{D}_{\text{II}}$ of dimensions n_{I} and n_{II} with basis ψ_{I} and ψ_{II}:

$$\mathscr{D} = \begin{Vmatrix} \mathscr{D}_{\text{I}} & 0 \\ 0 & \mathscr{D}_{\text{II}} \end{Vmatrix}. \tag{25.9}$$

Relative to this basis, the matrix $\mathscr{H}(\mathcal{K})$ has the form

$$\mathscr{H}(\mathcal{K}) = \begin{Vmatrix} \mathscr{H}^{\text{I I}} & \mathscr{H}^{\text{I II}} \\ \mathscr{H}^{\text{II I}} & \mathscr{H}^{\text{II II}} \end{Vmatrix}. \tag{25.10}$$

Since $\mathscr{D}$ (25.9) is a block-diagonal matrix, invariance of the Hamiltonian (25.1) yields the following relations for each block of the matrix $\mathscr{H}(\mathcal{K})$:

$$\begin{gathered} \mathscr{D}_{\text{I}}^{-1}(g)\,\mathscr{H}^{\text{I I}}(\mathcal{K})\,\mathscr{D}_{\text{I}}(g) = \mathscr{H}^{\text{I I}}(g^{-1}\mathcal{K}), \quad \mathscr{D}_{\text{I}}^{-1}(g)\,\mathscr{H}^{\text{I II}}(\mathcal{K})\,\mathscr{D}_{\text{II}}(g) = \mathscr{H}^{\text{I II}}(g^{-1}\mathcal{K}), \\ \mathscr{D}_{\text{II}}^{-1}(g)\,\mathscr{H}^{\text{II II}}(\mathcal{K})\,\mathscr{D}_{\text{II}}(g) = \mathscr{H}^{\text{II II}}(g^{-1}\mathcal{K}), \quad \mathscr{D}_{\text{II}}^{-1}(g)\,\mathscr{H}^{\text{II I}}(\mathcal{K})\,\mathscr{D}_{\text{I}}(g) = \mathscr{H}^{\text{II I}}(g^{-1}\mathcal{K}); \end{gathered} \tag{25.11}$$

note that for the diagonal blocks these relations naturally coincide with (25.1). In this case each block of $\mathscr{H}$ is a linear combination of the corresponding basis matrices $\boldsymbol{X}_l^{\mathrm{I\,I}}$, $\boldsymbol{X}_l^{\mathrm{II\,II}}$, $\boldsymbol{X}_l^{\mathrm{I\,II}}$, $\boldsymbol{X}_l^{\mathrm{II\,I}}$. The basis matrices $\boldsymbol{X}^{\mathrm{I\,I}}$ and $\boldsymbol{X}^{\mathrm{II\,II}}$ are square matrices of orders $n_{\mathrm{I}} \times n_{\mathrm{I}}$ and $n_{\mathrm{II}} \times n_{\mathrm{II}}$, respectively. $X^{\mathrm{I\,II}}$ has n_{I} rows and n_{II} columns; conversely, $X^{\mathrm{II\,I}}$ has n_{II} rows and n_{I} columns. If $n_{\mathrm{I}} = n_{\mathrm{II}}$, all the basis matrices are square.

It is easy to see that both the diagonal and the off-diagonal matrices, taken separately, form a basis for a representation of the group G, with characters

$$\chi_{\mathrm{I\,I}}(g) = |\chi_{\mathrm{I}}(g)|^2, \quad \chi_{\mathrm{II\,II}}(g) = |\chi_{\mathrm{II}}(g)|^2, \tag{25.12a}$$

$$\chi_{\mathrm{I\,II}}(g) = \chi_{\mathrm{I}}^{*}(g)\,\chi_{\mathrm{II}}(g), \quad \chi_{\mathrm{II\,I}}(g) = \chi_{\mathrm{I}}(g)\,\chi_{\mathrm{II}}^{*}(g) = \chi_{\mathrm{I\,II}}^{*}(g). \tag{25.12b}$$

Let us specialize these general considerations to construction of the Hamiltonian $\mathscr{H}_{\boldsymbol{k}_0}(\mathscr{K})$ at the point $\boldsymbol{k}_0$. Here G is the little group $G_{\boldsymbol{k}_0}$, and $\mathscr{D} = \mathscr{D}_{\mu}^{k_0}$. By (25.4), the character $\chi_{\boldsymbol{k}_0}^{X}$ of $\mathscr{D}^{X}$ at $\boldsymbol{k}_0$ is

$$\chi_X^{\boldsymbol{k}_0} = |\chi_{\mu}^{\boldsymbol{k}_0}(g)|^2, \tag{25.13}$$

where $\chi_{\mu}^{\boldsymbol{k}_0}(g)$ is the character of $\mathscr{D}_{\mu}^{\boldsymbol{k}_0}$.

It follows from (25.5) that $\mathscr{D}^{X}$ is independent of translations by the Bravais lattice periods and is thus determined only by the rotational elements r of the point group $F_{\boldsymbol{k}_0}$, i.e., $\mathscr{D}^{X}$ is always a vector representation of $F_{\boldsymbol{k}_0}$ and may be expanded in vector representations of the point group $F_{\boldsymbol{k}_0}$. Hence the basis matrices X always transform according to vector representations of $F_{\boldsymbol{k}_0}$. Consequently, by equation (19.43), which follows from (19.6), the number of sets $X^{\varkappa}$ which transform according to a representation $\mathscr{D}_{\varkappa}$ of the point group $F_{\boldsymbol{k}_0}$ is equal to the number of times $\mathscr{D}_{\varkappa}$ is contained in the product $\mathscr{D}_{\mu}^{*\boldsymbol{k}_0} \times \mathscr{D}_{\mu}^{\boldsymbol{k}_0}$:

$$N_{\varkappa} = \frac{1}{h} \sum_{g \in G'_{\boldsymbol{k}_0}} |\chi_{\mu}^{\boldsymbol{k}_0}(g)|^2 \chi^{\varkappa}(g). \tag{25.14}$$

Here the summation extends over all the fundamental elements of the little group, which contain no primitive translations. It is easy to see that formulas (25.1), (25.4), (25.5), (25.11) and (25.14) remain valid if, using (12.26), we go over from representations of the little group $\mathscr{D}_{\mu}^{\boldsymbol{k}_0}$ to projective representations of the point group, $\mathscr{D}_{\mu}^{\boldsymbol{k}_0}(r) = \mathscr{D}_{\mu}^{\boldsymbol{k}_0}(g)\, e^{i\boldsymbol{k}_0 \boldsymbol{\tau}}$. Therefore, the last equation may be rewritten as

$$N_{\varkappa} = \frac{1}{h} \sum_{r \in F_{\boldsymbol{k}_0}} |\chi_{\mu}^{\boldsymbol{k}_0}(r)|^2 \chi^{\varkappa}(r). \tag{25.14a}$$

Here the summation extends over all the elements of the crystal point group $F_{\boldsymbol{k}_0}$, and $\chi_{\mu}^{\boldsymbol{k}_0}(r)$ runs through the characters of the corresponding projective representations.

For the case of two combined representations $\mathscr{D}_{\mu}$ and $\mathscr{D}_{\nu}$, it follows from (25.12a) that the number $N_{\varkappa}$ for the diagonal submatrices X^{II} and $X^{\mathrm{II\,II}}$ is defined by a formula similar to (25.14); by (19.31), the corresponding formula for the off-diagonal submatrices $X^{\mathrm{I\,II}}$ is

$$N_{\varkappa} = \frac{1}{h} \sum_{g \in G'_{\boldsymbol{k}_0}} \chi_{\mu}^{*\boldsymbol{k}_0}(g)\, \chi_{\nu}^{\boldsymbol{k}_0}(g)\, \chi^{\varkappa}(g). \tag{25.15}$$

Hitherto we have considered the conditions imposed on $\mathcal{H}(\mathscr{K})$ by symmetry alone.

Additional conditions imposed on $\mathcal{H}(\mathscr{K})$ by time reversal symmetry of the Hamiltonian

As mentioned in §19, equation (19.43) is applicable when the points $\boldsymbol{k}_0$ and $-\boldsymbol{k}_0$ are in different stars (case (b_3)). In all other cases we must take into account additional conditions on $\mathcal{H}(\mathscr{K})$, due to time reversal symmetry.

Now the time reversal operation K takes $\boldsymbol{k}_0$ into $-\boldsymbol{k}_0$; therefore, whenever the points $\boldsymbol{k}_0$ and $-\boldsymbol{k}_0$ are not equivalent (cases (a_2), (b_2), and (c_2)) we shall first assume that the operator $\mathcal{H}(\mathscr{K})$ is constructed simultaneously for both points $\boldsymbol{k}_0$ and $-\boldsymbol{k}_0$, so that the basis functions are the functions ψ, i.e., $\psi_{\boldsymbol{k}_0}$, $\psi_{-\boldsymbol{k}_0}=R\psi_{\boldsymbol{k}_0}$, and in cases (b_2) and (c_2) also the functions $K\psi$, i.e., $K\psi_{\boldsymbol{k}_0}$ and $KR\psi_{\boldsymbol{k}_0}$, where R is any operation taking $\boldsymbol{k}_0$ into $-\boldsymbol{k}_0$.

We first consider case (a), when $\mathcal{H}(\mathscr{K})$ is constructed relative to basis functions ψ_i such that ψ_i and $K\psi_i$ obey a linear relation (18.7):

$$K\psi_i=\sum_l T_{li}\psi_l.$$

Upon transformation from the basis ψ to the basis $K\psi$, the matrix $\mathcal{H}(\mathscr{K})$ goes into $\boldsymbol{T}^{-1}\mathcal{H}(\mathscr{K})\boldsymbol{T}$.

As shown in §18, the variables $\mathscr{K}_l$ may always be classified as even or odd with respect to time reversal, so that $K\mathscr{K}_l=f\mathscr{K}_l$, $f=\pm 1$. Thus, for the components of the wave vector and the magnetic field, which are odd with respect to time reversal, $f=-1$, while for the electric field, space coordinates and components of the strain tensor, $f=1$.

The invariance of $\mathcal{H}(\mathscr{K})$ under time reversal means that time reversal takes $\mathcal{H}(\mathscr{K})$ into $\mathcal{H}^*(\mathscr{K})$; thus, in case (a), the matrix $\mathcal{H}(\mathscr{K})$ must satisfy the additional condition:

$$\boldsymbol{T}^{-1}\mathcal{H}(f\mathscr{K})\boldsymbol{T}=\mathcal{H}^*(\mathscr{K})=\tilde{\mathcal{H}}(\mathscr{K}). \tag{25.16}$$

Equation (25.16) shows that the basis matrices $\boldsymbol{X}_l^{\varkappa}$ must satisfy the relations

$$K\boldsymbol{X}_l^{\varkappa}=\boldsymbol{T}^{-1}\boldsymbol{X}_l^{\varkappa}\boldsymbol{T}=f\boldsymbol{X}_l^{\varkappa *}=f\tilde{\boldsymbol{X}}_l^{\varkappa}. \tag{25.17}$$

It is assumed here that all the components $\mathscr{K}$ are real, as are the constants $a_{\varkappa}$, and so hermiticity of the operator $\mathcal{H}$ implies that the matrices $\boldsymbol{X}$ are also hermitian.

Thus, in case (a) all the basis matrices $\boldsymbol{X}_l^{\varkappa}$ can be classified as even or odd with respect to time reversal, in the sense of (25.17), matrices of the same parity transforming independently according to the representation $\mathscr{D}_{\varkappa}$. Any given matrix $\boldsymbol{X}_l^{\varkappa}$ can be decomposed into even and odd matrices, defining matrices $\boldsymbol{X}_{l\,1,2}^{\varkappa}$ by

$$\boldsymbol{X}_{l\,1,2}^{\varkappa}=\frac{1}{2}(\boldsymbol{X}_l^{\varkappa}\pm K^2\boldsymbol{T}\tilde{\boldsymbol{X}}_l^{\varkappa}\boldsymbol{T}^{-1}). \tag{25.18}$$

Indeed,

$$\begin{aligned}K\boldsymbol{X}_{l\,1,2}^{\varkappa}&=\frac{1}{2}(\boldsymbol{T}^{-1}\boldsymbol{X}_l^{\varkappa}\boldsymbol{T}\pm K^2\tilde{\boldsymbol{X}}_l^{\varkappa})=\pm\frac{K^2}{2}(\tilde{\boldsymbol{X}}_l^{\varkappa}\pm K^2\boldsymbol{T}^{-1}\boldsymbol{X}_l^{\varkappa}\boldsymbol{T})=\\&=\pm\frac{K^2}{2}\widetilde{(\boldsymbol{X}_l^{\varkappa}\pm K^2\boldsymbol{T}\tilde{\boldsymbol{X}}_l^{\varkappa}\boldsymbol{T}^{-1})}=\pm K^2\tilde{\boldsymbol{X}}_{l\,1,2}^{\varkappa}.\end{aligned}$$

Here we have used the unitarity of the matrix $\boldsymbol{T}$ and the fact that, by (18.26), $\tilde{\boldsymbol{T}} = K^2\boldsymbol{T}$. Thus, if $K^2 = 1$, the matrices $\boldsymbol{X}^{\varkappa}_{l1}$ are even and the $\boldsymbol{X}^{\varkappa}_{l2}$ odd with respect to time reversal (if $K^2 = -1$, conversely, the $\boldsymbol{X}^{\varkappa}_{l1}$ are odd and the $\boldsymbol{X}^{\varkappa}_{l2}$ even). Condition (25.16) implies that $\mathscr{H}(\mathscr{K})$ includes products of matrices $\boldsymbol{X}^{\varkappa}_{l}$ and components $\mathscr{K}^{\varkappa}_{l}$ of the same parity with respect to time reversal.

In expanded form, equation (25.17) is

$$X^{\varkappa}_{l,\,ij} = f \sum_{i'j'} T^{*}_{i'i} X^{\varkappa}_{l,\,i'j'} T_{j'i} \tag{25.19}$$

This equation coincides with (19.8). As shown in §19, the formula (19.13) for $N_{\varkappa}$ follows from (19.8) and (19.4), which is similar to (25.8). It should be borne in mind, however, that since we are constructing $\mathscr{H}(\mathscr{K})$ for only one point of the star, the relations (25.8) are required to hold only for the elements of the little group $G_{\boldsymbol{k}_0}$, whereas equations (19.4) are valid for all operations in the space group G. Now it would be possible to construct $\mathscr{H}(\mathscr{K})$ for all points of the star rather than for just one point, i.e., for a complete representation of the space group. In this representation $\mathscr{H}(\mathscr{K})$ would be in block-diagonal form with N matrices of order $n_\mu \times n_\mu$ along the diagonal, where N is the number of points in the star and n_μ the dimension of the representation $\mathscr{D}^{\boldsymbol{k}_0}_{\mu}$. Conditions (25.1) and the resulting relation (25.8) would then define the form of $\mathscr{H}_{\boldsymbol{k}_0}(\mathscr{K})$ if $g \in G_{\boldsymbol{k}_0}$ and of the operator $\mathscr{H}_{\boldsymbol{k}}(\mathscr{K})$ for another point of the star if $g \notin G_{\boldsymbol{k}_0}$. Thus equation (25.8) is entirely equivalent to (19.4) and, consequently, the number of linearly independent even and odd sets of matrices $X^{\varkappa}$ which transform according to the representation $\mathscr{D}_{\varkappa}$ is given by (19.45) in case (a_1) and by (19.46) in (a_2).

Relations (25.16), (25.17), (25.18) are directly applicable to case (a_1) when there is a linear relation between the functions $\psi_{\boldsymbol{k}_0}$ and $K\psi_{\boldsymbol{k}_0}$. In case (a_2), as indicated above, they are valid for a matrix $\mathscr{H}(\mathscr{K})$ constructed simultaneously for the points $\boldsymbol{k}_0$ and $-\boldsymbol{k}_0$, i.e., relative to the basis functions $\psi_{i\boldsymbol{k}_0}$ and $R\psi_{i\boldsymbol{k}_0} = \psi_{i,-\boldsymbol{k}_0}$. Relative to this basis $\mathscr{H}(\mathscr{K})$ is

$$\mathscr{H}(\mathscr{K}) = \begin{Vmatrix} \mathscr{H}_{\boldsymbol{k}_0}(\mathscr{K}) & 0 \\ 0 & \mathscr{H}_{-\boldsymbol{k}_0}(\mathscr{K}) \end{Vmatrix}. \tag{25.20}$$

In this case we have a linear relation between the functions $K\psi_{i\boldsymbol{k}_0}$ and $R\psi_{i\boldsymbol{k}_0}$: $K\psi_{i\boldsymbol{k}_0} = \boldsymbol{T}R\psi_{i\boldsymbol{k}_0}$. Under the transformation R, the transformation matrix relative to this basis is

$$\mathscr{D}(R) = \begin{Vmatrix} 0 & \mathscr{D}^{\boldsymbol{k}_0}_{\mu}(R^2) \\ \boldsymbol{I} & 0 \end{Vmatrix}. \tag{25.21}$$

Substituting (25.20) and (25.21) into (25.1), we establish a relation between the blocks $\mathscr{H}_{\boldsymbol{k}_0}(\mathscr{K})$ and $\mathscr{H}_{-\boldsymbol{k}_0}(\mathscr{K})$ in (25.20) which is a consequence of the crystal symmetry:

$$\mathscr{H}_{-\boldsymbol{k}_0}(\mathscr{K}) = \mathscr{H}_{\boldsymbol{k}_0}(R^{-1}\mathscr{K}). \tag{25.22}$$

Under the transformation K, the set of functions $\psi_{i\boldsymbol{k}_0}$, $R\psi_{i\boldsymbol{k}_0}$ goes into $K\psi_{i\boldsymbol{k}_0}$, $KR\psi_{i\boldsymbol{k}_0} = \boldsymbol{T}R^2\psi_{i\boldsymbol{k}_0}$, so that the transformation matrix is

$$\boldsymbol{T}_1 = \begin{Vmatrix} 0 & \boldsymbol{T} \\ \boldsymbol{T}\mathscr{D}^{\boldsymbol{k}_0}_{\mu}(R^2) & 0 \end{Vmatrix}. \tag{25.23}$$

Substituting (25.23) and (25.20) into (25.16), we obtain the following relation between $\mathscr{H}_{\boldsymbol{k}_0}(\mathscr{K})$ and $\mathscr{H}_{-\boldsymbol{k}_0}(\mathscr{K})$:

$$\boldsymbol{T}^{-1}\mathscr{H}_{-\boldsymbol{k}_0}(\mathfrak{f}\mathscr{K})\boldsymbol{T}=\mathscr{H}^{*}_{\boldsymbol{k}_0}(\mathscr{K})=\tilde{\mathscr{H}}_{\boldsymbol{k}}\,(\mathscr{K}). \tag{25.24}$$

Equations (25.24) and (25.22) imply the condition

$$\boldsymbol{T}^{-1}\mathscr{H}_{\boldsymbol{k}_0}(R^{-1}\mathscr{K})\boldsymbol{T}=\mathscr{H}^{*}_{\boldsymbol{k}_0}(\mathfrak{f}\mathscr{K})=\tilde{\mathscr{H}}_{\boldsymbol{k}}\,(\mathfrak{f}\mathscr{K}), \tag{25.25}$$

which must be satisfied by $\mathscr{H}_{\boldsymbol{k}_0}(\mathscr{K})$ in case (a$_2$). In other words, in this case the matrices $\boldsymbol{X}_l^{\varkappa}$ must satisfy the relation

$$K\boldsymbol{X}_l^{\varkappa}=\boldsymbol{T}^{-1}\boldsymbol{X}_l^{\varkappa}\boldsymbol{T}=\mathfrak{f}\sum_{l'}\mathscr{D}^{\varkappa}_{l'l}(R)\,\tilde{\boldsymbol{X}}^{\varkappa}_{l'}, \tag{25.26}$$

where $\mathscr{D}_{\varkappa}(R)$ is the transformation matrix of the components $\mathscr{K}_l^{\varkappa}$ under the operation R.

Using (25.26) and (25.8), we can obtain the formula (19.46) for $N_{\varkappa}$ which, as shown above, determines the number of sets of even and odd matrices $\boldsymbol{X}^{\varkappa}$ which transform according to $\mathscr{D}_{\varkappa}$. The actual construction of the even and the odd matrices for $\mathscr{D}^{\boldsymbol{k}_0}$ in this case makes use of (25.26).

As noted in §§18 and 19, in cases (b) and (c) time reversal causes the representations $\mathscr{D}_{\mu}^{\boldsymbol{k}_0}$ and $\mathscr{D}_{\mu}^{*\boldsymbol{k}_0}$ to combine. Relative to the combined basis, the matrix $\mathscr{H}(\mathscr{K})$ has the form (25.10). Time reversal takes the corresponding basis ψ_i, $K\psi_i$ into itself, via the matrix $\boldsymbol{T}_2$

$$\boldsymbol{T}_2=\left\|\begin{matrix}0 & K^2\boldsymbol{I}\\ \boldsymbol{I} & 0\end{matrix}\right\|, \tag{25.27}$$

where I is the identity matrix of order n_{μ}. Relative to this complete basis, $\mathscr{H}(\mathscr{K})$ must satisfy condition (25.16) with $\boldsymbol{T}=\boldsymbol{T}_2$; hence the following relation for the diagonal and off-diagonal blocks:

$$\mathscr{H}_{\mathrm{II\,II}}(\mathscr{K})=\mathscr{H}^{*}_{\mathrm{I\,I}}(\mathfrak{f}\mathscr{K}), \tag{25.28a}$$

$$\mathscr{H}_{\mathrm{II\,I}}(\mathscr{K})=K^2\mathscr{H}^{*}_{\mathrm{I\,II}}(\mathfrak{f}\mathscr{K})=K^2\tilde{\mathscr{H}}_{\mathrm{II\,I}}(\mathfrak{f}\mathscr{K}). \tag{25.28b}$$

In (25.28b) we have used the hermiticity of the Hamiltonian $\mathscr{H}(\mathscr{K})$.

Thus, conditions (25.28) do not impose additional restrictions on $\mathscr{H}_{\mathrm{I\,I}}(\mathscr{K})$ but only establish the relation between $\mathscr{H}_{\mathrm{II\,II}}(\mathscr{K})$ and $\mathscr{H}_{\mathrm{I\,I}}(\mathscr{K})$. Therefore, equation (25.14) defines the numbers $N_{\varkappa}$ for the diagonal blocks $\mathscr{H}_{\mathrm{I\,I}}(\mathscr{K})$ and $\mathscr{H}_{\mathrm{II\,II}}(\mathscr{K})$. By (25.28b), the off-diagonal blocks $\boldsymbol{X}_l^{\varkappa\,\mathrm{I\,II}}$ must satisfy the condition

$$\boldsymbol{X}_l^{\varkappa\,\mathrm{I\,II}}=\mathfrak{f}K^2\tilde{\boldsymbol{X}}_l^{\varkappa\,\mathrm{I\,II}}. \tag{25.29}$$

Consequently, as in (25.18), the off-diagonal blocks can be classified as even and odd, setting

$$\boldsymbol{X}_{l\;1,2}^{\mathrm{I\,II}}=\frac{1}{2}\left(\boldsymbol{X}_{l\;1,2}^{\mathrm{I\,II}}\pm K^2\tilde{\boldsymbol{X}}_{l\;1,2}^{\mathrm{I\,II}}\right). \tag{25.30}$$

In its expanded form, equation (25.29) coincides with (19.10). In other words, the number of linearly independent even and odd off-diagonal blocks in cases (b$_1$), (b$_2$), and (c) is also given by (19.13), which in cases (b$_1$) and (c$_1$) reduces to (19.45) and in cases (b$_2$) and (c$_2$) to (19.44). As in case (a), the relations (25.27)–(25.30) apply directly to cases (b$_1$) and (c$_1$). In cases (b$_2$) and (c$_2$) they are valid for the matrix $\mathscr{H}(\mathscr{K})$ relative to the basis $\psi_{i\boldsymbol{k}_0}$, $KR\psi_{i\boldsymbol{k}_0}$, $R\psi_{i\boldsymbol{k}_0}$, $KR^2\psi_{i\boldsymbol{k}_0}$. In this basis, the first pair of functions belongs to the point $\boldsymbol{k}_0$ and forms a basis for $\mathscr{H}_{\boldsymbol{k}_0}$, the second belongs to $-\boldsymbol{k}_0$ and forms a basis for $\mathscr{H}_{-\boldsymbol{k}_0}$, so that $\mathscr{H}(\mathscr{K})$ has the form (25.20) and, as is easily seen, satisfies condition (25.22).

Under time reversal K, the above basis goes into itself via the matrix

$$\boldsymbol{T}_2' = \begin{Vmatrix} 0 & \boldsymbol{T} \\ K^2\tilde{\boldsymbol{T}} & 0 \end{Vmatrix}, \qquad \text{where} \qquad \boldsymbol{T} = \begin{Vmatrix} 0 & K^2\mathscr{D}_\mu^{\boldsymbol{k}_0}(R^2) \\ \boldsymbol{I} & 0 \end{Vmatrix}. \tag{25.31}$$

Substituting (25.31) and (25.20) into the invariance condition (25.16), we obtain the following relations for the diagonal and off-diagonal blocks of $\mathscr{H}_{\boldsymbol{k}_0}(\mathscr{K})$ in cases (b_2) and (c_2):

$$\mathscr{H}_{\boldsymbol{k}_0}^{\mathrm{I\,I}}(\mathscr{K}) = \mathscr{H}_{\boldsymbol{k}_0}^{*\mathrm{II\,II}}(fR\mathscr{K}) = \tilde{\mathscr{H}}^{\mathrm{II\,II}}(fR\mathscr{K}), \tag{25.32a}$$

$$\mathscr{H}_{\boldsymbol{k}_0}^{\mathrm{I\,II}}(\mathscr{K}) = K^2\mathscr{H}_{\boldsymbol{k}_0}^{*\mathrm{II\,I}}(fR\mathscr{K})\,\mathscr{D}_\mu^{*\boldsymbol{k}_0}(R^2) = K^2\tilde{\mathscr{H}}_{\boldsymbol{k}_0}^{\mathrm{I\,II}}(fR\mathscr{K})\,\mathscr{D}_\mu^{*\boldsymbol{k}_0}(R^2). \tag{25.32b}$$

Since in this case $\mathscr{D}_{\mathrm{II}}(g) = \mathscr{D}_{\mathrm{I}}^{*}(R^{-1}gR)$, it follows from (25.11) that the off-diagonal blocks $\mathscr{H}^{\mathrm{I\,II}}$ must also satisfy the relation

$$\mathscr{D}^{-1}(g)\,\mathscr{H}_{\boldsymbol{k}_0}^{\mathrm{I\,II}}(\mathscr{K})\,\mathscr{D}^{*}(R^{-1}gR) = \mathscr{D}_{\boldsymbol{k}_0}^{\mathrm{I\,II}}(g^{-1}\mathscr{K}). \tag{25.33}$$

Condition (25.32a) establishes the connection between $\mathscr{H}_{\boldsymbol{k}_0}^{\mathrm{I\,I}}(\mathscr{K})$ and $\mathscr{H}_{\boldsymbol{k}_0}^{\mathrm{II\,II}}(\mathscr{K})$ in cases (b_2) and (c_2). Condition (25.32b) imposes additional requirements on the off-diagonal blocks $\boldsymbol{X}^{\mathrm{I\,II}}$, which have the form

$$\boldsymbol{X}_l^{\mathrm{I\,II}} = K^2 f \sum_{l'} \tilde{\boldsymbol{X}}_{l'}^{\mathrm{I\,II}}\,\mathscr{D}_{l'l}(R^{-1})\,\mathscr{D}_\mu^{*\boldsymbol{k}_0}(R^2), \tag{25.34}$$

where $\mathscr{D}_{l'l}(R^{-1})$ is the transformation matrix for the components $\mathscr{K}_l$ (25.7):

$$R\mathscr{K}_l = \sum_{l'} \mathscr{D}_{l'l}(R^{-1})\,\mathscr{K}_{l'}.$$

In view of (25.11), these conditions yield the formula (19.44) which, as shown previously, defines the number $N_\varkappa$ for the off-diagonal blocks $\boldsymbol{X}_l^\varkappa$ in cases (b_2) and (c_2).

Thus, the number of linearly independent sets of matrices $\boldsymbol{X}_l^\varkappa$ which transform according to $\mathscr{D}_\varkappa$ is defined in case (b_3) by the general equation (25.14). The same formula defines the number of diagonal submatrices $\boldsymbol{X}^{\mathrm{I\,I}}$ and $\boldsymbol{X}^{\mathrm{II\,II}}$ in cases (b) and (c). In these cases the only result of time reversal is a relation between $\mathscr{H}_{\boldsymbol{k}_0}(\mathscr{K})$ and $\mathscr{H}_{-\boldsymbol{k}_0}(\mathscr{K})$:

$$\mathscr{H}_{-\boldsymbol{k}_0}(\mathscr{K}) = \mathscr{H}_{\boldsymbol{k}_0}(f\mathscr{K}). \tag{25.35}$$

In case (a_1) the number $N_\varkappa$ is given by (19.45),

$$N_\varkappa = \frac{1}{2h} \sum_{g \in G'_{\boldsymbol{k}_0}} \chi_\varkappa(g)\left[(\chi_\nu^{\boldsymbol{k}_0}(g))^2 + K^2 f \chi_\nu^{\boldsymbol{k}_0}(g^2)\right], \tag{25.36}$$

and in case (a_2) by (19.46):

$$N_\varkappa = \frac{1}{2h} \sum_{g \in G'_{\boldsymbol{k}_0}} \left\{\chi_\varkappa(g)\left|(\chi_\nu^{\boldsymbol{k}_0}(g)\right|^2 + K^2 f \chi_\varkappa(gR)\,\chi_\nu^{\boldsymbol{k}_0}[(gR)^2]\right\}. \tag{25.37}$$

For the off-diagonal components $\mathscr{H}^{\mathrm{I\,II}} = \tilde{\mathscr{H}}^{\mathrm{II\,I}}$ the number $N_\varkappa$ is defined in cases (b_1) and (c_1) by (25.36) and in cases (b_2) and (c_2) by (19.44):

$$N_\varkappa = \frac{1}{2h} \sum_{g \in G'_{\boldsymbol{k}_0}} \left\{\chi_\varkappa(g)\,\chi_\nu^{\boldsymbol{k}_0}(g)\,\chi_\nu^{\boldsymbol{k}_0}(R^{-1}gR) + K^2 f \chi_\varkappa(gR)\,\chi_\nu^{\boldsymbol{k}_0}[(gR)^2]\right\}, \tag{25.38}$$

where R is any element which transforms $\boldsymbol{k}_0$ into $-\boldsymbol{k}_0$. Since the functions $\mathscr{K}$ must transform according to the same representations as $\boldsymbol{X}^\varkappa$, equations

(25.14) and (25.30)–(25.38) tell us which functions $f(\mathcal{K})$ may appear in $\mathcal{H}(\mathcal{K})$ for a given symmetry, i.e., for a specific little group.

The basis matrices $X_l^{\varkappa}$ in the expansion of $\mathcal{H}(\mathcal{K})$ may be determined, in principle, from (25.8), (25.19), (25.26), provided, of course, that the representations $\mathcal{D}(g)$ or the projective representations $\mathcal{D}(r)$ are known. In case conjugate representations are combined, these matrices are determined by (25.29) and (25.34). When this is done, it is sufficient to use these relations only for the generators of the point group, subsequently determining only "generating" matrices, for example, matrices which transform like x, y, z; suitable products of these matrices then yield all the other n_s^4 matrices $X_l^{\varkappa}$. If the matrices which transform as x, y, z vanish for some representation $\mathcal{D}$, a convenient choice of generating matrices are matrices which transform like the antisymmetrized products $\{x_i x_k\}$, i.e., like the components of an axial vector J_i; if these also vanish, we choose matrices which transform like the symmetrized products $[x_i x_k]$.

As we show in the following section, in most cases the basis matrices may be constructed without explicit knowledge of the matrices $\mathcal{D}(g)$ or $\mathcal{D}(r)$; all we need are the characters of the corresponding representations. Of course, the form of $\mathcal{H}(\mathcal{K})$ depends on the choice of representations, but when the methods of § 26 are used the form of the representation is not explicitly specified.

§ 26. METHODS FOR CONSTRUCTION OF BASIS MATRICES. APPLICATION OF THEORY OF INVARIANTS TO DETERMINE ELECTRON SPECTRA IN CUBIC CRYSTALS

In the present section we examine in detail various methods for constructing basis matrices in the theory of invariants, which will then be used to calculate the spectra of crystals with space groups O_h^7 and T_d^2. Other applications of the theory will be demonstrated in § 31, where the spectrum of a crystal with a wurtzite lattice and its change under the effect of strain will be discussed.

Twofold degeneracy

It is obvious that if the representation $\mathcal{D}(g)$ is one-dimensional, the only components of $\mathcal{K}$ that may appear in $\mathcal{H}(\mathcal{K})$ are those that transform according to the identity representation, and to determine the shape of the spectrum we need only find these components, i.e., determine the components for which $N_1 \neq 0$, using equation (25.14), (25.36) or (25.37). This conclusion is naturally in agreement with the results obtained in § 21 by kp-theory.

In the case of a two-dimensional representation, it is again sufficient to know the characters of the representation in order to construct $\mathcal{H}(\mathcal{K})$. To verify this, consider the condition imposed on the determinant $|\mathcal{H}(\mathcal{K}) - IE|$ by equation (25.1):

$$|\mathcal{H}(\mathcal{K}) - IE| = |\mathcal{H}(g\mathcal{K}) - IE|. \tag{26.1}$$

Using (26.1), we can determine the four components $\mathscr{H}_i$ which can appear in $\boldsymbol{\mathscr{H}}(\mathscr{K})$ (or four sums of inherently different but identically transforming components, such as k_x^2 and ε_{xx}, each of which may have different coefficients a). Denote the component which transforms according to the identity representation by $\mathscr{H}_0$ and the other three by $\mathscr{H}_i$. The direct product $\mathscr{D}\times\mathscr{D}^*$ may contain, apart from the identity representation, either a three-dimensional representation, two- and one-dimensional representations, or three one-dimensional representations. Since $\mathscr{D}\times\mathscr{D}^*$ has a real character, it follows that in the first two cases all the representations are real, while in the last, two of the three one-dimensional representations may be complex conjugates. We therefore assume that in the operator

$$\boldsymbol{\mathscr{H}}(\mathscr{K})=a_0\boldsymbol{X}_0\mathscr{H}_0+\sum_{i\varkappa}a_\varkappa\boldsymbol{X}_i^\varkappa\mathscr{H}_i^{\varkappa^*} \qquad (26.2)$$

the functions $\mathscr{H}_0$ and $\mathscr{H}_1$ are always real, while $\mathscr{H}_2$ and $\mathscr{H}_3$ are either real or complex conjugates.

The matrices $\boldsymbol{X}_i$ are linear combinations of $\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y, \boldsymbol{\sigma}_z$ and do not include the identity matrix, since if the $\boldsymbol{X}_i$ had nonzero trace the determinant (26.1) would contain a noninvariant term $E\mathscr{H}_i$. Via a suitable unitary transformation we can always make a hermitian matrix $\boldsymbol{X}_1$ with zero trace equal to $\boldsymbol{\sigma}_z$ (up to a constant) and then the matrices $\boldsymbol{X}_2$ and $\boldsymbol{X}_3$ will not include $\boldsymbol{\sigma}_z$, since otherwise the determinant (26.1) would contain a noninvariant product $\mathscr{H}_1\mathscr{H}_2$ or $\mathscr{H}_1\mathscr{H}_3$, whereas by (9.27) if $\mathscr{H}_1$ is real an invariant may include only $\mathscr{H}_1^2$. Next, applying a unitary transformation $\boldsymbol{S}=(1+\alpha^2)^{-1/2}(\boldsymbol{I}+\alpha\boldsymbol{\sigma}_z)$, which does not alter $\boldsymbol{X}_1$, we reduce the matrices $\boldsymbol{X}_2$ and $\boldsymbol{X}_3$ to the form $\boldsymbol{\sigma}_x$ and $\boldsymbol{\sigma}_y$ if $\mathscr{H}_2$ and $\mathscr{H}_3$ are real, or $\boldsymbol{\sigma}_+=\frac{1}{2}(\boldsymbol{\sigma}_x+i\boldsymbol{\sigma}_y)$ and $\boldsymbol{\sigma}_-=\frac{1}{2}(\boldsymbol{\sigma}_x-i\boldsymbol{\sigma}_y)$ if $\mathscr{H}_2=\mathscr{H}_3^*$. Indeed, it follows from (26.1) that if $\boldsymbol{X}_1=\boldsymbol{\sigma}_z$ and $\boldsymbol{X}_2=\boldsymbol{\sigma}_x$, then $\boldsymbol{X}_3$ must be proportional to $\boldsymbol{\sigma}_y$. Consequently, for two-dimensional representations $\mathscr{D}$ the matrices $\boldsymbol{X}_i$ in (26.2) which transform according to the irreducible constituents of $\mathscr{D}\times\mathscr{D}^*$ may always be taken as

$$\boldsymbol{X}_0=\boldsymbol{I},\ \boldsymbol{X}_1=\boldsymbol{\sigma}_z,\ \boldsymbol{X}_2=\boldsymbol{\sigma}_x,\ \boldsymbol{X}_3=\boldsymbol{\sigma}_y \quad \text{if} \quad \mathscr{H}_0,\ \mathscr{H}_1,\ \mathscr{H}_2,\ \mathscr{H}_3 \ \text{are real},$$

$$\boldsymbol{X}_2=\boldsymbol{\sigma}_+=\begin{Vmatrix}0 & 1\\ 0 & 0\end{Vmatrix},\ \boldsymbol{X}_3=\boldsymbol{\sigma}_-=\begin{Vmatrix}0 & 0\\ 1 & 0\end{Vmatrix} \quad \text{if} \quad \mathscr{H}_2=\mathscr{H}_3^*, \qquad (26.3)$$

regardless of the specific form of the representation $\mathscr{D}$.

As for the constants $a_\varkappa$ appearing in (26.2), it is evident from the general equation (25.6) that if $\mathscr{D}\times\mathscr{D}^*$ contains one three-dimensional representation (apart from the identity representation) then $\boldsymbol{\mathscr{H}}(\mathscr{K})$ contains only one real constant other than a_0. If $\mathscr{D}\times\mathscr{D}^*$ includes two-dimensional and one-dimensional representations the number of constants increases to three, all real. When all the representations $\mathscr{D}^\varkappa$ are one-dimensional, the four constants $a_\varkappa$ must be real if the representations are real (and the components $\mathscr{H}_i$ and matrices $\boldsymbol{X}_i$ are chosen as indicated above); but if two of these representations, say $\mathscr{D}_2$ and $\mathscr{D}_3$, are complex conjugates, hermiticity implies that $a_2=a_3^*$, and a_2 and a_3 are complex, so that the number of independent real constants remains equal to four.

As an example let us calculate the spectrum in a diamond lattice for the two-dimensional representations $\Gamma_{12}^\pm$ and X_1, X_3.

TABLE 26.1. Basis functions for representations of $\boldsymbol{D_{4h}=D_{2d}\times C_i}$

Representations		Odd	Even
Group D_{2d}	Group D_{4h}	(with respect to time reversal)	
A_1	A_1^+		k_x^2; $k_y^2+k_z^2$; ε_{xx}; $\varepsilon_{yy}+\varepsilon_{zz}$
	A_1^-		$\sigma_y k_y-\sigma_z k_z$
A_2	A_2^+	σ_x	
	A_2^-		$\sigma_y k_z+\sigma_z k_y$
B_1	B_1^+		$k_y^2-k_z^2$; $\varepsilon_{yy}-\varepsilon_{zz}$
	B_1^-		$\sigma_y k_y+\sigma_z k_z$; $\sigma_x k_x$
B_2	B_2^+		$k_y k_z$; ε_{yz}
	B_2^-	k_x	$\sigma_y k_z-\sigma_z k_y$
E	E^+	σ_y, σ_z	$k_x k_y$, $k_x k_z$; ε_{xz}, ε_{xy}
	E^-	k_y, k_z	$\sigma_x k_y$, $\sigma_x k_z$; $\sigma_y k_x$, $\sigma_z k_x$

To construct the spectrum, we first use the basis functions of Table 11.1 (p. 72) to construct functions of the components k_i which transform according to the irreducible representations of the groups O_h and $D_{4h}=D_{2d}\times C_i$ (the latter is the point group at X). These functions are given in Tables 26.1 and 26.2, which also list functions which transform according to the correspond ing representations of T_d and D_{2d}. Since all the representations $\Gamma_i^\pm$ and X_i belong to case (a$_1$), it follows from (25.36) and from the character table (Table 11.1) that for the representations $\Gamma_{12}^\pm$ the matrix $\mathscr{H}(\boldsymbol{k})$ must contain even functions of $\boldsymbol{k}_i$ which transform according to the representations Γ_1^+ and Γ_{12}^+, and odd functions which transform according to Γ_2^+. Since the basis functions of $\Gamma_{12}^\pm$ are complex conjugates, we have from (26.3) that

$$\mathscr{H}(\boldsymbol{k})=Ak^2+B\left[\sigma_+(k_x^2+\varepsilon_3k_y^2+\varepsilon_3^2k_z^2)+\sigma_-(k_x^2+\varepsilon_3^2k_y^2+\varepsilon_3k_z^2)\right]. \tag{26.4}$$

Solving the equation $|\mathscr{H}(\boldsymbol{k})-\boldsymbol{I}E|=0$, we obtain

$$E(\boldsymbol{k})=Ak^2\pm B\left\{k^4-3(k_x^2k_y^2+k_x^2k_z^2+k_y^2k_z^2)\right\}^{1/2}, \tag{26.5}$$

which is just (24.4). The same equation is valid for the representation Γ_{12} of T_d^2, since the components $\boldsymbol{k}_i$ that transform according to the representations Γ_1 and Γ_{12} in T_d are the same.

For the representations X_1 and X_3, we see from (25.36) and Table 11.1 that $\mathscr{H}(\boldsymbol{k})$ must contain $\boldsymbol{k}_i$ which transform according to the representations $[X_{1,3}^2]=A_1^++B_1^-+B_2^+$ of D_{4h}, and odd functions which transform according to $\{X_{1,3}^2\}=A_2^-$. Letting these functions be real, in accordance with (26.3), we obtain

$$\mathscr{H}(\boldsymbol{k})=A_1k_x^2+A_2(k_y^2+k_z^2)+2A_3\sigma_zk_yk_z. \tag{26.6}$$

TABLE 26.2. Basis functions for representations of $\boldsymbol{O_h}$ and $\boldsymbol{T_d}$

Representations		Odd	Even
Group T_d	Group O_h	(with respect to time reversal)	
A_1 (Γ_1)	A_1^+		k^2; $\mathrm{Sp}\,\varepsilon$
	A_1^-		$k_x\sigma_x + k_y\sigma_y + k_z\sigma_z$
A_2 (Γ_2)	A_2^+	$[[J_zJ_y]\,J_x]$	
	A_2^-		
E (Γ_{12})	E^+		$J_1 = J_x^2 + \varepsilon_3 J_y^2 + \varepsilon_3^2 J_z^2,\ J_2 = J_1^+$ $k_1 = k_x^2 + \varepsilon_3 k_y^2 + \varepsilon_3^2 k_z^2,\ k_2 = k_1^*$ $\varepsilon^{(1)} = \varepsilon_{xx} + \varepsilon_3\varepsilon_{yy} + \varepsilon_3^2\varepsilon_{zz},\ \varepsilon^{(2)} = \varepsilon^{(1)*}$
	E^-		$K_1 = k_x\sigma_x + \varepsilon_3 k_y\sigma_y + \varepsilon_3^2 k_z\sigma_z,\ K_2 = K_1^+$
F_1 (Γ_{25})	F_1^+	$\sigma_x,\ \sigma_y,\ \sigma_z$ $J_x,\ J_y,\ J_z;\ J_x^3,\ J_y^3,\ J_z^3$	
	F_1^-		$k_x\sigma_y + k_y\sigma_x,\ k_y\sigma_z + k_z\sigma_y,\ k_z\sigma_x + k_x\sigma_z$
F_2 (Γ_{15})	F_2^+	$V_x = [J_x(J_y^2 - J_z^2)]$ $V_y = [J_y(J_z^2 - J_x^2)]$ $V_z = [J_z(J_x^2 - J_y^2)]$	$[J_xJ_y],\ [J_yJ_z],\ [J_zJ_x]$ $k_xk_y,\ k_yk_z,\ k_zk_x;\ \varepsilon_{xy},\ \varepsilon_{yz},\ \varepsilon_{zx}$
	F_2^-	$k_x,\ k_y,\ k_z$	$k_x\sigma_y - k_y\sigma_x,\ k_y\sigma_z - k_z\sigma_y,\ k_z\sigma_x - k_x\sigma_z$

$\varepsilon_3 = e^{2\pi i/3}$

Since the two linear independent functions k_x^2 and $k_y^2 + k_z^2$ transform according to A_1^+, each of these appears in $\mathcal{H}(\boldsymbol{k})$ with a (possibly) different constant. Consequently, the electron spectrum for the representations $X_{1,3}$ is

$$E(\boldsymbol{k}) = A_1k_x^2 + A_2(k_y^2 + k_z^2) \pm 2A_3k_yk_z = A_1k_x^2 + (A_2 \pm A_3)\,k_{y'}^2 + (A_2 \mp A_3)\,k_{z'}^2, \qquad (26.7)$$

where $k_{y'} = \frac{1}{\sqrt{2}}(k_y + k_z),\ k_{z'} = \frac{1}{\sqrt{2}}(k_z - k_y)$.

The spectrum (26.7) is of course the same as (24.29), but whereas in §24 a knowledge of the representation matrices was required, in the theory of invariants all we need to construct $\mathcal{H}(\boldsymbol{k})$ are the characters.

Basis functions transform according to vector or spinor representations or the p-equivalent projective representations

If $\mathscr{D}(r)$ in equation (25.14a) is a vector or spinor representation, the corresponding basis functions, as noted in §11 and 16, transform like definite combinations of the functions Y_m^l for different m and a certain integer or half-integer weight j. The functions Y_m^l, which, as shown in §10, realize an irreducible representation of the rotation group, are eigenfunctions of the operator J_z. By (10.9) the matrices $\boldsymbol{J}_\alpha^l$ transform under any transformation in the rotation group like components of a vector:

$$\boldsymbol{\mathcal{D}}_l(r^{-1})\,\boldsymbol{J}_\alpha^l\boldsymbol{\mathcal{D}}_l(r)=\sum_{\alpha'}\mathscr{D}_1^{-1}{}_{\alpha'\alpha}(r)\,\boldsymbol{J}_{\alpha'}^l=r\boldsymbol{J}_\alpha^l, \tag{26.8}$$

where $\boldsymbol{\mathcal{D}}_1$ is the matrix of the representation $\mathscr{D}_1$, while inversion operations do not change their sign, since $\boldsymbol{\mathcal{D}}_l^{\pm}(i)=\pm\boldsymbol{I}$. Consequently, under any transformation in a finite point group which is a subgroup of the orthogonal group, the matrices $\boldsymbol{J}_\alpha^l$ transform like components of a pseudovector. In this case, then, the matrices $\boldsymbol{J}_x^l$, $\boldsymbol{J}_y^l$, $\boldsymbol{J}_z^l$ and their products, relative to the appropriate basis, may play the part of the matrices $\boldsymbol{X}_l^\varkappa$.

Since the representation $\mathscr{D}_j$ may be reducible for the point group, the matrices $\boldsymbol{J}_\alpha^l$ and their products decompose into smaller "blocks" which transform under the point group operations according to (26.8), independently of each other. The orders of these blocks are determined by the dimensions of the constituents of the representation $\mathscr{D}_j$. Thus the basis matrices $\boldsymbol{X}$ for a given representation $\mathscr{D}_\varkappa$ should be these blocks, relative to a suitable basis. Accordingly, of all the possible functions that can be built up from the components J_x, J_y, J_z and their products, we must keep n_s^2 functions $X_l^\varkappa$ which transform according to the representations $\mathscr{D}_\varkappa$ occurring in $\mathscr{D}\times\mathscr{D}^*$. We next construct the invariant operator (25.6) from products of $\mathscr{H}_l^\varkappa$ and $\boldsymbol{X}_l^\varkappa$ that transform according to complex conjugate representations, and replace the operators J_i throughout by matrices $\boldsymbol{J}_i$ relative to a suitable basis, as defined by (10.17). Since by (25.5) the matrices $\boldsymbol{X}_l^\varkappa$ may be assumed to coincide for all projective-equivalent representations, it is sufficient to construct the matrices for one such representation; the most convenient choice is a representation whose basis consists of the functions Y_m^l with the smallest j.

Recall (see §13) that two unitarily inequivalent representations $\mathscr{D}$ and $\mathscr{D}'$ belonging to one factor system are p-equivalent if one of them is obtained from the other by multiplying by a one-dimensional representation. Hence it follows that for all groups which direct products of some group with C_i, C_s or C_2, vector and spinor representations which are even and odd relative to the operations i, σ_h or c_2, respectively, are always p-equivalent. Thus, the three-dimensional representations F_1 and F_2 of the groups O and T_d are p-equivalent in all point groups which have several three-dimensional representations. Therefore, we may assume that the matrices $\boldsymbol{X}_l^\varkappa$ are the same for all three-dimensional representations. This is also true of four-dimensional spinor representations of point groups. Point groups do not have vector or spinor representations of higher dimensions, while for two-dimensional representations it is more convenient to use the simpler method presented above.

Since the functions Y_m^1 $(m=1, 0, -1)$ form a basis for one of the three-dimensional representations of each of these groups, while the functions $Y_m^{3/2}$ $(m=\pm 3/2, \pm 1/2)$ are a basis for one of the four-dimensional representations, it is sufficient in practice to construct only the matrices $\boldsymbol{J}_i^{3/2}$ and $\boldsymbol{J}_i^1$ and their products (see Table 26.3 at the end of this section).

In order to allow for the additional requirements imposed on $\mathscr{H}(\mathscr{K})$ by time reversal symmetry, we must classify the matrices $\boldsymbol{X}_l^{\varkappa}$ as even or odd with respect to time reversal. Using equations (25.36) or (25.37), we can at once find the representations according to which the even and odd matrices transform. In case (a$_1$) the classification is immediate. In fact, by (18.46) the matrices of the operator J_i, as they are odd with respect to time reversal, satisfy the condition $\boldsymbol{T}^{-1}\boldsymbol{J}_i\boldsymbol{T} = -\tilde{J}_i$, i.e., $f = -1$ in (25.17). Consequently, in case (a$_1$) all matrices which do not change sign when J_i is replaced by $-J_i$ and the order of components in the products $J_iJ_kJ_l...$ is reversed are even, while the matrices which change sign are odd.

Of course, this simple rule is not valid in the other cases; for example, it follows from (25.26) that in case (a$_2$) the matrix $\boldsymbol{X}_l^{\varkappa}(J)$ is even or odd according as

$$\boldsymbol{T}^{-1}\boldsymbol{X}(J)\boldsymbol{T} = f\tilde{\boldsymbol{X}}(R^{-1}J) \text{ with } f = +1 \text{ or } f = -1.$$

In this case construction of even and odd matrices is simpler if equation (25.37) is used directly.

As an example, we construct the spectrum for the representations $\Gamma_{15}^{\pm}$ and $\Gamma_{25}^{\pm}$ in an O_h lattice and the representations Γ_{15} and Γ_{25} in a T_d lattice.

For Γ_{15} and Γ_{25}, the matrices X_i are the nine linearly independent matrices

$$\boldsymbol{J}_x,\ \boldsymbol{J}_y,\ \boldsymbol{J}_z,\ \boldsymbol{J}_x^2,\ \boldsymbol{J}_y^2,\ \boldsymbol{J}_z^2,\ [\boldsymbol{J}_x\boldsymbol{J}_y],\ [\boldsymbol{J}_x\boldsymbol{J}_z],\ [\boldsymbol{J}_y\boldsymbol{J}_z]$$

in the representation Y_m^1.

In accordance with (25.36), $\mathscr{H}(\mathscr{K})$ must contain the products of even functions of $\mathscr{K}_l$ and J_l which transform according to the representations $\Gamma_1^+ + \Gamma_{12}^+ + \Gamma_{25}^+$ and odd functions which transform according to Γ_{15}^+. Consequently, the Hamiltonian $\mathscr{H}(\mathscr{K})$ is

$$\begin{aligned}\mathscr{H}(\boldsymbol{k}) = Ak^2\boldsymbol{I} - B\,&[(k_x^2 + \varepsilon_3 k_y^2 + \varepsilon_3^2 k_z^2)(\boldsymbol{J}_x^2 + \varepsilon_3^2\boldsymbol{J}_y^2 + \varepsilon_3\boldsymbol{J}_z^2) + \\ &+ (k_x^2 + \varepsilon_3^2 k_y^2 + \varepsilon_3 k_z^2)(\boldsymbol{J}_x^2 + \varepsilon_3\boldsymbol{J}_y^2 + \varepsilon_3^2\boldsymbol{J}_z^2)] - 2\sqrt{3}D\,(k_xk_y[\boldsymbol{J}_x\boldsymbol{J}_y] + \\ &+ k_xk_z[\boldsymbol{J}_x\boldsymbol{J}_z] + k_yk_z[\boldsymbol{J}_y\boldsymbol{J}_z]) = Lk^2\boldsymbol{I} - (L-M)(k_x^2\boldsymbol{J}_x^2 + k_y^2\boldsymbol{J}_y^2 + k_z^2\boldsymbol{J}_z^2) - \\ &- 2N\,(k_xk_y[\boldsymbol{J}_x\boldsymbol{J}_y] + k_xk_z[\boldsymbol{J}_x\boldsymbol{J}_z] + k_yk_z[\boldsymbol{J}_y\boldsymbol{J}_z]).\end{aligned} \tag{26.9}$$

We have used the identities $\boldsymbol{J}_x^2 + \boldsymbol{J}_y^2 + \boldsymbol{J}_z^2 = j(j+1)\boldsymbol{I} = 2\boldsymbol{I}$.

If we now write (26.9) in matrix form using Table 26.3, we obtain the matrix $\mathscr{H}(\boldsymbol{k})$ of (24.5). For the T_d lattice, the matrix $\mathscr{H}(\boldsymbol{k})$, as is clear from Table 26.2, is also given by (26.9), since it must include even functions of k_i which transform according to Γ_1, Γ_{12}, Γ_{25} and odd functions which transform according to Γ_{15}. In the presence of a magnetic field, the matrix (26.9) will also include a term

$$\mathscr{H}(\boldsymbol{H}) = \mu_0 g_0 \frac{3\varkappa - 1}{2}(\boldsymbol{J}\boldsymbol{H}), \tag{26.10}$$

since the components H_i transform according to Γ_{15}^+ and are odd with respect to time reversal.

Projective representations are not equivalent to vector or spinor representations

In the general case of projective representations with dimension higher than two, the basis matrices must be sought directly from the relations (25.5) for the generators. As in the preceding case, it is sufficient to construct basis matrices for one of the p-equivalent projective representations corresponding to each factor system with projective representations of dimension > 2, which are not p-equivalent to spinor representations. As we see from Tables 14.2 and 14.3, only the groups D_{4h}, D_{6h} and O_h have such representations. The group D_{4h} has one four-dimensional representation for each of the classes K_5 and K_6 of factor systems. Equation (25.14a) at once shows that for both these representations $\mathscr{D} \times \mathscr{D}^*$ contains each one-dimensional vector representation of the group D_{4h} once and each two-dimensional representation twice. Consequently, one of the matrices $\boldsymbol{X}_l^{\varkappa}$ transforms according to the representation A_2^-, i.e., like z, and two transform according to E^-, i.e., like x and y.

The group D_{6h} has one four-dimensional representation for each of classes K_3, K_4, K_5, K_6.

For all these representations the product $\mathscr{D} \times \mathscr{D}^*$ contains every vector representation once, i.e., one of the matrices $\boldsymbol{X}_l^{\varkappa}$ transforms like z (according to A_2^-) and one like x and y (according to E_1^-).

The group O_h has three four-dimensional and one six-dimensional projective representations which are not p-equivalent to each other or to spinor representations. These representations belong to classes K_2 and K_3. For the four-dimensional representations $Q_2^{(3)}$ and $Q_3^{(3)}$, $\mathscr{D} \times \mathscr{D}^*$ contains $A_1^+ + A_2^+ + E^- + F_1^+ + F_1^- + F_2^+ + F_2^-$, and for the representation $Q_1^{(3)}$ this product includes $A_1^\pm$, $A_2^\pm$, $F_1^\pm$, $F_2^\pm$; in both cases, then, one matrix transforms like x, y, z (according to F_1^-). For the six-dimensional representation $R^{(2)}$, $\mathscr{D} \times \mathscr{D}^*$ contains $A_1^\pm$ and $A_2^\pm$ once and $E^\pm$, $F_1^\pm$, $F_2^\pm$ twice, so that in the last case there are two linearly independent matrices which transform like x, y, z (according to F_1^-). These matrices are listed in Table 26.4 at the end of the section. The other linearly independent matrices $\boldsymbol{X}_l^{\varkappa}$ are easily obtained as products of these matrices.

Spinor representations of space groups

For spinor representations of space and point groups, the operator $\mathscr{H}(\mathscr{K})$ can be constructed in two ways. One approach is by direct use of the spinor representations $\mathscr{D}(g)$, in which case we can construct the matrices $\boldsymbol{X}_l^{\varkappa}$ by one of the methods considered above. It should be noted, though, when determining $N_\varkappa$, that for spinor representations $K^2 = -1$, as opposed to $K^2 = 1$ for vector representations. This method is equivalent to the construction of $\mathscr{H}(\mathscr{K})$ in perturbation theory by direct incorporation of the spin-orbit coupling operator $(\mathscr{H}_{so})_{k_0}$ in $\mathscr{H}_0$. A shortcoming of the method is that one must ascertain a priori which of the constants $a_\varkappa$ in $\mathscr{H}(\mathscr{K})$ are relativistically small.

This problem is automatically solved in the second approach to the construction of $\boldsymbol{\mathcal{H}}(\mathcal{K})$. Here we choose basis functions $\psi_{m\alpha} = \psi_m(\boldsymbol{x})\alpha$ which transform according to the direct product $\mathcal{D} = \mathcal{D}_l \times \mathcal{D}_{1/2}$ of the representation $\mathcal{D}_l$ according to which the coordinate functions $\psi_m(\boldsymbol{x})$ transform and the representation $\mathcal{D}_{1/2}$ according to which the spin functions α, β transform. In this representation the matrices $\boldsymbol{X}$ in (25.6) may be expressed as direct products of matrices $\boldsymbol{X}_l^\varkappa$ whose elements $X_{l,\,mm'}^\varkappa$ depend only on the coordinate functions ψ_m and the matrices of the operators $\sigma_x, \sigma_y, \sigma_z$, which act only on the spin variables. This means that the operator $\boldsymbol{\mathcal{H}}(\mathcal{K})$ may be presented as a product of the matrices $\boldsymbol{X}_l^\varkappa$, the components $\mathcal{K}_l^\varkappa$ and the matrices σ_i or the identity matrix. According to (10.9) and (10.26), the matrices σ_i transform like the components of a pseudovector, i.e.,

$$\mathcal{D}_{1/2}^{-1}(g)\,\sigma_i\,\mathcal{D}_{1/2}(g) = \sum_l \mathcal{D}^{+}_{1\ li}(g^{-1})\,\sigma_l = g\,\sigma_l,$$

and so condition (25.1) in the representation $\mathcal{D}(g) = \mathcal{D}_l(g)\mathcal{D}_{1/2}(g)$ becomes

$$\mathcal{D}_l^{-1}(g)\,\boldsymbol{\mathcal{H}}(g\sigma,\ g\mathcal{K})\,\mathcal{D}_l(g) = \boldsymbol{\mathcal{H}}(\sigma,\ \mathcal{K}). \tag{26.11}$$

Formally speaking, the only difference between equation (26.11) and (25.1) is that, together with the components of $\mathcal{K}$ ($\boldsymbol{k}$, $\boldsymbol{\varepsilon}$, $\boldsymbol{x}$, $\boldsymbol{H}$ etc.), we also have the components of the pseudovector $\boldsymbol{\sigma}$ with $f_\varkappa = -1$ and their products with the other components of $\mathcal{K}$. All these products can be expanded in terms of irreducible representations and denoted as before by $\mathcal{K}_l^\varkappa$. In all other respects, the construction of $\boldsymbol{\mathcal{H}}(\mathcal{K})$ remains unchanged. We stress that in this approach both the form of the basis matrices and the values of $N_\varkappa$ depend only on the representation $\mathcal{D}_l$ according to which the coordinate functions transform. Furthermore, in the final equations the products $X_l^\varkappa\sigma_i$ must be replaced by direct products of the matrices $\boldsymbol{X}_l^\varkappa$ and σ_i.

Essentially, this second approach means that the basis is made up of eigenfunctions of operator $\mathcal{H}_0$ without $\mathcal{H}_{so}$; the operator $\mathcal{H}_{so}$ and the other terms in $\mathcal{H}$ are all treated as a perturbation. One can thus ascertain at once which of the coefficients entering $\boldsymbol{\mathcal{H}}(\mathcal{K})$ are small.

Since without spin-orbit coupling the operator $\mathcal{H}$ and, consequently, $\boldsymbol{\mathcal{H}}(\mathcal{K})$ do not involve the components σ_i, all the coefficients of the terms containing the spin operators must be relativistic. However, we must bear in mind that, as noted in §21, the relativistic terms in $\boldsymbol{\mathcal{H}}(\mathcal{K})$ are not always small. If spin-orbit splitting is comparable with or greater than the band gap E_g or the separation δE to the nearest bands, the order of magnitude of the coefficients cannot be correctly determined unless we construct $\boldsymbol{\mathcal{H}}(\mathcal{K})$ with all close-lying terms incorporated.

In the second approach to constructing $\boldsymbol{\mathcal{H}}(\mathcal{K})$, all the irreducible constituents of $\mathcal{D} \times \mathcal{D}_{1/2}$ are included "en bloc." If we wish to determine $\boldsymbol{\mathcal{H}}(\mathcal{K})$ for only one of the constituents, it is necessary to diagonalize the resulting operator; this may be a highly tedious task, and then the first approach is preferable.

As an example, we consider construction of the spectrum at the points Γ and X for an O_h^7 lattice and at the point Γ for a T_d lattice, including spin-orbit coupling.

We first consider the point Γ. Using the first method, we can construct $\boldsymbol{\mathcal{H}}(\boldsymbol{k})$ at once for the fourfold degenerate spinor representations $\Gamma_8^\pm$ and Γ_8.

In this case we can take as basis the sixteen linearly independent matrices $\boldsymbol{J}_x, \boldsymbol{J}_y, \boldsymbol{J}_z, \boldsymbol{J}_x^2, \boldsymbol{J}_y^2, \boldsymbol{J}_z^2, [\boldsymbol{J}_x\boldsymbol{J}_y], [\boldsymbol{J}_x\boldsymbol{J}_z], [\boldsymbol{J}_y\boldsymbol{J}_z], \boldsymbol{J}_x^3, \boldsymbol{J}_y^3, \boldsymbol{J}_z^3, \boldsymbol{V}_x = [\boldsymbol{J}_x(\boldsymbol{J}_y^2 - \boldsymbol{J}_z^2)], \boldsymbol{V}_y, \boldsymbol{V}_z, [\boldsymbol{J}_x[\boldsymbol{J}_y\boldsymbol{J}_z]]$ in the representation $Y_m^{3/2}$ (see Table 26.3). For the representations $\Gamma_8^\pm$ of O_h, it follows from (25.36) that $\mathscr{H}(\boldsymbol{k})$ must contain even functions of $\boldsymbol{k}$ which transform according to the representations $\{\Gamma_8^{\pm^2}\} = \Gamma_1^+ + \Gamma_{12}^+ + \Gamma_{25}^+$ and odd functions which transform according to $[\Gamma_8^{\pm^2}] = \Gamma_2^+ + 2\Gamma_{15}^+ + \Gamma_{25}^+$ (for the appropriate functions, see Table 26.2). Thus the Hamiltonian is

$$\mathscr{H}(\boldsymbol{k}) = -\left(\gamma_1 + \frac{5}{2}\gamma_2\right)k^2\boldsymbol{I} + 2\gamma_2(\boldsymbol{J}_x^2k_x^2 + \boldsymbol{J}_y^2k_y^2 + \boldsymbol{J}_z^2k_z^2) + \\ + 4\gamma_3([\boldsymbol{J}_x\boldsymbol{J}_y]k_xk_y + [\boldsymbol{J}_x\boldsymbol{J}_z]k_xk_z + [\boldsymbol{J}_y\boldsymbol{J}_z]k_yk_z). \tag{26.12}$$

(We have transformed the product $J_1k_2 + J_2k_1$, as in (26.9).)

Substituting the matrices $\boldsymbol{J}_i$ from Table 26.3, we obtain the same matrix $\mathscr{H}(\boldsymbol{k})$ as in (24.13), giving the spectrum (24.14):

$$E_{1,2}(\boldsymbol{k}) = Ak^2 \pm [B^2k^4 + (D^2 - 3B^2)(k_x^2k_y^2 + k_x^2k_z^2 + k_y^2k_z^2)]^{1/2}.$$

The constants γ_1, γ_2 and γ_3 in (26.12), which coincide with the dimensionless Luttinger constants up to a dimension factor $\hbar^2/2m$, are related to the constants of (24.6), (24.12) and (26.9) by

$$\gamma_1 = -\frac{L + 2M}{3} = -A, \quad \gamma_2 = -\frac{L - M}{6} = -\frac{B}{2}, \\ \gamma_3 = -\frac{N}{6} = -\frac{D}{2\sqrt{3}}. \tag{26.13}$$

In a magnetic field, the Hamiltonian $\mathscr{H}(\boldsymbol{H})$ will include the products of the components H_i by $\boldsymbol{J}_i$ and $\boldsymbol{J}_i^3$, since the representation Γ_{15}^+ according to which the components H_i transform occurs twice in $[\Gamma_8^{\pm^2}]$. Therefore,

$$\mathscr{H}(\boldsymbol{H}) = \mu_0 g_0 \{\mathscr{k}(\boldsymbol{J}\boldsymbol{H}) + \mathscr{q}(\boldsymbol{J}_x^3H_x + \boldsymbol{J}_y^3H_y + \boldsymbol{J}_z^3H_z)\}. \tag{26.14}$$

If $\gamma_2 = \gamma_3$, then $\mathscr{H}(\boldsymbol{k}) = -\left(\gamma_1 + \frac{5}{2}\gamma_2\right)k^2\boldsymbol{I} + 2\gamma_2(\boldsymbol{J}\boldsymbol{k})^2$; this Hamiltonian has spherical symmetry. Its eigenfunctions are functions of $j = 3/2$: $\Psi_m^{3/2\boldsymbol{k}}$, $m = \pm 3/2, \pm 1/2$, with quantization axis in the $\boldsymbol{k}$ direction. The energies are $E_{\pm 3/2} = -\gamma_1 + 2\gamma_2$, $E_{\pm 1/2} = -\gamma_1 - 2\gamma_2$. Functions $\Psi_{m'}^{3/2\boldsymbol{k}'}$ and $\Psi_m^{3/2\boldsymbol{k}}$ with different quantization axes are related by

$$\Psi_{m'}^{3/2\boldsymbol{k}'} = \sum_m \mathscr{D}_{mm'}^{3/2}(g)\Psi_m^{3/2\boldsymbol{k}},$$

where $\mathscr{D}^{3/2}(g)$ are matrices of the representations of the rotation group with $j = 3/2$ for an element g which takes $\boldsymbol{k}'$ into $\boldsymbol{k}$: $\boldsymbol{k} = g\boldsymbol{k}'$. In particular, a relation of this kind connects the functions $\Psi_m^{3/2\boldsymbol{k}}$ for arbitrary $\boldsymbol{k}$ to the functions (23.2) with the quantization axis in the z direction. Therefore, when $\gamma_2 = \gamma_3$ we may take the matrix elements $\mathscr{D}_{mm'}^{3/2}$ to be the coefficients in (26.14).

As indicated above, the representation $\Gamma_8^\pm$ may originate either in splitting of the representation $\Gamma_{15}^\pm$ or $\Gamma_{25}^\pm$ due to spin-orbit coupling, or in the representation $\Gamma_{12}^\pm$, which does not split. If we construct the matrix $\mathscr{H}(\boldsymbol{k})$ by the second method for $\Gamma_{15}^\pm$ and $\Gamma_{25}^\pm$ and include the relativistic terms, it will have additional terms which are odd with respect to time reversal, transform according to Γ_{25}^+ and contain the components σ_i:

$$\begin{aligned}\mathcal{H}(\boldsymbol{k}) = {} & A_1 k^2 \boldsymbol{I} + \Delta_1(\sigma_x J_x + \sigma_y J_y + \sigma_z J_z) - A_2(J_x^2 k_x^2 + J_y^2 k_y^2 + J_z^2 k_z^2) - \\ & - 2A_3([J_x J_y] k_x k_y + [J_x J_z] k_x k_z + [J_y J_z] k_y k_z) + a_1 k^2 (\boldsymbol{J}\boldsymbol{\sigma}) + \\ & + a_2(J_x \sigma_x k_x^2 + J_y \sigma_y k_y^2 + J_z \sigma_z k_z^2) + \\ & + a_3([J_x \sigma_y] k_x k_y + [J_y \sigma_z] k_y k_z + [J_x \sigma_z] k_x k_z), \qquad (26.15)\end{aligned}$$

where $A_1 = L$, $A_2 = L - M$, and $A_3 = N$. The second term in this formula is the spin-orbit coupling, the last three are relativistic corrections to the constants A, B, D of (24.13)–(24.14a):

$$\delta A = a_1 - \frac{1}{3} a_2, \quad \delta B = -\frac{2}{3} a_2, \quad \delta D = -\frac{1}{\sqrt{3}} a_3. \qquad (26.15a)$$

For the band Γ_6 in (24.13b), $\delta A = -2a_1 + \frac{2}{3} a_2$. Of all the constants in (26.12) and (26.14), only q is relativistic and is consequently small compared to $\varkappa$.*

If the representation $\Gamma_8^{\pm}$ originates in $\Gamma_{12}^{\pm}$, a comparison of (24.4) and (24.14a) shows that $D = 0$ and $C^2 = -3B^2$ in the nonrelativistic approximation. If we construct $\mathcal{H}(\boldsymbol{k})$ for the representation $\Gamma_{12}^{\pm}$, including the relativistic terms, it follows from (25.15) that we must include products of the components σ_i and $k_l k_m$ which are odd with respect to time reversal and transform according to the representation Γ_2^{+}:

$$\mathcal{H}(\boldsymbol{k}) = Ak^2 + B(\rho_+ k_1 + \rho_- k_2) + a\rho_z(\sigma_x k_y k_z + \sigma_y k_x k_z + \sigma_z k_x k_y). \qquad (26.16)$$

Here ρ_l are the basis matrices $X_l^{\varkappa}$ defined by (26.3), and σ_l the spin matrices; $\rho_l \sigma_k$ denotes the direct product of the corresponding matrices. In matrix form,

$$\mathcal{H}(\boldsymbol{k}) = \begin{Vmatrix} Ak^2 + ak_x k_y & a(k_y - ik_x)k_z & Bk_1 & 0 \\ a(k_y + ik_x)k_z & Ak^2 - ak_x k_y & 0 & Bk_1 \\ Bk_2 & 0 & Ak^2 - ak_x k_y & -a(k_y - ik_x)k_z \\ 0 & Bk_2 & -a(k_y + ik_x)k_z & Ak^2 + ak_x k_y \end{Vmatrix}. \qquad (26.16a)$$

The equation $|\mathcal{H}(\boldsymbol{k}) - \boldsymbol{I}E| = 0$ yields a degeneracy for $E(\boldsymbol{k})$ in agreement with (24.14), $D^2 = a^2$.

The Hamiltonian $\mathcal{H}(\boldsymbol{k})$ for the representation Γ_8 of the group T_d includes even products of the components k_l which transform according to the representations $\Gamma_1 + \Gamma_{12} + \Gamma_{25}$ (i.e., the same as for the group O_h), and odd products which transform according to $\Gamma_2 + 2\Gamma_{15} + \Gamma_{25}$, i.e., as opposed to (26.12), $\mathcal{H}(\boldsymbol{k})$ will have linear k terms, containing products of the components k_l and V_l which transform (see Table 26.2) according to the representation Γ_{25}:

$$\mathcal{H}'(\boldsymbol{k}) = \frac{4}{\sqrt{3}} \varkappa_0 (k_x [J_x (J_y^2 - J_z^2)] + k_y [J_y (J_z^2 - J_x^2)] + k_z [J_z (J_x^2 - J_y^2)]). \qquad (26.17)$$

In matrix notation,

$$\mathcal{H}'(\boldsymbol{k}) = \varkappa_0 \begin{Vmatrix} 0 & k_y - ik_x & -2k_z & \sqrt{3}(k_y + ik_x) \\ k_y + ik_x & 0 & -\sqrt{3}(k_y - ik_x) & 2k_z \\ -2k_z & -\sqrt{3}(k_y + ik_x) & 0 & k_y - ik_x \\ \sqrt{3}(k_y - ik_x) & 2k_z & k_y + ik_x & 0 \end{Vmatrix}. \qquad (26.17a)$$

---

* For example, for p-Ge, according to the data in /31.4/, $\varkappa = 3.41 \pm 0.03$, $q = 0.06 \pm 0.01$.

For the representation Γ_{12} of the group T_d the linear k matrix $\mathcal{H}'(\boldsymbol{k})$ may also be constructed by the second approach. For this purpose we must include in $\mathcal{H}$ relativistically small components $K_1 = k_x\sigma_x + \varepsilon_3 k_y\sigma_y + \varepsilon_3^2 k_z\sigma_z$ and $K_2 = K_1^+$, which are even with respect to time reversal and transform according to the representation Γ_{12}:

$$\mathcal{H}(\boldsymbol{k}) = 2\varkappa_0(\rho_+ K_1 + \rho_- K_2). \tag{26.18}$$

The matrix (26.18) may be transformed into (26.17a) by a suitable unitary matrix.

Since the matrix (26.9) constructed for the representations Γ_{15} and Γ_{25} of the group T_d does not contain terms linear in $\boldsymbol{k}$, it is clear that the constants $\varkappa_0$ in (26.17) are also relativistically small for the representation Γ_8, originating in Γ_{15} or Γ_{25}.

By (24.12) and (26.17), the secular equation $|\mathcal{H}(\boldsymbol{k}) + \mathcal{H}'(\boldsymbol{k}) - IE| = 0$ becomes

$$\begin{aligned}(E - Ak^2)^4 - 2(E - Ak^2)^2(B^2k^4 + C^2\Theta + 4\varkappa_0^2k^2) &+ \\ + 16\sqrt{3}(E - Ak^2)D\varkappa_0^2\Theta + (B^2k^4 + C^2\Theta)^2 &+ \\ + (2\varkappa_0)^4(k^4 - 3\Theta) + 8B^2\varkappa_0^2(k_x^6 + k_y^6 + k_z^6) &- \\ - 4(3B^2 + 2D^2)\varkappa_0^2k^2\Theta + 84B^2\varkappa_0^2k_x^2k_y^2k_z^2 &= 0,\end{aligned} \tag{26.19}$$

$$\text{where } \Theta = k_x^2k_y^2 + k_x^2k_z^2 + k_y^2k_z^2.$$

This equation can be solved exactly only for a few special directions. Thus, if $\boldsymbol{k}$ is in the [100] direction, i.e., $k_y = k_z = 0$, or in the [010] or [011] direction, then

$$E_{100}(\boldsymbol{k}) = Ak^2 \pm (B^2k^4 + 4\varkappa_0^2k^2)^{1/2}. \tag{26.20}$$

Consequently, in these directions each term remains doubly degenerate.

In the [111] direction and those equivalent to it, one branch (corresponding to heavy holes for $A < 0$ and $D < 0$) splits completely:

$$E_{111}^- = \left(A - \frac{D}{\sqrt{3}}\right)k^2 \pm 2\sqrt{2}\,\varkappa_0 k, \tag{26.21a}$$

while the other branch (light holes) remains degenerate and is not changed:

$$E_{111}^+ = \left(A + \frac{D}{\sqrt{3}}\right)k^2. \tag{26.21b}$$

Because of the presence of linear k terms, the extremum is shifted from $\boldsymbol{k} = 0$ to points on the threefold axes. It is evident from (26.21) that the extremum point corresponds to $k_0 = \dfrac{\sqrt{2}\,\varkappa_0}{A - D/\sqrt{3}}$, and the energy at this point is lower than that at $k = 0$ by a quantity ΔE,

$$\Delta E = \frac{2\varkappa_0^2}{A - D/\sqrt{3}}. \tag{26.22}$$

The surfaces of constant energy near the extrema ($|\boldsymbol{k} - \boldsymbol{k}_0| \ll k_0$) are ellipsoids of revolution with axis of revolution in the [111] direction, and for holes near the extremum*

---

* Cyclotron resonance measurements have shown that for p-InSb

$$\Delta E = (6 \pm 3)\cdot 10^{-4}\ \text{ev},\ m_\perp^*/m_0 = 0.15 \pm 0.02,\ m_\parallel^*/m_0 = 0.57 \pm 0.10.$$

$$\frac{\hbar^2}{2m_{\parallel}^*} = -\left(A - \frac{D}{\sqrt{3}}\right), \quad \frac{\hbar^2}{2m_{\perp}^*} = -\frac{1}{2}\left(3A - 4B + \frac{D}{\sqrt{3}}\right). \qquad (26.23)$$

We now consider the spectrum at the point X of an O_h^7 lattice for the representation $X_5^{\pm}$. Since this fourfold degenerate representation originates in the representation X_1 or X_3, in order to construct $\mathcal{H}(\boldsymbol{k})$ we must incorporate relativistic terms in (26.6). These terms include products of $\boldsymbol{\sigma}_i$ and k_i, as given in Table 26.1; terms even with respect to time reversal transform according to $A_1^+ + B_1^- + B_2^+$ and odd ones transform according to A_2^-. Since the products $\sigma_i k_l$ transform according to odd representations and $\boldsymbol{\sigma}_i k_l k_m$ according to even representations, there remain only components which transform according to B_1^-, i.e., $\boldsymbol{\sigma}_x k_x$ and $\boldsymbol{\sigma}_y k_y + \sigma_z k_z$, so that

$$\mathcal{H}(\boldsymbol{k}) = Ak_x^2 + B(k_y^2 + k_z^2) + 2C\rho_z k_y k_z + \rho_x[a(\sigma_x k_x) + b(\boldsymbol{\sigma}_y k_y + \boldsymbol{\sigma}_z k_z)], \qquad (26.24)$$

whence

$$E(\boldsymbol{k}) = Ak_x^2 + B(k_y^2 + k_z^2) \pm [4C^2 k_y^2 k_z^2 + a^2 k_x^2 + b^2(k_y^2 + k_z^2)]^{1/2}. \qquad (26.25)$$

For the representation X_5, which originates in X_2 or X_4, the spectrum $E(\boldsymbol{k})$ naturally has the same form, but now the constant a in the linear k_x term is not relativistically small. Depending on the relations among the constants, the energy minimum for the spectrum defined by equation (26.25) is either on the axis k_x at the point $k_{x_0} = a/2A$, where $E = -a^2/4A$, or on axes k_y' and k_z' forming angles of 45° with the axes k_y and k_z.

Combination of several representations

Let us consider the construction of basis matrices for the two- and multiband models, when $\mathcal{H}(\mathcal{K})$ is constructed for several bands simultaneously. In this case the representation $\mathscr{D}(g)$ is reducible and combines representations $\mathscr{D}^{\mathrm{I}}(g)$, $\mathscr{D}^{\mathrm{II}}(g)$,... corresponding to the first, second, ... bands; the matrix $\mathscr{D}(g)$ is block-diagonal, as in the combination of complex conjugate representations in cases (b$_1$), (b$_2$) and (c). However, as opposed to those cases, the representations $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$ are unrelated. Therefore, for the off-diagonal parts $\mathcal{H}^{\mathrm{I\,II}}(\mathcal{K}) = \mathcal{H}^{\mathrm{II\,I}+}(\mathcal{K})$, which satisfy the condition

$$\mathscr{D}^{\mathrm{I}}(g^{-1})\,\mathcal{H}^{\mathrm{I\,II}}(\mathcal{K})\,\mathscr{D}^{\mathrm{II}}(g) = \mathcal{H}^{\mathrm{I\,II}}(g^{-1}\mathcal{K}), \qquad (26.26)$$

it follows from (25.12) that $N_\varkappa$ is given by an expression similar to (19.31):

$$N_\varkappa = \frac{1}{h}\sum_{g \in G'_{\boldsymbol{k}_0}} \chi_\varkappa(g)\,\chi^{\mathrm{II}*}(g)\,\chi^{\mathrm{I}}(g), \qquad (26.27)$$

i.e., the off-diagonal parts of $\mathcal{H}(\mathcal{K})$ may involve components $\mathcal{K}_l^\varkappa$ which transform according to the constituents $\mathscr{D}_\varkappa$ of the products $\mathscr{D}^{\mathrm{I}} \times \mathscr{D}^{\mathrm{II}*}$ and $\mathscr{D}^{\mathrm{I}*} \times \mathscr{D}^{\mathrm{II}}$. The additional requirements due to time reversal symmetry need not be taken into account, because, as noted in §19, they have no effect on the off-diagonal elements.

The number $N_\varkappa$ for the diagonal parts $\mathcal{H}^{\mathrm{I\,I}}(\mathcal{K})$ and $\mathcal{H}^{\mathrm{II\,II}}(\mathcal{K})$ is defined by one of the equations (25.36)–(25.38), depending on the behavior of the representations $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$ under time reversal. The matrices $\boldsymbol{X}_l^\varkappa$ can be classified as interband matrices, which contain only off-diagonal nonzero elements and transform according to the representations occurring in $\mathscr{D}^{\mathrm{I}} \times \mathscr{D}^{\mathrm{II}*}$ and $\mathscr{D}^{\mathrm{I}*} \times \times \mathscr{D}^{\mathrm{II}}$, and intraband matrices, which transform according to $\mathscr{D}^{\mathrm{I}} \times \mathscr{D}^{\mathrm{I}*}$ and $\mathscr{D}^{\mathrm{II}} \times \mathscr{D}^{\mathrm{II}*}$ and contain only diagonal elements. Since $\boldsymbol{\mathscr{D}}(g)$ is in block-diagonal form, these matrices transform independently under all transformations, including time reversal. They may be constructed by any of the methods described above.

If the projective representations $\mathscr{D}^{\mathrm{I}}(r)$ and $\mathscr{D}^{\mathrm{II}}(r)$ are p-equivalent and the corresponding matrices may be written as

$$\boldsymbol{\mathscr{D}}^{\mathrm{II}}(r) = \boldsymbol{\mathscr{D}}(r)\, \boldsymbol{\mathscr{D}}^{\mathrm{I}}(r), \tag{26.28}$$

where $\boldsymbol{\mathscr{D}}(r)$ is some one-dimensional representation, and if the matrices $\boldsymbol{X}_l^{\varkappa}$ which transform according to the constituents $\mathscr{D}_{\varkappa}$ of $\mathscr{D}^{I} \times \mathscr{D}^{I*}$ are known for one band, then the intraband matrices for the combined representation $\mathscr{D}(g)$ may be defined as

$$\boldsymbol{I}\boldsymbol{X}_l^{\varkappa} = \left\| \begin{matrix} \boldsymbol{X}_l^{\varkappa} & 0 \\ 0 & \boldsymbol{X}_l^{\varkappa} \end{matrix} \right\|, \quad \rho_z \boldsymbol{X}_l^{\varkappa} = \left\| \begin{matrix} \boldsymbol{X}_l^{\varkappa} & 0 \\ 0 & -\boldsymbol{X}_l^{\varkappa} \end{matrix} \right\|, \tag{26.29a}$$

which transform according to the same representation as $\boldsymbol{X}_l^{\varkappa}$. As interband matrices we then take

$$\rho_+ \boldsymbol{X}_l^{\varkappa} = \left\| \begin{matrix} 0 & \boldsymbol{X}_l^{\varkappa} \\ 0 & 0 \end{matrix} \right\|, \quad \rho_- \boldsymbol{X}_l^{\varkappa} = \left\| \begin{matrix} 0 & 0 \\ \boldsymbol{X}_l^{\varkappa} & 0 \end{matrix} \right\|, \tag{26.29b}$$

which transform according to the representations $\mathscr{D}_{\varkappa'} = \mathscr{D} \times \mathscr{D}_{\varkappa}$ occurring in $\mathscr{D}^{\mathrm{I}*} \times \mathscr{D}^{\mathrm{II}}$, and $\mathscr{D}_{\varkappa''} = \mathscr{D}^* \times \mathscr{D}_{\varkappa}$ occurring in $\mathscr{D}^{\mathrm{I}} \times \mathscr{D}^{\mathrm{II}*}$, respectively. The representations $\mathscr{D}_{\varkappa'}$ and $\mathscr{D}_{\varkappa''}$ are p-equivalent to the representation $\mathscr{D}_{\varkappa}$, but may be unitarily inequivalent if $\mathscr{D}^*(r) \neq \mathscr{D}(r)$. If $\mathscr{D}^* = \mathscr{D}$, we may define the interband matrices as

$$\rho_x \boldsymbol{X}_l^{\varkappa} = \left\| \begin{matrix} 0 & \boldsymbol{X}_l^{\varkappa} \\ \boldsymbol{X}_l^{\varkappa} & 0 \end{matrix} \right\|, \quad \rho_y \boldsymbol{X}_l^{\varkappa} = \left\| \begin{matrix} 0 & -i\boldsymbol{X}_l^{\varkappa} \\ i\boldsymbol{X}_l^{\varkappa} & 0 \end{matrix} \right\|. \tag{26.29c}$$

Moreover, one of these sets of matrices, say $\rho_x \boldsymbol{X}_l^{\varkappa}$, may be assumed to be even with respect to time reversal and the other, $\rho_y \boldsymbol{X}_l^{\varkappa}$, odd. This corresponds to a definite choice of phases for the basis functions.

The basis matrices (26.29) are particularly convenient when the representations $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$ are combined owing to time reversal. In case (c) the representations $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$ coincide, in case (b) they are p-equivalent. This choice of basis differs from (25.9); it corresponds to a unitary transformation of $\boldsymbol{\mathscr{D}}^{\mathrm{II}}$ in (25.9) under which $\boldsymbol{\mathscr{H}}^{\mathrm{II\,II}}(\mathscr{K})$ goes into $\tilde{\boldsymbol{\mathscr{H}}}^{\mathrm{II\,II}}(\mathscr{K})$. By (25.11) and (25.27), the matrices $\tilde{\boldsymbol{\mathscr{H}}}^{\mathrm{II\,II}}(\mathscr{K})$ and $\boldsymbol{\mathscr{H}}^{\mathrm{I\,I}}(\mathscr{K})$ must satisfy the following relations:

$$\begin{aligned} &\text{cases } (\mathrm{b}_1),\ (\mathrm{c}_1){:}\quad \tilde{\boldsymbol{\mathscr{H}}}^{\mathrm{II\,II}}(\mathscr{K}) = \boldsymbol{\mathscr{H}}^{\mathrm{I\,I}}(f\mathscr{K}), \\ &\text{cases } (\mathrm{b}_2),\ (\mathrm{c}_2){:}\quad \tilde{\boldsymbol{\mathscr{H}}}^{\mathrm{II\,II}}(\mathscr{K}) = \boldsymbol{\mathscr{H}}^{\mathrm{I\,I}}(fR^{-1}\mathscr{K}). \end{aligned} \tag{26.30}$$

It is clear that in the first case the diagonal matrices can always be chosen in the form (26.29a), but in the second this is possible only if the components of $\mathscr{K}$ are chosen so that $fR^{-1}\mathscr{K} = \pm \mathscr{K}$. The matrices $\boldsymbol{I}\boldsymbol{X}_l^{\varkappa}$ then appear in $\boldsymbol{\mathscr{H}}(\mathscr{K})$ with functions which do not change sign under the operations K or RK, and the matrices $\rho_z \boldsymbol{X}_l^{\varkappa}$ with functions which change sign, i.e., the former are even with respect to these operations and the latter odd. In case the representations $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$ are independent and no additional conditions are imposed on the off-diagonal parts of $\mathscr{H}(\mathscr{K})$, each function $\mathscr{K}_l^{\varkappa}$ appears in $\mathscr{H}(\mathscr{K})$ with both matrices with independent constants.

If the representations $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$ are p-equivalent to a vector representation and either coincide or differ only in parity, we can define the intraband matrices as the matrices $\boldsymbol{IJ}_i$ and $\boldsymbol{\rho}_z\boldsymbol{J}_i$ constructed by the method described above, which transform like the components of the pseudovector $\boldsymbol{J}_i$ (or their products), and the interband matrices as the matrices $\boldsymbol{\rho}_+\boldsymbol{J}_l$ and $\boldsymbol{\rho}_-\boldsymbol{J}_l$ or $\boldsymbol{\rho}_x\boldsymbol{J}_l$ and $\boldsymbol{\rho}_y\boldsymbol{J}_l$, which transform like the components of a pseudovector if $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$ have the same parity, or like the components of a polar vector otherwise. For one-dimensional representations, which are always p-equivalent, we can define the intraband matrices as the matrices $\boldsymbol{I}$ and $\boldsymbol{\rho}_z = \boldsymbol{\sigma}_z$, which transform according to the identity representation, and the interband matrices are $\boldsymbol{\rho}_+ = \boldsymbol{\sigma}_+$, which transforms according to the representation $\mathscr{D}^{\mathrm{I}*} \times \mathscr{D}^{\mathrm{II}}$, and $\boldsymbol{\rho}_- = \boldsymbol{\sigma}_-$, which transforms according to $\mathscr{D}^{\mathrm{I}} \times \mathscr{D}^{\mathrm{II}*}$. If the representations $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$ are real or coincide, we can take $\boldsymbol{\rho}_x = \boldsymbol{\sigma}_x$ and $\boldsymbol{\rho}_y = \boldsymbol{\sigma}_y$ as interband matrices.

For two-dimensional p-equivalent representations, the sixteen basis matrices may be chosen as

$$\begin{aligned} \boldsymbol{I\sigma}_l &= \begin{Vmatrix} \boldsymbol{\sigma}_l & 0 \\ 0 & \boldsymbol{\sigma}_l \end{Vmatrix}, \quad & \boldsymbol{\rho}_z\boldsymbol{\sigma}_l &= \begin{Vmatrix} \boldsymbol{\sigma}_l & 0 \\ 0 & -\boldsymbol{\sigma}_l \end{Vmatrix}, \\ \boldsymbol{\rho}_+\boldsymbol{\sigma}_l &= \begin{Vmatrix} 0 & \boldsymbol{\sigma}_l \\ 0 & 0 \end{Vmatrix}, \quad & \boldsymbol{\rho}_-\boldsymbol{\sigma}_l &= \begin{Vmatrix} 0 & 0 \\ \boldsymbol{\sigma}_l & 0 \end{Vmatrix}, \end{aligned} \tag{26.31}$$

where $\boldsymbol{\sigma}_l$ runs through the Pauli matrices, including the identity matrix. For real representations, $\boldsymbol{\rho}_+\boldsymbol{\sigma}_l$ and $\boldsymbol{\rho}_-\boldsymbol{\sigma}_l$ must be replaced by the "even" matrices $\boldsymbol{\rho}_x\boldsymbol{\sigma}_l$ and the "odd" matrices $\boldsymbol{\rho}_y\boldsymbol{\sigma}_l$. The matrices $\boldsymbol{I}$ and $\boldsymbol{\rho}_z$ then transform according to the identity representation, while $\boldsymbol{I\sigma}_l$ and $\boldsymbol{\rho}_z\boldsymbol{\sigma}_l$ transform according to the other constituents $\mathscr{D}_{\varkappa}$ of $\mathscr{D}^{\mathrm{I}} \times \mathscr{D}^{\mathrm{I}*} = \mathscr{D}^{\mathrm{II}} \times \mathscr{D}^{\mathrm{II}*}$. The matrices $\boldsymbol{\rho}_+$ and $\boldsymbol{\rho}_-$ transform according to $\mathscr{D}(r)$ and $\mathscr{D}^{\times}(r)$, the matrices $\boldsymbol{\rho}_+\boldsymbol{\sigma}_l$ and $\boldsymbol{\rho}_-\boldsymbol{\sigma}_l$ according to the representations $\mathscr{D}_{\varkappa} \times \mathscr{D}$ and $\mathscr{D}_{\varkappa} \times \mathscr{D}^*$ occurring in $\mathscr{D}^{\mathrm{I}} \times \mathscr{D}^{\mathrm{II}*}$ and $\mathscr{D}^{\mathrm{I}*} \times \mathscr{D}^{\mathrm{II}}$. In this case, $\mathscr{H}(\mathscr{K})$ may be constructed knowing only the characters of the representations $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$.

If the representations $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$ are not p-equivalent and at least one of them has dimension >2, determination of the matrices $\boldsymbol{X}_l^{\varkappa}$ requires knowledge either of the basis or of the representation matrices $\boldsymbol{\mathscr{D}}(g)$ for the generators.

Thus, if $\mathscr{D}^{\mathrm{I}}(r)$ and $\mathscr{D}^{\mathrm{II}}(r)$ are p-equivalent to vector or spinor representations, their basis functions may be chosen as superpositions of the functions Y^j_m; the values j^{I} and j^{II} of j for the representations $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$ are generally different. Here the interband matrices $\boldsymbol{X}_l^{\varkappa}$ may be defined as matrices $\boldsymbol{A}_m$, which transform in the general case like components of a tensor of suitable rank.

The general expressions for the nonzero elements of the corresponding matrices are given, for example, in /I.7/ (see §29). Thus, if $j^{\mathrm{I}} = j^{\mathrm{II}}$, these elements are defined by (10.17), up to arbitrary constants $C_{\mathrm{I}}^{\mathrm{II}} = C_{\mathrm{II}}^{\mathrm{I}*}$. If $|j^{\mathrm{I}} - j^{\mathrm{II}}| = 1$, the matrices $\boldsymbol{R}_m$, which transform like components of an axial vector if the representations $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$ have the same parity or like components of a polar vector otherwise, are defined by

$$\begin{aligned} (R_z)^{jm}_{j-1,\, m} &= \left[(R_z)^{j-1,\, m}_{jm}\right]^{*} = C^{j}_{j-1}\left(j^2 - m^2\right)^{1/2}, \\ (R_-)^{j,\, m-1}_{j-1,\, m} &= \left[(R_+)^{j-1,\, m}_{j,\, m-1}\right]^{*} = C^{j}_{j-1}\left[(j - m + 1)(j - m)\right]^{1/2}, \\ (R_+)^{jm}_{j-1,\, m-1} &= \left[(R_-)^{j-1,\, m-1}_{jm}\right]^{*} = -\, C^{j*}_{j-1}\left[(j + m - 1)(j + m)\right]^{1/2}. \end{aligned} \qquad (26.32)$$

For any two sets of matrices $\boldsymbol{R}_m$, the constants $C^{\mathrm{II}}_{\mathrm{I}} = C^{\mathrm{I}*}_{\mathrm{II}}$ or $C^{j}_{j-1} = (C^{j-1}_{j})^{*}$ must be chosen so that these matrices are linearly independent; it is sufficient to stipulate that the matrices of one set be real and those of the other pure imaginary, setting

$$C^{\mathrm{II}(1)}_{\mathrm{I}} = C^{\mathrm{I}(1)}_{\mathrm{II}} = 1, \quad C^{\mathrm{II}(2)}_{\mathrm{I}} = -\, C^{\mathrm{I}(2)}_{\mathrm{II}} = -\, i. \qquad (26.33)$$

As an example, Table 26.5 presents the matrices $\boldsymbol{R}_x$, $\boldsymbol{R}_y$, $\boldsymbol{R}_z$ relative to the basis Y^0, Y^1_m. These matrices form bases for pairs of representations: A and F of the group T, A_1 or A_2 and F_1 or F_2 of the group T_d, and $A^{\pm}_1$ and $F^{\pm}_1$ (or the p-equivalent pairs $A^{\pm}_2$ and $F^{\pm}_2$) of the groups O and O_h. Suitable products of these matrices yield all the remaining sixteen matrices, including the matrices $\{\boldsymbol{R}_x \boldsymbol{R}_y\}$ which transform like $\boldsymbol{J}_z$, etc., and coincide with the corresponding matrices $\boldsymbol{J}_z$, $\boldsymbol{J}_x$, $\boldsymbol{J}_y$ of Table 26.3 up to constants.

Allowance for spin in the two-band model may be achieved via either of the two approaches cited above. In the first approach, the irreducible spinor representations play the role of the representations $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$, in the second – the representations $\mathscr{D}^{\mathrm{I}}_{l} \times \mathscr{D}_{1/2}$ and $\mathscr{D}^{\mathrm{II}}_{l} \times \mathscr{D}_{1/2}$. The components $\boldsymbol{\sigma}_i$ are then incorporated in $\mathscr{H}$ in accordance with the above rules.

As an example, we consider the two-band model for classes O_h and T_d (the point Γ).

The simplest case is that both representations $\mathscr{D}^{\mathrm{I}}$ and $\mathscr{D}^{\mathrm{II}}$ are two-dimensional; here, when spin is included, $\mathscr{H}(\boldsymbol{k})$ contains interband linear k terms only for the pair of representations Γ_6 and Γ_7, and in the O_h lattice for the pairs Γ^{+}_{6}, Γ^{-}_{7} and Γ^{-}_{6}, Γ^{+}_{7}. In this case $\mathscr{H}(\boldsymbol{k})$ will contain interband terms which transform according to $\Gamma^{\pm}_{6} \times \Gamma^{\mp}_{7} = \Gamma^{-}_{2} + \Gamma^{-}_{15}$, even intraband terms which transform according to $\{\Gamma^{2}_{6}\} = \{\Gamma^{2}_{7}\} = \Gamma^{+}_{1}$ and odd intraband terms which transform according to $[\Gamma^{2}_{6}] = [\Gamma^{2}_{7}] = \Gamma^{+}_{25}$. Since these representations are p-equivalent: $\Gamma_6 = \Gamma_7 \Gamma_2$, the even intraband matrices are necessarily $\boldsymbol{I}$ and $\boldsymbol{\rho}_z$, the odd ones $\boldsymbol{\sigma}_i$ and $\boldsymbol{\rho}_z \boldsymbol{\sigma}_i$. The interband matrices $\boldsymbol{\rho}_x$ and $\boldsymbol{\rho}_y$ transform according to Γ_2, while $\boldsymbol{\rho}_x \boldsymbol{\sigma}_i$ and $\boldsymbol{\rho}_y \boldsymbol{\sigma}_i$ transform according to Γ_{15}. We shall assume that the matrices $\boldsymbol{\rho}_x \boldsymbol{\sigma}_i$ are even, the matrices $\boldsymbol{\rho}_y \boldsymbol{\sigma}_i$ odd. As noted above, this choice of basis functions makes the functions $\mathscr{K}^{\varkappa}_{l}$ real. These functions are given in Table 26.2.

For crystals of class O_h

$$\mathscr{H}(\boldsymbol{k}) = \boldsymbol{\rho}_z \frac{E_g}{2} + \frac{\hbar k^2}{2m_1} + \boldsymbol{\rho}_z \frac{\hbar^2 k^2}{2m_2} + A \boldsymbol{\rho}_y (\boldsymbol{\sigma}_x k_x + \boldsymbol{\sigma}_y k_y + \boldsymbol{\sigma}_x k_z). \qquad (26.34)$$

For crystals of class T_d, the Hamiltonian $\mathscr{H}(\boldsymbol{k})$ will include the additional term

$$\mathscr{H}''(\boldsymbol{k}) = a \boldsymbol{\rho}_x (\boldsymbol{\sigma}_x k_y k_z + \boldsymbol{\sigma}_y k_x k_z + \boldsymbol{\sigma}_z k_x k_y), \qquad (26.35)$$

since in T_d the components k_i and products $k_i k_j$ transform according to the same representation Γ_{15}. The spectrum defined by (26.34) is

$$E(\boldsymbol{k}) = \frac{\hbar^2 k^2}{2m_1} \pm \left\{ \left(\frac{E_g}{2} + \frac{\hbar^2 k^2}{2m_2} \right)^2 + A^2 k^2 \right\}^{1/2}. \qquad (26.36)$$

If $1/m_1 = 1/m_2 = 0$, the equation $\mathcal{H}\psi = i\hbar\frac{\partial\psi}{\partial t}$ coincides in form with the Dirac equation, with the speed of light replaced by $s = A/\hbar$ and m by $m^* = E_g/2s^2$. In this case the spectrum near the band edges is

$$E(\boldsymbol{k}) = \pm\left(\frac{E_g}{2} + \frac{\hbar^2k^2}{2m^*}\right).$$

The analogy between the two-band equation and the Dirac equation is frequently useful. For instance, it follows at once from this analogy that the two-band equation (26.34) for $1/m_1 = 1/m_2 = 0$ is invariant under the transformation obtained from the Lorentz transformation of relativity theory when c is replaced by s.

The constant A in (26.34) is relativistically small if the representations Γ_6 and Γ_7 originate in nondegenerate representations Γ_1 and Γ_2. If one of these representations is generated by splitting of the terms $\Gamma_{15}\times\mathcal{D}_{1/2} = \Gamma_7 + \Gamma_8$ or $\Gamma_{25}\times\mathcal{D}_{1/2} = \Gamma_6 + \Gamma_8$, and the other from Γ_1 or Γ_2, respectively, this constant is not small and in this case, in the nonrelativistic approximation, the interband elements for the pairs Γ_{15}, Γ_1 or Γ_{25}, Γ_2 will also include terms linear in k.

If it is necessary to construct $\mathcal{H}(\boldsymbol{k})$ for all terms occurring in these pairs of representations, we can take as basis the matrices $\boldsymbol{R}_m$ in the representation $\mathcal{D}_0 + \mathcal{D}_1$, as shown in Table 26.5: six interband matrices $\boldsymbol{R}_x, \boldsymbol{R}_y, \boldsymbol{R}_z$ (even) and $\boldsymbol{R}'_x, \boldsymbol{R}'_y, \boldsymbol{R}'_z$, and ten intraband matrices $\boldsymbol{I}, \boldsymbol{R}^2_x, \boldsymbol{R}^2_y, \boldsymbol{R}^2_z, \{\boldsymbol{R}_x\boldsymbol{R}_y\}, \{\boldsymbol{R}_x\boldsymbol{R}_z\}, \{\boldsymbol{R}_y\boldsymbol{R}_z\}$, $[\boldsymbol{R}_x\boldsymbol{R}_y], [\boldsymbol{R}_y\boldsymbol{R}_z], [\boldsymbol{R}_z\boldsymbol{R}_x]$. If we include only nonrelativistic terms independent of $\boldsymbol{k}$, the Hamiltonian $\mathcal{H}(\boldsymbol{k})$ for a T_d lattice will be

$$\begin{aligned}\mathcal{H}(\boldsymbol{k}) = \Delta_1 R^2 + 2i\Delta_2\sum_i \sigma_i\{\boldsymbol{R}_{i+1}\boldsymbol{R}_{i+2}\} + A_1\boldsymbol{R}^2k^2 + A_2\sum_i \boldsymbol{R}_i^2k_i^2 + \\ + A_3\sum_{i\neq j}[\boldsymbol{R}_i\boldsymbol{R}_j]k_ik_j + A_4\sum_i \boldsymbol{R}'_ik_i + A_5\sum_i \boldsymbol{R}_ik_{i+1}k_{i+2}.\end{aligned} \tag{26.37}$$

For an O_h lattice the last term is dropped. The matrix (26.37) can be brought to the form of Table 24.2 (up to constants) by a unitary transformation diagonalizing the second term in $\mathcal{H}(\boldsymbol{k})$ (the spin-orbit coupling).

TABLE 26.3. The matrices $\boldsymbol{J}_i$ and their products

Representation D_1, basis x, y, z

$\boldsymbol{J}_x = \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{Vmatrix}$	$\boldsymbol{J}_y = \begin{Vmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{Vmatrix}$	$\boldsymbol{J}_z = \begin{Vmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}$
$\boldsymbol{J}_x^2 = \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}$	$\boldsymbol{J}_y^2 = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{Vmatrix}$	$\boldsymbol{J}_z^2 = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{Vmatrix}$
$2[\boldsymbol{J}_x\boldsymbol{J}_y] = \begin{Vmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}$	$2[\boldsymbol{J}_y\boldsymbol{J}_z] = \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{Vmatrix}$	$2[\boldsymbol{J}_x\boldsymbol{J}_z] = \begin{Vmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{Vmatrix}$

Representation $\mathcal{D}_{3/2}$, basis $Y_m^{3/2}$ (23.2)

$$J_x = \begin{Vmatrix} 0 & i\frac{\sqrt{3}}{2} & 0 & 0 \\ -i\frac{\sqrt{3}}{2} & 0 & i & 0 \\ 0 & -i & 0 & i\frac{\sqrt{3}}{2} \\ 0 & 0 & -\frac{i\sqrt{3}}{2} & 0 \end{Vmatrix} \qquad J_y = \begin{Vmatrix} 0 & \frac{\sqrt{3}}{2} & 0 & 0 \\ \frac{\sqrt{3}}{2} & 0 & 1 & 0 \\ 0 & 1 & 0 & \frac{\sqrt{3}}{2} \\ 0 & 0 & \frac{\sqrt{3}}{2} & 0 \end{Vmatrix} \qquad J_z = \begin{Vmatrix} \frac{3}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{3}{2} \end{Vmatrix}$$

$$J_x^2 = \begin{Vmatrix} \frac{3}{4} & 0 & -\frac{\sqrt{3}}{2} & 0 \\ 0 & \frac{7}{4} & 0 & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & 0 & \frac{7}{4} & 0 \\ 0 & -\frac{\sqrt{3}}{2} & 0 & \frac{3}{4} \end{Vmatrix} \qquad J_y^2 = \begin{Vmatrix} \frac{3}{4} & 0 & \frac{\sqrt{3}}{2} & 0 \\ 0 & \frac{7}{4} & 0 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 0 & \frac{7}{4} & 0 \\ 0 & \frac{\sqrt{3}}{2} & 0 & \frac{3}{4} \end{Vmatrix} \qquad J_z^2 = \begin{Vmatrix} \frac{9}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{9}{4} \end{Vmatrix}$$

$$[J_xJ_y] = \begin{Vmatrix} 0 & 0 & i\frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 0 & i\frac{\sqrt{3}}{2} \\ -i\frac{\sqrt{3}}{2} & 0 & 0 & 0 \\ 0 & -i\frac{\sqrt{3}}{2} & 0 & 0 \end{Vmatrix} \qquad [J_xJ_z] = \begin{Vmatrix} 0 & i\frac{\sqrt{3}}{2} & 0 & 0 \\ -i\frac{\sqrt{3}}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -i\frac{\sqrt{3}}{2} \\ 0 & 0 & i\frac{\sqrt{3}}{2} & 0 \end{Vmatrix} \qquad [J_yJ_z] = \begin{Vmatrix} 0 & \frac{\sqrt{3}}{2} & 0 & 0 \\ \frac{\sqrt{3}}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{\sqrt{3}}{2} \\ 0 & 0 & -\frac{\sqrt{3}}{2} & 0 \end{Vmatrix}$$

$$J_x^3 = \left\| \begin{matrix} 0 & i\,\frac{7\sqrt{3}}{8} & 0 & -i\,\frac{3}{4} \\ -i\,\frac{7\sqrt{3}}{8} & 0 & i\,\frac{10}{4} & 0 \\ 0 & -i\,\frac{10}{4} & 0 & i\,\frac{7\sqrt{3}}{8} \\ i\,\frac{3}{4} & 0 & -i\,\frac{7\sqrt{3}}{8} & 0 \end{matrix} \right\| \quad J_y^3 = \left\| \begin{matrix} 0 & \frac{7\sqrt{3}}{8} & 0 & \frac{3}{4} \\ \frac{7\sqrt{3}}{8} & 0 & \frac{10}{4} & 0. \\ 0 & \frac{10}{4} & 0 & \frac{7\sqrt{3}}{8} \\ \frac{3}{4} & 0 & \frac{7\sqrt{3}}{8} & 0 \end{matrix} \right\| \quad J_z^3 = \left\| \begin{matrix} \frac{27}{8} & 0 & 0 & 0 \\ 0 & \frac{1}{8} & 0 & 0 \\ 0 & 0 & -\frac{1}{8} & 0 \\ 0 & 0 & 0 & -\frac{27}{8} \end{matrix} \right\|$$

$$2V_x = \left\| \begin{matrix} 0 & -i\,\frac{\sqrt{3}}{2} & 0 & i\,\frac{3}{2} \\ i\,\frac{\sqrt{3}}{2} & 0 & i\,\frac{3}{2} & 0 \\ 0 & -i\,\frac{3}{2} & 0 & -i\,\frac{\sqrt{3}}{2} \\ -i\,\frac{3}{2} & 0 & i\,\frac{\sqrt{3}}{2} & 0 \end{matrix} \right\| \quad 2V_y = \left\| \begin{matrix} 0 & \frac{\sqrt{3}}{2} & 0 & \frac{3}{2} \\ \frac{\sqrt{3}}{2} & 0 & -\frac{3}{2} & 0 \\ 0 & -\frac{3}{2} & 0 & \frac{\sqrt{3}}{2} \\ \frac{3}{2} & 0 & \frac{\sqrt{3}}{2} & 0 \end{matrix} \right\| \quad 2V_z = \left\| \begin{matrix} 0 & 0 & -\sqrt{3} & 0 \\ 0 & 0 & 0 & \sqrt{3} \\ -\sqrt{3} & 0 & 0 & 0 \\ 0 & \sqrt{3} & 0 & 0 \end{matrix} \right\|$$

$$2\,[J_x\,[J_yJ_z]] = \left\| \begin{matrix} 0 & 0 & i\,\frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 0 & -i\,\frac{\sqrt{3}}{2} \\ -i\,\frac{\sqrt{3}}{2} & 0 & 0 & 0 \\ 0 & \frac{i\sqrt{3}}{2} & 0 & 0 \end{matrix} \right\|$$

$$2\,[J_iJ_k] = J_iJ_k + J_kJ_i, \qquad V_x = \left[J_x\left(J_y^2 - J_z^2\right)\right]$$

TABLE 26.4. Basis matrices for projective representations

Group	Class	Representation	X	Y	Z
D_{4h}	K_5	$Q^{(5)}$	$\begin{Vmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{Vmatrix}$	$\begin{Vmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{Vmatrix}$	$\begin{Vmatrix} \sigma_x & 0 \\ 0 & -\sigma_x \end{Vmatrix}$
			$\begin{Vmatrix} \sigma_y & 0 \\ 0 & \sigma_y \end{Vmatrix}$	$-\begin{Vmatrix} \sigma_z & 0 \\ 0 & \sigma_z \end{Vmatrix}$	
	K_6	$Q^{(6)}$	$\begin{Vmatrix} \sigma_y & 0 \\ 0 & -\sigma_y \end{Vmatrix}$	$\begin{Vmatrix} -\sigma_z & 0 \\ 0 & \sigma_z \end{Vmatrix}$	$\begin{Vmatrix} \sigma_x & 0 \\ 0 & -\sigma_x \end{Vmatrix}$
			$\begin{Vmatrix} 0 & -i\sigma_y \\ i\sigma_y & 0 \end{Vmatrix}$	$\begin{Vmatrix} 0 & -i\sigma_z \\ i\sigma_z & 0 \end{Vmatrix}$	
D_{6h}	K_3	$Q^{(3)}$	$\begin{Vmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{Vmatrix}$	$-\begin{Vmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{Vmatrix}$	$\begin{Vmatrix} \sigma_z & 0 \\ 0 & -\sigma_z \end{Vmatrix}$
	K_4	$Q^{(4)}$	$\begin{Vmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{Vmatrix}$	$\begin{Vmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{Vmatrix}$	$\begin{Vmatrix} \sigma_z & 0 \\ 0 & -\sigma_z \end{Vmatrix}$
	K_5	$Q^{(5)}$	$\begin{Vmatrix} \sigma_x & 0 \\ 0 & -\sigma_x \end{Vmatrix}$	$\begin{Vmatrix} -\sigma_y & 0 \\ 0 & \sigma_y \end{Vmatrix}$	$\begin{Vmatrix} \sigma_z & 0 \\ 0 & -\sigma_z \end{Vmatrix}$
	K_6	$Q^{(6)}$	$\begin{Vmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{Vmatrix}$	$-\begin{Vmatrix} \sigma_y & 0 \\ 0 & \sigma_y \end{Vmatrix}$	$\begin{Vmatrix} \sigma_z & 0 \\ 0 & -\sigma_z \end{Vmatrix}$

TABLE 26.4 (continued)

Group	Class	Representation	X	Y	Z
O_h	K_2	$R^{(2)}$	$\begin{Vmatrix} \varphi & 0 \\ 0 & -\varphi \end{Vmatrix}, \quad \varphi = \begin{Vmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}$	$\begin{Vmatrix} \gamma & 0 \\ 0 & -\gamma \end{Vmatrix}, \quad \gamma = \begin{Vmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{Vmatrix}$	$\begin{Vmatrix} \varkappa & 0 \\ 0 & -\varkappa \end{Vmatrix}, \quad \varkappa = \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & i \\ 0 & -i & 0 \end{Vmatrix}$
			$\begin{Vmatrix} 0 & \lambda \\ \lambda & 0 \end{Vmatrix}, \quad \lambda = \begin{Vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}$	$\begin{Vmatrix} 0 & \delta \\ \delta & 0 \end{Vmatrix}, \quad \delta = \begin{Vmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{Vmatrix}$	$\begin{Vmatrix} 0 & \eta \\ \eta & 0 \end{Vmatrix}, \quad \eta = \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{Vmatrix}$
	K_3	$Q_1^{(3)}$	$\begin{Vmatrix} \sigma_x & 0 \\ 0 & -\sigma_x \end{Vmatrix}$	$\begin{Vmatrix} \sigma_y & 0 \\ 0 & -\sigma_y \end{Vmatrix}$	$\begin{Vmatrix} \sigma_z & 0 \\ 0 & -\sigma_z \end{Vmatrix}$
		$Q_2^{(3)}$ $Q_3^{(3)}$	$\begin{Vmatrix} 0 & \varepsilon_6 \sigma_x \\ \varepsilon_6^{-1} \sigma_x & 0 \end{Vmatrix}, \quad \varepsilon_6 = e^{\frac{2\pi i}{6}}$	$\begin{Vmatrix} 0 & \varepsilon_6^{-1} \sigma_y \\ \varepsilon_6 \sigma_y & 0 \end{Vmatrix}$	$\begin{Vmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{Vmatrix}$

For groups D_{nh}, the z-axis lies along c_n, the x-axis along u_2.

TABLE 26.5. Matrices of components of polar vector in representation $(Y^0, Y^1_1, Y^1_0, Y^1_1)$

$$\boldsymbol{R}_x = \left\| \begin{matrix} 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 \end{matrix} \right\| \qquad \boldsymbol{R}_y = \left\| \begin{matrix} 0 & -\frac{i}{\sqrt{2}} & 0 & -\frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{i}{\sqrt{2}} & 0 & 0 & 0 \end{matrix} \right\| \qquad \boldsymbol{R}_z = \left\| \begin{matrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} \right\|$$

$$\boldsymbol{R}'_x = \left\| \begin{matrix} 0 & -\frac{i}{\sqrt{2}} & 0 & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{i}{\sqrt{2}} & 0 & 0 & 0 \end{matrix} \right\| \qquad \boldsymbol{R}'_y = \left\| \begin{matrix} 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 \end{matrix} \right\| \qquad \boldsymbol{R}'_z = \left\| \begin{matrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} \right\|$$

§27. SHALLOW IMPURITY CENTERS AND EXCITONS IN SEMICONDUCTORS

Shallow impurity centers

As in §22, to determine the wave functions and energy E of a shallow impurity center in effective mass theory we must solve the system of equations (22.26):

$$\sum_{s} \mathcal{H}_{mm'}(\boldsymbol{k}) f_{m'} = E f_m, \tag{27.1}$$

or, in matrix notation,

$$\mathcal{H}f = Ef, \text{ where } f = \begin{Vmatrix} f_1 \\ f_2 \\ \vdots \\ f_N \end{Vmatrix}.$$

Here m and m' run from 1 to N, where N is the band degeneracy at the extremum, and

$$\mathcal{H}(\boldsymbol{k}) = \mathcal{H}_0(\boldsymbol{k}) + U(\boldsymbol{r})\boldsymbol{I}. \tag{27.2}$$

$\mathcal{H}_0(\boldsymbol{k})$ is the matrix defining the spectrum near the band extremum, $k_a = -i\frac{\partial}{\partial r_a}$, $U = -Ze^2/\varkappa r$ is the Coulomb potential, $\varkappa$ the dielectric constant,* Z the impurity charge.**

Each solution f^s of system (27.1), which is a column matrix with components f^s_m, determines an impurity center wave function:

$$\Psi^s = \sum_{m=1}^{N} f^s_m \psi_{m\boldsymbol{k}_0}, \tag{27.3}$$

where $\psi_{m\boldsymbol{k}_0}$ are the Bloch functions at the bottom of the band. Let us consider how the impurity center wave functions f transform.

By (25.1b),

$$\mathcal{D}^{-1}(g)\mathcal{H}(g\boldsymbol{k})\mathcal{D}(g) = \mathcal{H}(\boldsymbol{k}), \tag{27.4}$$

where $\mathcal{D}(g)$ is the transformation matrix of the functions $\psi_{m\boldsymbol{k}_0}$ at the bottom of the band. If the crystallographic point group contains elements g with nonprimitive translations, then in general $U(g\boldsymbol{r}) \neq U(\boldsymbol{r})$. However, at large distances (compared to the lattice constant), when the effective mass approximation is valid, we may assume that $U(g\boldsymbol{r}) = U(\boldsymbol{r})$. Therefore, each operator (27.1) satisfies (27.4).

Suppose there are μ linearly independent solutions of equation (27.1), f^s $(s = 1, 2, \ldots, \mu)$. Using (27.4) and the equation

$$\mathcal{H}(g^{-1}\boldsymbol{k}) f^s(g^{-1}\boldsymbol{r}) = E f^s(g^{-1}\boldsymbol{r}),$$

---

[*] With an eye to our subsequent discussion of impurity centers in cubic crystals, we shall assume that $\varkappa$ is a scalar.

[**] In the theory of excitons discussed in this section, the spin-dependence of ψ is significant; here, therefore, as opposed to §22, we shall assume from the start that the basis functions depend also on spin. Therefore, the space coordinates will be denoted throughout by $\boldsymbol{r}$, while $\boldsymbol{x}$ will include the spin indices.

we obtain the equality

$$\mathcal{H}(\boldsymbol{k})\boldsymbol{\mathcal{D}}(g)\boldsymbol{f}^s(g^{-1}\boldsymbol{r}) = E\boldsymbol{\mathcal{D}}(g)\boldsymbol{f}^s(g^{-1}\boldsymbol{r}),$$

which implies that the function $\boldsymbol{\mathcal{D}}(g)\boldsymbol{f}^s(g^{-1}\boldsymbol{r})$ with components

$$(\boldsymbol{\mathcal{D}}(g)\boldsymbol{f}^s(g^{-1}\boldsymbol{r}))_m = \sum_{m'} \mathcal{D}_{m'm}(g) f^s_{m'}(g^{-1}\boldsymbol{r}) \tag{27.5}$$

corresponds to the same energy E as the functions $\boldsymbol{f}^s(\boldsymbol{r})$ $(s = 1, 2, \ldots, \mu$. Hence there is a linear relation between these functions:

$$\boldsymbol{\mathcal{D}}(g)\boldsymbol{f}^s(g^{-1}\boldsymbol{r}) = \sum_{s'} T^{s's}(g)\boldsymbol{f}^{s'}(\boldsymbol{r}), \tag{27.6}$$

whence

$$\begin{gathered} \boldsymbol{f}^s(g^{-1}\boldsymbol{r}) = \boldsymbol{\mathcal{D}}^{-1}(g)\sum_{s'} T^{s's}(g)\boldsymbol{f}^{s'}(\boldsymbol{r}), \\ f^s_m(g^{-1}\boldsymbol{r}) = \sum_{s'm'} T^{s's}(g)\mathcal{D}^{*}_{m'm}(g) f^{s'}_{m'}(\boldsymbol{r}). \end{gathered} \tag{27.7}$$

Thus, the transformation of the functions $f^s_m(g^{-1}\boldsymbol{r})$ involves both a linear transformation of the columns $\boldsymbol{f}^s$ into each other, defined by the matrix $\boldsymbol{T}$, and a transformation of the components within each column, defined by the matrix $\boldsymbol{\mathcal{D}}$.

A direct check readily shows that the matrices $\boldsymbol{T}(g)$ form representations of the little group $G_{\boldsymbol{k}_0}$, defining a transformation of the functions Ψ^s (27.3):

$$\Psi^s(g^{-1}\boldsymbol{r}) = \sum_{s'} T^{s's}(g)\Psi^{s'}(\boldsymbol{r}). \tag{27.8}$$

Thus, the wave functions Ψ (27.3) of an impurity center near $\boldsymbol{k}_0$ transform according to a representation of the little group $G_{\boldsymbol{k}_0}$.

It follows from (27.7) that the functions $\boldsymbol{f}$ transform according to a representation $\mathcal{D}_{\boldsymbol{f}}$ of the point group:

$$\mathcal{D}_{\boldsymbol{f}} = \mathcal{D}^{*} \times T, \tag{27.9}$$

where $\mathcal{D}$ is the representation of $G_{\boldsymbol{k}_0}$ according to which the functions at the bottom of the band transform, and T the representation according to which the impurity center wave function transforms. This equation specifies the representations according to which the envelope functions may transform for a given symmetry T of the total wave function.

Since the potential energy in equation (27.1) is minimal near the impurity center, we should expect the ground state of the impurity center to belong to functions $\boldsymbol{f}$ which are finite at the origin; in other words, the set of functions $\boldsymbol{f}$ of the ground state must contain functions which transform according to the identity representation of $G_{\boldsymbol{k}_0}$. It follows from (27.9) that the representation $\boldsymbol{\mathcal{D}}_{\boldsymbol{f}}$ will contain the identity representation only if $\mathcal{D} = T$. We may therefore expect the degeneracy of the ground state of a shallow impurity center to coincide with the band degeneracy at the extremum point; this has been confirmed by experiment in all cases investigated to date.

Until now we have considered the state of an impurity center near an extremum point $\boldsymbol{k}_0$. If there are several equivalent extrema $\boldsymbol{k}_l$, $l = 1, 2, \ldots, n$, an additional degeneracy of the shallow impurity level appears, since in effective mass theory the wave function Ψ of the center is an arbitrary linear combination of functions Ψ_l with the same energy E (27.3), corresponding to different extrema:

$$\Psi = \sum_{l=1}^{n} c_l \Psi_l. \tag{27.10}$$

Therefore, in effective mass theory the degeneracy of the ground state of a shallow impurity level equals the band degeneracy at an extremum point multiplied by the number of equivalent extrema, so that it is simply the dimension of the irreducible representation of the group with star $\{\boldsymbol{k}_0\}$.

However, this high degeneracy is rooted in the effective mass approximation and is not a consequence of symmetry. In fact, the degeneracy of the impurity center ground state is generally lower, since the possible degeneracy of a term is determined by the dimension of the irreducible representations of the local symmetry point group at the site of the impurity atom. Therefore, the additional degeneracy due to the effective mass approximation is removed by interactions which are neglected in effective mass theory; these will be discussed below.

In effective mass theory, different impurity atoms with the same effective charge have identical energy spectra, since according to (27.1) the spectrum depends only on the band parameters near the extremum and on the dielectric constant of the crystal. The corrections to effective mass theory, which alter the ground state energy and lead to a splitting of the degeneracy characteristic for the many valley model (chemical shift*), are different for unlike impurity atoms.

If the chemical shift is significantly less than the ground state energy (measured from the bottom of the band), the correct wave functions are certain linear combinations of the functions (27.10), and the nature of the splitting and the functions themselves may be derived by group-theoretic considerations, without a detailed investigation of the chemical shift mechanism. For this purpose we must expand the representation $\mathscr{D}^{(\boldsymbol{k}_0)}$ of the space group according to which the functions (27.10) transform in terms of irreducible representations of the local symmetry point group G_i at the site of the impurity ion. Note that only rotational elements of the space group which do not occur together with nonprimitive translations may appear in the local symmetry group.

For any element $g \in G_i$ the character of the representation $\mathscr{D}^{(\boldsymbol{k}_0)}$ of the space group is (see (12.24))

$$\chi(g) = \sum_l \chi_l(g) = \sum_l \chi(g_l g g_l^{-1}). \tag{27.11}$$

Here χ is the character of the representation according to which the wave functions Ψ_1 (27.3) of the impurity center at $\boldsymbol{k}_1$ transform, and g_l is an element taking $\boldsymbol{k}_1$ into $\boldsymbol{k}_l$, $g_l \boldsymbol{k}_1 = \boldsymbol{k}_l$. The summation extends over all points of the star $\{\boldsymbol{k}_0\}$ in the symmetry group $G_{\boldsymbol{k}_l}$ which contain g.

If the wave functions Ψ_1 at the extremum point transform according to the identity representation, then $\chi_l(g) = 1$, and it follows from (27.10) that in this case the character of the space group representation is

$$\chi(g) = N_g, \tag{27.12}$$

where N_g is the number of points of the star which are fixed under g.

---

* In the Western literature, this splitting, which involves a transition from one valley to another, is also known as "valley-orbit splitting."

Character theory thus makes it possible to determine the splitting of the impurity level; on the other hand, using projection operators it is not difficult directly to determine the linear combinations of the functions (27.10) which correspond to a state with one energy.

Note that in the effective mass approximation the impurity center wave function in a crystal possessing a center of inversion is always even or odd. However, if the local symmetry group does not contain inversion, the total impurity center wave function does not possess a definite parity when corrections are introduced in the effective mass equations.

We now consider the wave functions of shallow impurity centers at a few points of the Brillouin zone in cubic crystals.

Spherical band at $\boldsymbol{k}_0=0$. In this case equation (27.1) has the form

$$\left(\frac{\hbar^2}{2m^*}\Delta+\frac{e^2}{\varkappa r}+E\right)f=0. \tag{27.13}$$

Equation (27.13) is simply the Schrödinger equation for the hydrogen atom with m replaced by m^* and e^2 by $e^2/\varkappa$. Such centers are said to be hydrogenic. The ground state of a hydrogenic center is not degenerate; the corresponding solution is $f_0=\frac{1}{\sqrt{\pi a_B^3}}\exp(-r/a_B)$, where $a_B=\hbar^2\varkappa/m^*e^2$ is the Bohr radius of the impurity center. The ground state energy for a hydrogenic center is

$$E_0=-\frac{m^*e^4}{2\hbar^2\varkappa^2}.$$

The energy of the excited states is $E_n=-E_0/n^2$, where n is the principal quantum number; $n=n_r+l+1$, n_r is the radial quantum number and l the azimuthal quantum number. The parity of hydrogenic states is $P=(-1)^l$.

Impurity center in the many-valley model. Suppose that the band extrema lie at a point $k_0\neq 0$ on the symmetry axes, so that the constant energy surfaces are ellipsoids of revolution:

$$E(\boldsymbol{k})=\frac{\hbar^2k_z^2}{2m_\parallel}+\frac{\hbar^2\left(k_x^2+k_y^2\right)}{2m_\perp}, \tag{27.14}$$

with the z-axis directed along the axis of revolution of the ellipsoid. In this case equation (27.1) has the form

$$\left(\frac{\hbar^2}{2m_\parallel}\frac{\partial^2}{\partial z^2}+\frac{\hbar^2}{2m_\perp}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)+\frac{e^2}{\varkappa r}+E\right)f=0. \tag{27.15}$$

We introduce a dimensionless parameter $\gamma=m_\perp/m_\parallel$, dimensionless coordinates measured in Bohr radii $a_B=\hbar^2\varkappa/m_\perp e^2$, and dimensionless energy in units of $E_B=m_\perp e^4/2\hbar^2\varkappa^2$. Then equation (27.15) is rewritten as

$$\left(\gamma\frac{\partial^2}{\partial z^2}+\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{2}{r}+E\right)f=0. \tag{27.16}$$

In the general case the variables in this equation are not separable and we cannot solve in closed form, except in the case $\gamma=1$ discussed above, and in the "adiabatic" case $\gamma\ll 1$, in which a series solution may be obtained in powers of the small parameter γ /21.2/.

For the ground state energy, one obtains good results by a direct variational method, using a trial function

$$f = (\pi a_\perp^2 a_\parallel)^{-1/2} \exp\left\{ -\left(\frac{\rho^2}{a_\perp^2} + \frac{z^2}{a_\parallel^2}\right)^{1/2}\right\}, \tag{27.17}$$

where $\rho^2 = x^2 + y^2$, $a_\parallel$ and $a_\perp$ are variational parameters defined by the condition that they minimize the functional

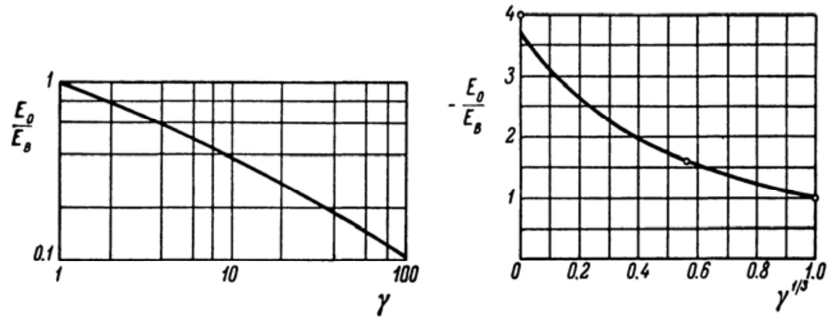
$$E = \int f \mathscr{H} f \, dr = \frac{1}{3a_\perp^2}\left(2 + \frac{\gamma}{1+\alpha^2}\right) - \frac{2}{a_\perp}\frac{\operatorname{arcsinh}\alpha}{\alpha}, \quad a_\parallel = a_\perp(1+\alpha^2). \tag{27.18}$$

Minimization of this functional with the trial function (27.17) yields the following equations for $a_\perp$ and α:

$$a_\perp = \frac{1}{3}\left(2 + \frac{\gamma}{1+\alpha^2}\right)\frac{\alpha}{\operatorname{arcsinh}\alpha}, \qquad \gamma = 2(1+\alpha^2)^{3/2}\frac{\sqrt{1+\alpha^2}\operatorname{arcsinh}\alpha - \alpha}{\alpha\sqrt{1+\alpha^2} - \operatorname{arcsinh}\alpha}. \tag{27.19}$$

Equations (27.18) and (27.19) refer to the case $\gamma > 1$; if $\gamma < 1$, $1+\alpha^2$ is replaced by $1-\alpha^2$ and arcsinh α by arcsin α.

Comparison of the energies obtained by the variational method with those given by exact solution at $\gamma = 0$ reveal an error of the order of 7.5% at $\gamma = 0$ and less at other γ values; this is comparatively low. When $\gamma = 1$, (27.17) is an exact solution of equation (27.13). The ground state energy as a function of γ, calculated from (27.18) and (27.19), is shown in Figure 28.



a b

FIGURE 28. Ground state energy E_0/E_B of an impurity center vs. anisotropy parameter $\gamma = m_\perp/m_{\parallel}$: a) $\gamma \geqslant 1$ /21.10/; b) $\gamma \leqslant 1$ /21.2/. $E_B = m_\perp e^4/2\hbar^2\varkappa^2$.

A few words about classification of excited states for equations (27.15), (27.16). When $m_\parallel = m_\perp$, i.e., $\gamma = 1$, the l quantum number is degenerate; this is characteristic of a Coulomb center. But if $\gamma \neq 1$, this degeneracy is removed. The effective mass Hamiltonian (27.15), (27.16) possesses axial symmetry $C_{\infty v} \times C_i$, and therefore its eigenfunctions are characterized by a magnetic quantum number $m_z = m$ and parity P. When spin is neglected, $m = 0$ states are onefold degenerate, while $m \neq 0$ states are twofold degenerate in $\pm m$. Rotation about the symmetry axis through the angle φ multiplies the degenerate functions by $e^{\pm im\varphi}$, while a reflection σ_v takes them into one another. The eigenfunctions of the Hamiltonian (27.15) for given m and parity p are thus superpositions of states of a Coulomb center with different n but given m value and parity. Thus, the wave function with $m = 0$ and positive parity is a superposition of states $1s, 2s, 3s, 3d, 4s, 4d$, etc. with $m = 0$.

If the spacing of the excited levels in the Coulomb spectrum exceeds the perturbation due to the discrepancy of the masses $m_\parallel$ and $m_\perp$, the states of the operator (27.15) are built up in the first approximation from excited states of a Coulomb center with the same n. Thus, states with $n = 1, 2$ transform as follows: $1s \rightarrow 1s(0, +)$; $2p \rightarrow 2p_0(0, -) + 2p_\pm(\pm 1, -)$; $2s \rightarrow 2s(0, +)$ (the magnetic quantum number m and parity are indicated in parentheses). The excited states for a many-valley band, like the ground state, have additional degeneracies depending on the number of extrema. When the effective mass equations are corrected this degeneracy is removed, and the nature of the splitting may be determined by (27.11). However, owing to the large radius of the wave functions, the corrections to the effective mass equations for the excited states are small, in fact negligible for states with $f(0) = 0$.

Let us consider in detail the structure of donor impurities in Ge and Si, whose conduction band has a many-valley structure.

The conduction band of Si has six minima inside the Brillouin zone on the fourfold axes; $m_\parallel = 0.98\, m$, $m_\perp = 0.19\, m$ and $\varkappa = 11$. Figure 28 shows that the ground state energy is $E_0 = 0.029$ ev. Experimental values of the ionization energy for various donors are given in Table 27.1.

TABLE 27.1

Donor	Ionization energy, ev		Donor	Ionization energy, ev	
	Si	Ge		Si	Ge
P	0.044	0.013	Li	0.033	—
As	0.049	0.014	Bi	—	0.013
Sb	0.039	0.010			

It is evident from equation (27.1) that effective mass theory gives the ionization energy the correct order of magnitude, but for heavy atoms the discrepancy between theoretical and experimental values may reach 50%.

In effective mass theory the $1s$ ground state of the donor in Si is sixfold degenerate, according to the number of minima in the conduction band. The donor ions in Si and Ge are substitutional impurities, whose local symmetry group is T_d. Using (27.12), we easily check that because of chemical shifts the sixfold degenerate $1s$ state splits into a singlet state A_1, a twofold degenerate state E and a threefold degenerate state F_2:

$$1s = A_1 + E + F_2.$$

Indexing the functions (27.3) corresponding to the extrema

$$k_0(100),\quad k_0(\bar{1}00),\quad k_0(010),\quad k_0(0\bar{1}0),\quad k_0(001),\quad k_0(00\bar{1})$$

from 1 to 6 in order, we see that the correct wave functions (27.10), determined by the coefficients c_l, have the form

$$\begin{aligned} \Psi_{A_1} &= \frac{1}{\sqrt{6}}(111111), & \Psi^1_{F_2} &= \frac{1}{\sqrt{2}}(1\bar{1}0000), \\ \Psi^1_E &= \frac{1}{2}(11\bar{1}\bar{1}00), & \Psi^2_{F_2} &= \frac{1}{\sqrt{2}}(001\bar{1}00), \\ \Psi^2_E &= \frac{1}{2\sqrt{3}}(1111\bar{2}\bar{2}), & \Psi^3_{F_2} &= \frac{1}{\sqrt{2}}(00001\bar{1}). \end{aligned} \tag{27.20}$$

The corrections to the effective mass equations for the donor ground state in Si are described in general by three constants: Δ_0, which determines the downward shift of the singlet A_1, and Δ_2, Δ_1, which define the separation of the singlet A_1 from the doublet E and triplet F_2.

Experimental results, including examination of optical and paramagnetic spectra of donor impurities in strained crystals, show that for all donors in Si, except Li, the orbital singlet A_2 is low and the E and F_2 states have practically the same energy $\Delta_1 \approx \Delta_2$.

The energy of the excited donor states in Si has been calculated by Luttinger and Kohn /21.2/. They use a variational method, corrected by consideration of the exact values of the energy at $\gamma = 1$ and $\gamma = 0$, which can be determined analytically. The results for a few neighboring excited states are given in Table 27.2, along with the form of the trial function.

TABLE 27.2

State	Trial function	Si	
		E_{var}	E_{corr}
$1s$	$c \exp\left[-\left(\frac{\rho^2}{a^2} + \frac{z^2}{b^2}\right)^{1/2}\right]$	0.029	
$2p_0$	$cz \exp\left[-\left(\frac{\rho^2}{a^2} + \frac{z^2}{b^2}\right)^{1/2}\right]$	0.0107	0.0113
$2s$	$(c_1 + c_2\rho^2 + c_3 z^2) \exp\left[-\left(\frac{\rho^2}{a^2} + \frac{z^2}{b^2}\right)^{1/2}\right]$	0.0082	0.0088
$2p_\pm$	$cx \exp\left[-\left(\frac{\rho^2}{a^2} + \frac{z^2}{b^2}\right)^{1/2}\right]$	0.0057	0.0059
$3p_0$	$(c_1 + c_2\rho^2 + c_3 z^2) z \exp\left[-\left(\frac{\rho^2}{a^2} + \frac{z^2}{b^2}\right)^{1/2}\right]$	0.0052	0.0057

In the effective mass approximation, each $m = 0$ state is sixfold degenerate, while each $m \neq 0$ excited state is twelvefold degenerate. Using equation (27.11), we readily see that chemical shifts split the degenerate states as follows:

$$\begin{array}{lll} m = 0: & & A_1 + E + F_2; \\ m \neq 0, & \text{even:} & A_1 + A_2 + 2E + F_1 + F_2; \\ m & \text{odd:} & 2F_1 + 2F_2. \end{array} \tag{27.21}$$

However, this splitting may be quite small for the excited states, except possibly the $2s$ state, so that in practice it may be neglected. According to the estimates in /21.12/ for the $2s$ state in Si, the energy difference between the onefold state A_1 and the fivefold state $E + F_2$ is of the order of 0.14–0.12 mev.

In Ge, the conduction band minima lie on threefold axes, on the boundary of the Brillouin zone. The components of the effective mass tensor are $m_\parallel/m_0 = 1.4$, $m_\perp/m_0 = 0.083$ and the dielectric constant $\varkappa = 16$, giving $E_0 = 0.008$ ev for the ground state in the effective mass approximation. Experimental E values are given in Table 27.1. It is clear that effective mass theory yields a better agreement with experiment for Ge than for Si, owing to the fact that the donor centers in Ge are "shallower." In the effective mass approximation, the impurity ground state in Ge is fourfold degenerate.

Chemical shifts split the degenerate state into a singlet A_1 and a threefold degenerate state F_2; relative to the basis functions (27.3), with extrema on the [111], [11$\bar{1}$], [1$\bar{1}$1], [$\bar{1}$11] axes, the corresponding wave functions are

$$\begin{gathered}\Psi_{A_1}=\frac{1}{2}(1111);\\ \Psi^1_{F_2}=\frac{1}{2}(11\bar{1}\bar{1}),\quad \Psi^2_{F_2}=\frac{1}{2}(1\bar{1}1\bar{1}),\quad \Psi^3_{F_2}=\frac{1}{2}(1\bar{1}\bar{1}1).\end{gathered} \tag{27.22}$$

The above classification of terms arising from the $1s$ ground state for Ge and Si does not allow for spin-orbit coupling, and therefore each of the states obtained possesses an additional twofold degeneracy due to spin. Counting spin, these representations become:

$$A_1\rightarrow E_1'(\Gamma_6),\quad E\rightarrow G'(\Gamma_8),\quad F_2\rightarrow E_2'(\Gamma_7)+G'(\Gamma_8).$$

Therefore, as a result of spin-orbit coupling the sixfold degenerate (counting spin) state F_2 splits into twofold and fourfold degenerate states E_2' and G'.

Degenerate band. We consider the structure of a shallow impurity center in a degenerate band such as the valence band of Ge and Si, in which the wave functions transform according to the representation Γ_8^+ at an extremum point. Since the spin-orbit splitting of the valence band, which is 0.29 ev in Ge, greatly exceeds the acceptor ionization energy, this approximation is quite good for Ge. The wave functions f of a shallow acceptor center are defined by equations (27.1) with the matrix $\mathscr{H}$ of (24.13).

System (27.1) cannot be solved explicitly for Ge, and therefore the variational method is again used to find the wave functions and ground state energy E_0. The form of the trial functions is not arbitrary but must satisfy certain conditions dictated by symmetry.

For the ground state functions of a shallow acceptor center in the $\Gamma_8^\pm$ (or Γ_8) band, which also transform according to $\Gamma_8^\pm$ (Γ_8), it follows from (27.9) that the functions f transform according to the representations

$$\Gamma_8^\pm\times\Gamma_8^\pm=2\Gamma_{15}^+ +2\Gamma_{25}^+ +\Gamma_{12}^+ +\Gamma_2^+ +\Gamma_1^+.$$

Functions f transforming according to these representations may be chosen as suitable cubic harmonics, built up from harmonic polynomials of degree l multiplied by exponential factors e^{-r/r_l}. When this is done, each cubic harmonic transforming according to a given representation of the cubic group may be constructed from harmonic polynomials corresponding to different values of l. However, harmonics with larger l give a smaller contribution to the ground state energy, since the corresponding functions vanish rapidly at $z=0$, and so the computational work may be reduced by confining attention to harmonics with small l values. Thus, Schechter /21.4/ restricts himself to harmonics with $l\leqslant 2$, i.e., to s- and d-type functions, which transform according to the representations Γ_1^+ ($l=0$), Γ_{12}^+ ($l=2$), Γ_{25}^+ ($l=2$).

Using the selected cubic harmonics and the projection operators, we can construct trial functions f for the impurity center, which transform according to the representation Γ_8^-:

$$\begin{aligned} f_1 = c_1\varphi_1 \begin{Vmatrix} 1 \\ 0 \\ 0 \\ 0 \end{Vmatrix} + c_2\varphi_2 \begin{Vmatrix} \chi_1 \\ 0 \\ -\chi_2 \\ 0 \end{Vmatrix} + c_3\varphi_3 \begin{Vmatrix} 0 \\ (i/\sqrt{3})(e_2 + ie_1) \\ ie_3/\sqrt{3} \\ 0 \end{Vmatrix} + \\ + c_4\varphi_4 \begin{Vmatrix} 0 \\ (2/\sqrt{15})(e_1 - ie_2) \\ (1/\sqrt{15})\, ie_3 \\ -(1/\sqrt{5})(e_1 + ie_2) \end{Vmatrix}, \\ f_2 = c_1\varphi_1 \begin{Vmatrix} 0 \\ 1 \\ 0 \\ 0 \end{Vmatrix} + c_2\varphi_2 \begin{Vmatrix} 0 \\ -\chi_1 \\ 0 \\ -\chi_2 \end{Vmatrix} + c_3\varphi_3 \begin{Vmatrix} -(e_1 + ie_2)/\sqrt{3} \\ 0 \\ 0 \\ ie_3/\sqrt{3} \end{Vmatrix} + \\ + c_4\varphi_4 \begin{Vmatrix} 0 \\ 0 \\ -(1/\sqrt{5})(e_1 - ie_2) \\ -i\sqrt{3/5}\, e_3 \end{Vmatrix}, \end{aligned} \tag{27.23}$$

where

$$\begin{aligned} &\left.\begin{aligned} \chi_1 &= z^2 - \frac{1}{2}(x^2 + y^2) \\ \chi_2 &= \frac{\sqrt{3}}{2}(x^2 - y^2) \end{aligned}\right\} \Gamma_{12}^{+}, \quad \left.\begin{aligned} e_1 &= yz \\ e_2 &= xz \\ e_3 &= xy \end{aligned}\right\} \Gamma_{25}^{+}, \\ &\varphi_1 = \frac{1}{\sqrt{\pi}} e^{-r/r_1} r_1^{-3/2}, \quad \varphi_2 = \frac{e^{-r/r_2} r_2^{-7/2}}{3\sqrt{\pi}}, \\ &\varphi_3 = e^{-r/r_2} r_2^{-7/2} \sqrt{\frac{2}{3\pi}}, \quad \varphi_4 = e^{-r/r_3} r_3^{-7/2} \sqrt{\frac{2}{3\pi}}, \end{aligned} \tag{27.24}$$

and c_1, c_2, c_3, c_4 are numerical coefficients satisfying the normalization condition $c_1^2 + c_2^2 + c_3^2 + c_4^2 = 1$. The functions f_4 and f_3 are Kramers conjugates of f_1 and f_2, and we should keep in mind the transformations (24.18) of the functions of the bottom of the band under time reversal. The variational parameters in these trial functions are $r_1, r_2, r_3, c_1, c_2, c_3$, c_4.

Schechter /21.4/ uses "truncated" functions (27.23), setting $c_4 = 0$. Variational calculations for Ge yield $E_0 = 0.008$ ev for band parameters $A = -13.2\,\hbar^2/2m_0$, $B = -8.9\,\hbar^2/2m_0$, $C = 12.5\,\hbar^2/2m_0$. Table 27.3 presents the variational parameters.

TABLE 27.3

	E, ev	r_1, Å	r_2, Å	c_1	c_2	c_3
Ge	0.0088	44.0	34.4	0.86	−0.28	0.43
Si	0.031	17.71	13.21	0.92	−0.14	0.32

Experimental ionization energies for acceptors in Ge are given in Table 27.4.

TABLE 27.4

Acceptor	Ionization energy, ev		Acceptor	Ionization energy, ev	
	Ge	Si		Ge	Si
B	0.104	0.045	Ga	0.0108	0.065
Al	0.0102	0.057	In	0.112	0.16

It is clear that effective mass theory yields good agreement with the experimental ionization energy in Ge for all acceptors except In.

Similar calculations for Si yield $E_0 = 0.031$ using band parameters $A = -4.04\,\hbar^2/2m_0$, $B = -1.24\,\hbar^2/2m_0$, $C = 4.0\,\hbar^2/2m_0$. The variational parameters and experimental acceptor ionization energies are given in Tables 27.3–27.5.

TABLE 27.5

$B>0$						$B<0$					
E, ev	r_1, Å	r_2, Å	c_1	c_2	c_3	E, ev	r_1, Å	r_2, Å	c_1	c_2	c_3
0.054	9.12	7.09	0.83	−0.33	0.42	0.035	15.61	12.0	0.93	0.10	0.32

In Si the ionization energy of acceptor impurities is of the order of or greater than the valence band spin-orbit splitting $\Delta = 0.04$ ev; for Si, therefore, the approximation $\Delta \gg E_0$ used above is not valid.

Schechter /21.4/ also calculates the ionization energy of an acceptor center in the opposite limiting case of zero spin-orbit splitting, i. e., for band Γ_{15}. Trial wave functions which transform according to the representation Γ_{15} are chosen as follows:

$$
\begin{aligned}
f_1 &= c_1\varphi_1(r)\begin{Vmatrix} 1 \\ 0 \\ 0 \end{Vmatrix} + c_2\varphi_2(r)\begin{Vmatrix} \sqrt{3}\,\chi_2 - \chi_1 \\ 0 \\ 0 \end{Vmatrix} + c_3\varphi_3(r)\begin{Vmatrix} 0 \\ \varepsilon_3 \\ \varepsilon_2 \end{Vmatrix}, \\
f_2 &= c_1\varphi_1(r)\begin{Vmatrix} 0 \\ 1 \\ 0 \end{Vmatrix} + c_2\varphi_2(r)\begin{Vmatrix} 0 \\ -(\chi_1 + \sqrt{3}\,\chi_2) \\ 0 \end{Vmatrix} + c_3\varphi_3(r)\begin{Vmatrix} \varepsilon_3 \\ 0 \\ \varepsilon_1 \end{Vmatrix}, \qquad (27.25) \\
f_3 &= c_1\varphi_2(r)\begin{Vmatrix} 0 \\ 0 \\ 1 \end{Vmatrix} + c_2\varphi_2(r)\begin{Vmatrix} 0 \\ 0 \\ 2\chi_1 \end{Vmatrix} + c_3\varphi_3(r)\begin{Vmatrix} \varepsilon_2 \\ \varepsilon_1 \\ 0 \end{Vmatrix},
\end{aligned}
$$

where

$$
\varphi_1(r) = \frac{1}{\sqrt{\pi}}\,e^{-r/r_1} r_1^{-7/2}, \quad \varphi_2(r) = \frac{e^{-r/r_2}\, r_2^{-7/2}}{3\sqrt{2\pi}}, \quad \varphi_3(r) = \frac{e^{-r/r_2}\, r_2^{-7/2}}{\sqrt{3\pi}}, \qquad (27.26)
$$

and the coefficients c_1, c_2, c_3 satisfy the normalization condition $c_1^2 + c_2^2 + c_3^2 = 1$. Table 27.5 presents the variational calculations for Si for the case $\Delta = 0$, with the same band parameters as before.

The calculations show that in Si, where the impurity centers are deeper, effective mass theory, though it gives the correct order of magnitude for the acceptor ionization energy, yields results which deviate from the experimental data.

The excited acceptor states in effective mass theory are also characterized by irreducible representations of the group O_h, i. e., they transform according to the representations $\Gamma_6^{\pm}$, $\Gamma_7^{\pm}$, $\Gamma_8^{\pm}$. The states $\Gamma_6^{\pm}$, $\Gamma_7^{\pm}$ are Kramers doublets, and $\Gamma_8^{\pm}$ is fourfold degenerate.

When $B=0$ and $D=0$, the energy operator (24.12) in the effective mass approximation would have spherical symmetry, and envelope functions of an acceptor center in the Γ_8 band would correspond to states of a Coulomb center with mass $m^* = \hbar^2/2A$, each state possessing additional fourfold degeneracy due to valence band degeneracy.

If $B = D/\sqrt{3}$, i. e., $\gamma_2 = \gamma_3$, the Hamiltonian (27.1) with $\mathscr{H}(\boldsymbol{k})$ defined by (24.2) or (26.12) retains its spherical symmetry and commutes with the operator $\boldsymbol{F} = \boldsymbol{L} + \boldsymbol{J}$, where $\boldsymbol{L}$ is the orbital angular momentum operator (4.4). For given $\boldsymbol{F}$, the state Ψ_m^F is m-fold degenerate, $m = -F, -F+1, \ldots, F$, and the angular part represents the set of spherical harmonics $Y_{m'}^{l}$, with $l = F \pm {}^3/_2$ and $F \pm {}^1/_2$ (when $F = {}^1/_2$, $l = 1, 2$) /21.14/.

If $B \neq D/\sqrt{3}$, when the symmetry is reduced to cubic, this degeneracy is partially removed. The nature of this splitting is easily determined by group-theoretic considerations.

If B and D are not small, each state $\Gamma_i^{\pm}$ $(i = 6, 7, 8)$ is a mixture of hydrogenic states with different n and l for which the product $\mathscr{D}_l^{\pm} \times \Gamma_8^{+}$ contains $\Gamma_i^{\pm}$. Thus, in the general case the slow ground state function can be considered as a superposition of hydrogenic $1s$, $3d$, $5g$, etc. states, and the neighboring excited states Γ_8^-, Γ_7^-, Γ_6^- are a mixture of $2p$, $4f$, etc. states.

Schechter's variational calculations /21.4/ (see equations (27.23) and Table 27.4) show that the contribution of d-like functions to the ground state wave function is considerable.

Schechter carried out a variational calculation for the neighboring excited states Γ_8^- and Γ_6^-, Γ_7^-, using p- and f-type envelope functions. A shortcoming of Schechter's trial functions for the excited states is that their radial parts have no nodes besides $r=0$, while hydrogenic p-, d- and f-type functions have such nodes if $n_r > 1$.

Mendelson and James /21.5/ express the angular part of the trial function as a linear combination of spherical harmonics, guaranteeing the correct symmetry properties and the prescribed parity of the function; they determine the radial functions and the coefficients by solving a system of differential equations. In their approach, the impurity states are classified not only according to representation index and parity, but also according to the number of nodes of the radial wave function.

The ground state $\Gamma_8^+(0)$ of Γ_8^+, which is nodeless, is a superposition of $1s$ and $3d$ states of a hydrogenic center. The ionization energy is 0.0093 ev.* One more state $\Gamma_8^+(1)$, whose radial function has one node (not considered by Schechter) is primarily a superposition of $2s$ and $3d$ hydrogenic states; the state $\Gamma_6^-(0)$ is a superposition of $2p$ and $4f$ hydrogenic states. Of the Γ_7^-

---

* Mendelson and James base their calculations on the band parameter data of Levinger and Frankl /27.21/. They remark that if the data from Dresselhaus, Kip and Kittel /19.1/ are used, the discrepancy in energies is of the order of 1%, that in the wave functions of the order of 4%.

states, Mendelson and James considered $\Gamma_7^-(0)$ which, like $\Gamma_6^-(0)$, is a superposition of $2p$ and $4f$ states, and $\Gamma_7^-(1)$, whose radial function has one node, which is a superposition of $3p$ and $5f$ states.

Mendelson and James calculated the energy of several Γ_8^- states. Two of these, $\Gamma_8^-(0, 1)$ and $\Gamma_8^-(0, 2)$, have a nodeless radial function but angular functions of different types. They may be interpreted as two different superpositions of $2p$ and $4f$ hydrogenic states. The state $\Gamma_8^-(1)$, with one node in the radial function and an angular part very similar to the function $\Gamma_8^-(0, 1)$, is a superposition of $3p$ and $5f$ states.

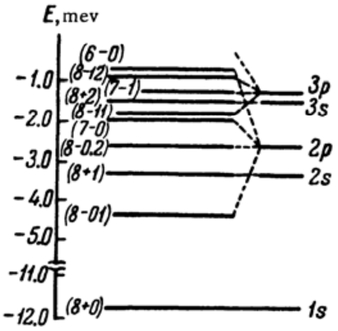


FIGURE 29. Excited acceptor states in Ge.

Figure 29 illustrates the pattern of excited states for an In shallow acceptor in Ge, based on the experimental data of /41.6/. These data are in good agreement with the theoretical calculations of /21.5/.

Let us consider the asymptotic behavior of the shallow impurity wave functions far away from the impurity nucleus. We shall approximate the wave function asymptotically by $f \sim e^{-kr}$. Equation (27.15) for the ellipsoidal constant energy surfaces then yields the relationship between k and the bound state energy $E < 0$ as a function of the direction of $\boldsymbol{r}$:

$$k = \sqrt{-\frac{2E}{\hbar^2}\left(m_\perp (n_x^2 + n_y^2) + m_\parallel n_z^2\right)}, \quad \text{where} \quad n_i = r_i/r. \tag{27.27}$$

Thus, the wave function is asymptotically anisotropic. According to (27.27), the wave function extends farthest in the direction of least effective mass. Note that the asymptotic form of the wave function has the same form as the asymptotic trial function, except that $a_\parallel^{-2}$ and $a_\perp^{-2}$ are replaced by $-2m_\parallel E/\hbar^2$ and $-2m_\perp E/\hbar^2$, respectively.

The case of a degenerate band is somewhat more complicated. Here we consider the asymptotic behavior of the acceptor states in Ge and Si, i.e., in the case of the Γ_8 band. Approximating the solution of the effective mass equation (27.1) asymptotically by

$$f \sim \left\| \begin{matrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{matrix} \right\| e^{-kr}, \tag{27.28}$$

we obtain a homogeneous system for the coefficients c_s:

$$\sum_s \mathcal{H}_{is}(k^2 n_\alpha n_\beta)\, c_s = E c_i, \tag{27.29}$$

where $\mathcal{H}_{is}(k^2 n_\alpha n_\beta)$ is the matrix (24.13) with $k_\alpha k_\beta$ replaced by $k^2 n_\alpha n_\beta$. The solvability condition for the system,

$$|\mathcal{H} - EI| = 0 \ , \tag{27.29a}$$

gives a dependence of $\boldsymbol{k}^2$ on E similar to (24.14a):

$$\boldsymbol{k}^2\left(|A| \pm \sqrt{B^2(n_x^4+n_y^4+n_z^4)+C^2(n_x^2n_y^2+n_x^2n_z^2+n_y^2n_z^2)}\right) = -E. \tag{27.30}$$

Equation (27.29a) has two roots (27.30), corresponding to the two branches of the hole spectrum. This means that the asymptotic behavior of the wave function is determined by two terms of type (27.28), with different solutions $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ of equation (27.30) and different solutions c_i and c_i' of equation (27.29) with $\boldsymbol{k} = \boldsymbol{k}_1$ and $\boldsymbol{k} = \boldsymbol{k}_2$, respectively. Since the Γ_8 levels are fourfold degenerate, we have four functions of type (27.28), which are pairs of Kramers conjugates. As follows from (27.30), the asymptotic behavior of the wave function depends essentially on the direction and reproduces the form of the constant energy surface of the holes in $\boldsymbol{k}$-space. The coefficients c_i and c_i' also depend on n_α: they are determined by equations (24.19) with k_α replaced by kn_α. At large distances, only one exponent, corresponding to the light hole band, "survives."

Thus, the asymptotic behavior of the wave functions of a shallow impurity center in a degenerate band is determined by the light carrier spectrum, despite the fact that this branch of the spectrum may give only an insignificant contribution to the energy. Since the ratio of the masses of heavy and light holes may be quite high, allowance for this situation is important for various effects which depend on the behavior of the impurity center wave functions at large distances.

Corrections to effective mass theory. We now consider corrections to the Hamiltonian (27.2) due to higher order terms in $\boldsymbol{k}$ and also to inclusion of matrix elements of the potential $U_{\boldsymbol{k}}$ with large $\boldsymbol{k} \approx \boldsymbol{b}_M$, which were omitted in (22.8). The former terms yield corrections of the order of E_0^2/E_g, which may be significant for comparatively deep centers; the latter terms are responsible for chemical shifts.

To calculate the energy corrections of the order of E_0^2/E_g in the effective mass equations (22.15), we must retain the fourth order terms in $\boldsymbol{k}$, which are easily obtained by the usual kp-theory or the theory of invariants,* and also incorporate corrections ΔU to the Coulomb potential, which appear in third order perturbation theory when quadratic terms in $\mathscr{H}_2 = \hbar(\boldsymbol{k}\boldsymbol{\pi})/m$ and linear terms in $\mathscr{H}_1 = -U(\boldsymbol{r})$ are included. Since by (22.10b) and (22.10c) the operator $\mathscr{H}_2$ is diagonal in $\boldsymbol{k}$ and $\mathscr{H}_1$ is diagonal in n, we obtain the following expression for ΔU from the general equation (15.47):

$$\begin{aligned}\Delta U_{m'\boldsymbol{k}',\, m\boldsymbol{k}} &= -\frac{1}{2}\frac{\hbar^2}{m^2}\sum_{s\alpha\beta}\frac{\pi^\alpha_{m's}\pi^\beta_{sm}}{(E_s-E_m)^2}\,U_{\boldsymbol{k}'-\boldsymbol{k}}\left(k'_\alpha k'_\beta - 2k'_\alpha k_\beta + k_\alpha k_\beta\right) = \\ &= -\left[\sum_{\alpha\beta}(k'-k)_\alpha(k'-k)_\beta\, G^{\alpha\beta}_{m'm} - \left([\boldsymbol{k}'\boldsymbol{k}]\,\boldsymbol{g}_{m'm}\right)\right]U_{\boldsymbol{k}'-\boldsymbol{k}},\end{aligned} \tag{27.31}$$

where

$$G^{\alpha\beta}_{m'm} = \frac{1}{2}\frac{\hbar^2}{m^2}\sum_s\frac{\left[\pi^\alpha_{m's}\pi^\beta_{sm}\right]}{(E_s-E_m)^2} = \frac{1}{4}\frac{\hbar^2}{m^2}\sum_s\frac{\pi^\alpha_{m's}\pi^\beta_{sm}+\pi^\beta_{m's}\pi^\alpha_{sm}}{(E_s-E_m)^2}, \tag{27.32}$$

$$g^\gamma_{m'm} = \frac{1}{2}\frac{\hbar^2}{m^2}\sum_{s\alpha\beta}\delta_{\alpha\beta\gamma}\frac{\pi^\alpha_{m's}\pi^\beta_{sm}}{(E_s-E_m)^2} = \frac{1}{2}\frac{\hbar^2}{m^2}\sum_{\gamma s}\frac{\pi^{\gamma+1}_{m's}\pi^{\gamma+2}_{sm}-\pi^{\gamma+2}_{m's}\pi^{\gamma+1}_{sm}}{(E_s-E_m)^2}, \tag{27.33}$$

---

* If the group $G_{\boldsymbol{k}_0}$ does not contain inversion, the operator $\mathscr{H}(\boldsymbol{k})$ will also include third order $\boldsymbol{k}$ terms.

$\delta_{\alpha\beta\gamma}$ is the antisymmetric unit tensor. To find ΔU in the coordinate representation, we use (22.14a), multiplying ΔU by $e^{i(\boldsymbol{k}'-\boldsymbol{k})\boldsymbol{r}}$ and summing over $\boldsymbol{q}=\boldsymbol{k}'-\boldsymbol{k}$ for a fixed vector $\boldsymbol{k}$, which is then replaced by the operator $\boldsymbol{k}=-i\nabla$ The first term in (27.31) gives

$$\Delta U^{1}_{m'm}=-\sum_{\alpha\beta} G^{\alpha\beta}_{m'm}\sum_{\boldsymbol{q}} q_\alpha q_\beta U_{\boldsymbol{q}} e^{i\boldsymbol{q}\boldsymbol{r}}=\sum_{\alpha\beta} G^{\alpha\beta}_{m'm}\left(\frac{\partial}{\partial r_\alpha}\frac{\partial}{\partial r_\beta}U(\boldsymbol{r})\right).$$

In calculating this term we must remember that

$$\nabla^2 U(\boldsymbol{r})=\frac{4\pi Z e^2}{\varkappa}\delta(r)$$

is singular at $\boldsymbol{r}=0$, and so the corresponding term must be separated. The result is

$$\Delta U^{1}_{m'm}=\left[\sum_{\alpha\beta}\left(G^{\alpha\beta}_{m'm}-\frac{1}{3}\operatorname{Tr} G_{m'm}\delta_{\alpha\beta}\right)\frac{\partial^2}{\partial r_\alpha \partial r_\beta}+\frac{1}{3}\operatorname{Tr} G_{m'm}\nabla^2\right]U(\boldsymbol{r}).$$

Evaluating $\nabla^2 U$ and performing the differentiation in the first term, we finally obtain

$$\Delta U^{1}_{m'm}=-\frac{Ze^2}{\varkappa}\left[\sum_{\alpha\beta} G^{\alpha\beta}_{m'm}\left(\frac{3r_\alpha r_\beta}{r^5}-\frac{1}{r^3}\delta_{\alpha\beta}\right)-\frac{4\pi}{3}\operatorname{Tr} G_{m'm}\delta(\boldsymbol{r})\right], \tag{27.34}$$

where

$$\operatorname{Tr} G_{m'm}=\sum_{\alpha} G^{\alpha\alpha}_{m'm}.$$

In calculating the second term we use the fact that $[\boldsymbol{k}'\boldsymbol{k}]=[\boldsymbol{q}\boldsymbol{k}]$, and so

$$\Delta U^{2}_{m'm}=\sum_{\boldsymbol{q}}([\boldsymbol{q}\boldsymbol{k}]\,\boldsymbol{g}_{m'm})\,U_{\boldsymbol{q}} e^{i\boldsymbol{q}\boldsymbol{r}}=-([\nabla U\cdot\nabla]\,\boldsymbol{g}_{m'm})=-\frac{Ze^2}{\varkappa r^3}([\boldsymbol{r}\nabla]\,\boldsymbol{g}_{m'm}).$$

Consequently,

$$\Delta U^{2}_{m'm}=-i\frac{Ze^2}{\varkappa r^3}(\boldsymbol{M}\boldsymbol{g}_{m'm}), \tag{27.35}$$

where $\boldsymbol{M}=-i[\boldsymbol{r}\nabla]$ is the angular momentum operator.

The first term ΔU^1 represents the effective dipole-dipole and contact interaction. The matrix $\boldsymbol{G}^{\alpha\beta}$ defining it is similar to the matrix $\mathscr{H}^{\alpha\beta}$ of (21.19) defining the spectrum and has the same number of linearly independent elements. It can be obtained by formal differentiation of the latter:

$$G^{\alpha\beta}_{m'm}=-\frac{1}{2}\frac{\partial \mathscr{H}^{\alpha\beta}_{m'm}}{\partial E_m}. \tag{27.36}$$

The second term ΔU^2 describes the interaction of the orbital momenta of the slow and fast functions, causing splitting of terms in cases when the representation $\Gamma_f\times\mathscr{D}_{1/2}$ is reducible counting spin. The matrix $\boldsymbol{g}^{\gamma}$ defining it is similar to the matrix $\boldsymbol{L}^{\gamma}$ of (22.22) defining the effective magnetic moment:

$$g^{\gamma}_{m'm}=-\frac{i\hbar^2}{m}\frac{\partial L^{\gamma}_{m'm}}{\partial E_m}. \tag{27.37}$$

In the simplest case, the two-band spherical model, the sums (27.32) and (27.33) contain one interband term with $s=n$, and the indices m and n take two values $m=1, 2$ and $n=1, 2$, corresponding to spin ± 1/2. In this model, by (26.34),

$$E_m - E_n = E_g = 2m^* s^2, \quad \pi_{mn} = m^* s \sigma_{mn}. \tag{27.38}$$

Then

$$G^{\alpha\beta}_{m'm} = \frac{1}{2}\frac{\hbar^2 s^2}{E_g^2}\sum_n \left(\sigma^\alpha_{m'n}\sigma^\beta_{nm}\right) = \frac{1}{2}\frac{\hbar^2 s^2}{E_g^2}\delta_{m'm}\delta_{\alpha\beta}, \tag{27.39a}$$

$$g^\gamma_{m'm} = \frac{1}{2}\frac{\hbar^2 s^2}{E_g^2}\sum_n \delta_{\alpha\beta\gamma}\sigma^\alpha_{m'n}\sigma^\beta_{nm} = i\frac{\hbar^2 s^2}{E_g^2}\sigma^\gamma_{m'm} \tag{27.39b}$$

and so

$$\Delta U^1 = \frac{2\pi Z e^2 \hbar^2 s^2}{E_g^2}\delta(\boldsymbol{r}), \tag{27.40}$$

$$\Delta U^2 = \frac{Z e^2 \hbar^2 s^2}{E_g^2}\frac{1}{r^3}(\sigma \boldsymbol{M}). \tag{27.41}$$

In this model, the fourth order $\boldsymbol{k}$ terms in $\mathscr{H}$, which make contributions to the energy of the same order as ΔU, have the form

$$\Delta\mathscr{H} = -\frac{1}{E_g}\left(\frac{\hbar^2\nabla^2}{2m^*}\right)^2. \tag{27.42}$$

The corrections to the energy of the hydrogenic ground state $f_0 = \frac{1}{\sqrt{\pi a^3}}e^{-r/a}$ due to ΔU^1 and $\Delta\mathscr{H}$ are respectively $4E_0^2/E_g$ and $-5E_0^2/E_g$, so that the total correction is $-E_0/E_g$.*

For the corrections due to the short-range part of the potential $U(\boldsymbol{r})$, it is sufficient to retain in (22.8) only the intraband terms

$$U_{m'\boldsymbol{k}', m\boldsymbol{k}} = \langle\varphi_{m'\boldsymbol{k}'}|U(\boldsymbol{r})|\varphi_{m\boldsymbol{k}}\rangle = \mathscr{V}\sum_M U_{\boldsymbol{k}'-\boldsymbol{k}+\boldsymbol{b}_M} B^{m'm}_M, \tag{27.43}$$

since the interband terms yield corrections which involve an additional small parameter E_0/E_g. From the matrix element (27.43) we have kept in (27.2) only the term with $M=0$. Consequently, the correction is

$$\delta U_{m'\boldsymbol{k}', m\boldsymbol{k}} = \mathscr{V}\sum_{M\neq 0} U_{\boldsymbol{k}'-\boldsymbol{k}+\boldsymbol{b}_M} B^{m'm}_M \approx \mathscr{V}\sum_{M\neq 0} U_{\boldsymbol{b}_M} B^{m'm}_M = \mathscr{V}\sum_M U_{\boldsymbol{b}_M} B^{m'm}_M - U_0\delta_{m'm}. \tag{27.44}$$

Here we have neglected the small quantity $\boldsymbol{k}'-\boldsymbol{k}$ in comparison with the reciprocal lattice vectors $\boldsymbol{b}_M$ and used the equation

$$\mathscr{V} B^{m'm}_0 U_0 = U_0\delta_{m'm} = \frac{1}{\mathscr{V}}\delta_{m'n}\int U(\boldsymbol{r})\,d\boldsymbol{r}.$$

A glance at (27.43) shows that the first term in (27.44) is $U_{m'\boldsymbol{k}_0, n'\boldsymbol{k}_0} = U_{m'n'}$, and so

---

* For this model, the two-band equation (26.34) coincides with the Dirac equation (up to constants). Therefore, the problem of an impurity center may be solved accurately in the two-band approximation in the same way as that of the hydrogen atom spectrum in the Dirac model /21.9/. The exact expression for the ground state energy is

$$E_0 = m^* s^2\left\{\left[1-\left(\frac{Ze^2}{\varkappa\hbar s}\right)^2\right]^{1/2} - 1\right\}.$$

$$\delta U_{m'\boldsymbol{k},\, m\boldsymbol{k}} = U_{m'm} - U_0\delta_{m'm} = \frac{1}{\mathscr{V}}\int U(\boldsymbol{r})\,\psi^*_{m'\boldsymbol{k}_0}\psi_{m\boldsymbol{k}_0}\,d\boldsymbol{x} - \delta_{m'm}\frac{1}{\mathscr{V}}\int U(\boldsymbol{r})\,d\boldsymbol{r} = \\ = \frac{1}{\mathscr{V}}\int U(\boldsymbol{r})\left[u^*_{m'\boldsymbol{k}_0}u_{m\boldsymbol{k}_0} - \delta_{m'm}\right]d\boldsymbol{x}. \qquad (27.45)$$

As in (22.14), in order to transform to the $\boldsymbol{r}$-representation we must multiply (27.45) by $e^{i\boldsymbol{q}\boldsymbol{r}}$ and sum over $\boldsymbol{q}$. Noting that $\delta U_{m'\boldsymbol{k}',\, m\boldsymbol{k}}$ does not depend on $\boldsymbol{k}$ and $\boldsymbol{k}'$ (i.e., on $\boldsymbol{q}$), while $\sum_{\boldsymbol{q}} e^{i\boldsymbol{q}\boldsymbol{r}} = \mathscr{V}\delta(\boldsymbol{r})$ since $\int e^{i\boldsymbol{q}\boldsymbol{r}}\,d\boldsymbol{r} = \mathscr{V}\delta_{\boldsymbol{q}0}$, we find that

$$\delta U_{m'm}(\boldsymbol{r}) = \mathscr{V}(U_{m'm} - U_0\delta_{m'm})\,\delta(\boldsymbol{r}), \qquad (27.46)$$

and consequently the matrix element between functions f^s_m (27.1) is

$$\delta U_{s's} = \mathscr{V}\sum_{mm'}(U_{m'm} - U_0\delta_{m'm})\,f^{*s'}_{m'}(0)\,f^s_m(0). \qquad (27.47)$$

Obviously, the main contribution to the matrix element (27.45) comes from integration over the cell near the impurity center. The corrections (27.47) yield a change of energy and may also cause additional splitting of a degenerate level at a given extremum $\boldsymbol{k}_0$. The operator δU has symmetry group $\overline{G} = G_i \cap G_{\boldsymbol{k}_0}$, which is in the general case lower than $G_{\boldsymbol{k}_0}$. In particular, it does not contain nonprimitive translations, since G_i does not contain them. If the representation $\mathscr{D}^{\boldsymbol{k}_0}_\mu$ of $G_{\boldsymbol{k}_0}$ according to which the wave functions $\psi_{m\boldsymbol{k}_0}$ transform is reducible in $\overline{G}$, the matrix $\delta\boldsymbol{U}$ (27.47) is not a multiple of the identity matrix, and the perturbation (27.46) causes splitting of terms.

The corrections to the effective mass equations for acceptor centers in Ge and Si do not split the fourfold degenerate ground state, since for symmetries $G_0 = O_h$ and $\overline{G} = T_d$ the levels have the same degeneracies.

When corrections are included in effective mass theory, the impurity center wave functions no longer have a definite parity, since T_d does not contain inversion; the corrections may therefore lead to observation of certain effects in the acceptor centers which are not detected in the effective mass approximation because O_h contains inversion /40.4/.

If the band has several equivalent extrema, at points $\boldsymbol{k}_{0i}$, then besides the intraband terms (27.43) we must introduce matrix elements between functions belonging to different extrema:

$$\delta U_{m'\boldsymbol{k}_{0i},\, m\boldsymbol{k}_{0j}} = \mathscr{V}\sum_M B^{m'_i m_j}_M U_{\boldsymbol{k}_{0i}-\boldsymbol{k}_{0j}+\boldsymbol{b}_M} = \frac{1}{\mathscr{V}}\int \psi^*_{m'\boldsymbol{k}_{0i}}U(\boldsymbol{r})\,\psi_{m\boldsymbol{k}_{0j}}\,d\boldsymbol{x} = \\ = \frac{1}{\mathscr{V}}\int u^*_{m'\boldsymbol{k}_{0i}}(\boldsymbol{x})\,U(\boldsymbol{r})\,u_{m\boldsymbol{k}_{0j}}(\boldsymbol{x})\,e^{i(\boldsymbol{k}_{0j}-\boldsymbol{k}_{0i})\boldsymbol{r}}\,d\boldsymbol{x}. \qquad (27.48)$$

Here, as in (27.44), we are neglecting the small deviation of the vectors $\boldsymbol{k}'_i - \boldsymbol{k}_j$, measured from extrema $\boldsymbol{k}_{0i}$ and $\boldsymbol{k}_{0j}$, from $\boldsymbol{k}_{0i} - \boldsymbol{k}_{0j} + \boldsymbol{b}_M$.

In the $\boldsymbol{r}$-representation, the operator is similar in form to (27.46):

$$\delta U_{m'_i m_j}(\boldsymbol{r}) = \mathscr{V}U_{m'\boldsymbol{k}_{0i},\, m\boldsymbol{k}_{0j}}\,\delta(\boldsymbol{r}). \qquad (27.49)$$

This operator determines chemical shift in many-valley semiconductors and lifts many-valley degeneracy. The number of linearly independent elements in this matrix and in the matrix (27.46) is easily determined, using the general equation of §19 and the fact that the potential $U(\boldsymbol{r})$ transforms according to the identity representation of the symmetry group G_i of the impurity center.

It follows from equation (27.49) that, in order to describe the splitting of many-valley degeneracy of an impurity center, we must use the general equations of §19 to expand the representation of the space group with star $\{\boldsymbol{k}_0\}$, according to which the wave functions $\psi_{m\boldsymbol{k}_{0l}}$ transform, in terms of irreducible representations of the group G_i (see (27.11) and (27.12)).

Excitons

Shallow excitons, i.e., excitons whose band energy is small compared to the band gap, may be regarded as bound states of an electron-hole pair.[*] The equation defining the wave functions and spectrum of shallow excitons is similar to the equation for an impurity center. To derive this equation, we first construct the effective mass Hamiltonian for the interaction of two electrons, one in a conduction band m near an extremum $\boldsymbol{k}_c$, the other in a valence band n with extremum $\boldsymbol{k}_v$; each of the states may be degenerate and the energy difference $E_m(\boldsymbol{k}_c) - E_n(\boldsymbol{k}_v) = E_m^0 - E_n^0$ is significantly greater than the binding energy of the exciton.

We start from the Schrödinger equation for a system of two electrons in a medium with dielectric constant $\varkappa$:

$$[\mathcal{H}(\boldsymbol{x}_1, \boldsymbol{x}_2) - E]\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = [(\mathcal{H}_0(\boldsymbol{x}_1) + \mathcal{H}_0(\boldsymbol{x}_2) + U(\boldsymbol{r}_1 - \boldsymbol{r}_2) - E]\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = 0, \quad (27.50)$$

where $\mathcal{H}_0(\boldsymbol{x})$ is defined by (22.1a) and $U(\boldsymbol{r}) = e^2/\varkappa r$.

We first consider a direct exciton, when $\boldsymbol{k}_v = \boldsymbol{k}_c = \boldsymbol{k}_0$. We write the wave function $\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ for the two electrons in a form similar to (22.2), (22.5):

$$\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{st} F_{st}(\boldsymbol{r}_1, \boldsymbol{r}_2)\psi_{s\boldsymbol{k}_0}(\boldsymbol{x}_1)\psi_{t\boldsymbol{k}_0}(\boldsymbol{x}_2) = \sum_{st\boldsymbol{k}_1\boldsymbol{k}_2} C_{s\boldsymbol{k}_1,\, t\boldsymbol{k}_2}\varphi_{s\boldsymbol{k}_1}(\boldsymbol{x}_1)\varphi_{t\boldsymbol{k}_2}(\boldsymbol{x}_2), \quad (27.51)$$

where

$$F_{st}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{1}{\mathcal{V}}\sum_{\boldsymbol{k}_1\boldsymbol{k}_2} C_{s\boldsymbol{k}_1,\, t\boldsymbol{k}_2} e^{i(\boldsymbol{k}_1\boldsymbol{r}_1 + \boldsymbol{k}_2\boldsymbol{r}_2)}, \quad (27.51\text{a})$$

$$C_{s\boldsymbol{k}_1,\, t\boldsymbol{k}_2} = \frac{1}{\mathcal{V}}\int F_{st}(\boldsymbol{r}_1, \boldsymbol{r}_2)\, e^{-i(\boldsymbol{k}_1\boldsymbol{r}_1 + \boldsymbol{k}_2\boldsymbol{r}_2)}\, d\boldsymbol{r}_1\, d\boldsymbol{r}_2. \quad (27.51\text{b})$$

The wave function $\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ must be antisymmetric, since the electrons obey the Pauli principle: $\Psi(\boldsymbol{x}_2, \boldsymbol{x}_1) = -\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$. Hence it follows that

$$C_{s\boldsymbol{k}_1,\, t\boldsymbol{k}_2} = -C_{t\boldsymbol{k}_2,\, s\boldsymbol{k}_1} \quad \text{or} \quad F_{st}(\boldsymbol{r}_1, \boldsymbol{r}_2) = -F_{ts}(\boldsymbol{r}_2, \boldsymbol{r}_1). \quad (27.52)$$

Substituting (27.51) into equation (27.50), multiplying on the left by $\varphi^*_{s'\boldsymbol{k}_1'}(\boldsymbol{x}_1)\varphi^*_{t'\boldsymbol{k}_2'}(\boldsymbol{x}_2)$, integrating over $\boldsymbol{r}_1, \boldsymbol{r}_2$ and summing over the spin variable α, we obtain a system of equations similar to (22.11):

$$\sum_{st\boldsymbol{k}_1\boldsymbol{k}_2}\Big[(E^0_{st}(\boldsymbol{k}_1, \boldsymbol{k}_2) - E)\delta_{ss'}\delta_{tt'}\delta_{\boldsymbol{k}_1\boldsymbol{k}_1'}\delta_{\boldsymbol{k}_2\boldsymbol{k}_2'} + \mathcal{H}_{1\,\substack{\boldsymbol{k}_1'\boldsymbol{k}_2'\\ \boldsymbol{k}_1\boldsymbol{k}_2}}\delta_{ss'}\delta_{tt'} + \mathcal{H}_{2\,\substack{s't'\\ st}}\delta_{\boldsymbol{k}_1\boldsymbol{k}_1'}\delta_{\boldsymbol{k}_2\boldsymbol{k}_2'}\Big]C_{s\boldsymbol{k}_1,\, t\boldsymbol{k}_2} = 0. \quad (27.53)$$

---

[*] This model of the exciton is due to Wannier and Mott; shallow excitons are therefore known as Wannier-Mott excitons.

Here, as opposed to (22.10), we have included the term $\hbar^2k^2/2m$ in E^0:

$$E^0_{st}(\mathbf{k}_1,\ \mathbf{k}_2) = E^0_{st} + \frac{\hbar^2k_1^2}{2m} + \frac{\hbar^2k_2^2}{2m}, \tag{27.54}$$

where

$$E^0_{st} = E^0_m + E^0_n = E_{m\mathbf{k}_0} + E_{n\mathbf{k}_0}.$$

As in (22.8), (22.10b), to calculate the matrix element of the Coulomb potential to a first approximation, we retain only terms with $M=0$ in the expansion

$$\sum_{\alpha} \psi^*_{s\mathbf{k}_0}\psi_{s'\mathbf{k}_0} = \mathscr{V}\sum_{M} B^{ss'}_M e^{-i\mathbf{b}_M\mathbf{r}},$$

where the sum over α stands for summation over the spin variables:

$$\mathscr{H}_{1\substack{s'\mathbf{k}'_1,\ t'\mathbf{k}'_2\\ s\mathbf{k}_1,\ t\mathbf{k}_2.}} = \mathscr{V}^2\sum_{ML} B^{s's}_M B^{t't}_L U_{\mathbf{k}'_1-\mathbf{k}_1+\mathbf{b}_M}\delta_{\mathbf{k}'_1-\mathbf{k}_1+\mathbf{b}_M,\ \mathbf{k}_2-\mathbf{k}'_2-\mathbf{b}_L} \approx \delta_{s's}\delta_{t't}U_{\mathbf{k}'_1-\mathbf{k}_1}\delta_{\mathbf{k}'_1+\mathbf{k}'_2,\ \mathbf{k}_2+\mathbf{k}_1}. \tag{27.55}$$

Note that here only states with small $\mathbf{k}_1$ and $\mathbf{k}_2$ need be considered, since we are concerned with excitons with a large radius near an extremum of the exciton band. The operator $\mathscr{H}_2$ in (27.53) is similar in form to (22.6), (22.10c):

$$\mathscr{H}_{2\substack{s'\mathbf{k}'_1,\ t'\mathbf{k}'_2\\ s\mathbf{k}_1,\ t\mathbf{k}_2}} = \frac{\hbar}{m}\left[(\mathbf{k}_1\boldsymbol{\pi}_{s's})\delta_{t't} + (\mathbf{k}_2\boldsymbol{\pi}_{t't})\delta_{s's}\right]\delta_{\mathbf{k}'_1\mathbf{k}_1}\delta_{\mathbf{k}'_2\mathbf{k}_2}. \tag{27.56}$$

Next, as in §22, we must derive the effective mass equation containing only the functions $\tilde{C}_{m\mathbf{k}_1,\ n\mathbf{k}}$, describing a state in which one electron is in the m band and the other in the n band; the other interband matrix elements are dropped. The basis of this "truncated" equation are the $2N_mN_n$ functions $\psi_m(\mathbf{x}_1)\psi_n(\mathbf{x}_2)$ and $\psi_m(\mathbf{x}_2)\psi_n(\mathbf{x}_1)$. Here $\psi_n = \psi_{n\mathbf{k}_0}$, and N_m and N_n are the degeneracies of the m and n bands at $\mathbf{k}_0$. Dropping the interband matrix elements, we obtain a system of equations for the coefficients $\bar{\bar{C}}_{m\mathbf{k}_1,\ n\mathbf{k}_2}$, $\bar{\bar{C}}_{n\mathbf{k}_2,\ m\mathbf{k}_1}$, resembling (22.12):

$$\sum_{mn\mathbf{k}_1\mathbf{k}_2} \bar{\mathscr{H}}_{\substack{m'\mathbf{k}'_1,\ n'\mathbf{k}'_2\\ m\mathbf{k}_1,\ n\mathbf{k}_2}}\bar{\bar{C}}_{m\mathbf{k}_1,\ n\mathbf{k}_2} + \bar{\bar{\mathscr{H}}}_{\substack{m'\mathbf{k}'_1,\ n'\mathbf{k}'_2\\ n\mathbf{k}_2,\ m\mathbf{k}_1}}\bar{\bar{C}}_{n\mathbf{k}_2,\ m\mathbf{k}_1} = E\bar{\bar{C}}_{m'\mathbf{k}'_1,\ n'\mathbf{k}'_2},$$

which, in view of (27.52), we can rewrite

$$\sum_{mn\mathbf{k}_1\mathbf{k}_2}\left(\bar{\mathscr{H}}_{\substack{m'\mathbf{k}'_1,\ n'\mathbf{k}'_2\\ m\mathbf{k}_1,\ n\mathbf{k}_2}} - \bar{\mathscr{H}}_{\substack{m'\mathbf{k}'_1,\ n'\mathbf{k}'_2\\ n\mathbf{k}_2,\ m\mathbf{k}_1}}\right)\bar{\bar{C}}_{m\mathbf{k}_1,\ n\mathbf{k}_2} = E\bar{\bar{C}}_{m'\mathbf{k}'_1,\ n'\mathbf{k}'_2}. \tag{27.57}$$

Here the energy E is measured from $E^0_{mn} = E^0_m + E^0_n$.

A similar system of equations may be derived for the coefficients $\bar{\bar{C}}_{n\mathbf{k}_2,\ m\mathbf{k}_1}$, but these equations are not independent, since they follow from (27.57), (27.52) and the condition

$$\bar{\mathscr{H}}_{\substack{m'\mathbf{k}'_1,\ n'\mathbf{k}'_2\\ m\mathbf{k}_1,\ n\mathbf{k}_2}} = \bar{\mathscr{H}}_{\substack{n\mathbf{k}_2,\ m\mathbf{k}_1\\ n'\mathbf{k}'_2,\ m'\mathbf{k}'_1}},$$

which is a consequence of the symmetry of the Hamiltonian (27.50): $\mathscr{H}(\mathbf{x}_1,\ \mathbf{x}_2) = \mathscr{H}(\mathbf{x}_2,\ \mathbf{x}_1)$. It is therefore sufficient to consider the system of $N_m \times N_n$ equations (27.57).

Using the general equations (15.47), we can use (27.55), (27.56) to find explicit expressions for $\bar{\mathscr{H}}$ in the required order of perturbation theory, as in §22. If we retain only first and second order terms, then

$$\overline{\mathscr{H}}_{\substack{m'\boldsymbol{k}_1',\, n'\boldsymbol{k}_2' \\ m\boldsymbol{k}_1,\, n\boldsymbol{k}_2}} = \left[\mathscr{H}(\boldsymbol{k}_1)_{m'm}\,\delta_{n'n} + \mathscr{H}(\boldsymbol{k}_2)_{n'n}\,\delta_{m'm}\right]\delta_{\boldsymbol{k}_1'\boldsymbol{k}_1}\delta_{\boldsymbol{k}_2'\boldsymbol{k}_2} + \\ + U_{\boldsymbol{k}_1'-\boldsymbol{k}_1}\,\delta_{\boldsymbol{k}_1'+\boldsymbol{k}_2',\, \boldsymbol{k}_1+\boldsymbol{k}_2}\,\delta_{m'm}\delta_{n'n}, \tag{27.58}$$

where $\mathscr{H}(\boldsymbol{k})$ is defined by (21.19) and $\overline{\mathscr{H}}^{(2)}_{\substack{m'\boldsymbol{k}_1',\, n'\boldsymbol{k}_2' \\ n\boldsymbol{k}_2,\, m\boldsymbol{k}_1}} = 0$.

In the third approximation,

$$\Delta U_{\substack{m'\boldsymbol{k}_1',\, n'\boldsymbol{k}_2' \\ m\boldsymbol{k}_1,\, n\boldsymbol{k}_2}} = \Bigl(\Delta U_{\substack{m'\boldsymbol{k}_1' \\ m\boldsymbol{k}_1}}\,\delta_{n'n} + \Delta U_{\substack{n'\boldsymbol{k}_2' \\ n\boldsymbol{k}_2}}\,\delta_{m'm}\Bigr)\delta_{\boldsymbol{k}_1'+\boldsymbol{k}_2',\, \boldsymbol{k}_1+\boldsymbol{k}_2}; \tag{27.59}$$

here $\Delta U_{\substack{m'\boldsymbol{k}' \\ m\boldsymbol{k}}}$ is determined by (27.31), and

$$\overline{\mathscr{H}}^{(3)}_{\substack{m'\boldsymbol{k}_1',\, n'\boldsymbol{k}_2' \\ n\boldsymbol{k}_2,\, m\boldsymbol{k}_1}} = \sum_{\alpha\beta} Q^{\alpha\beta}_{\substack{m'n' \\ nm}}\, U_{\boldsymbol{k}_1'-\boldsymbol{k}_2}(\boldsymbol{k}_1 - \boldsymbol{k}_2')_\alpha(\boldsymbol{k}_1 - \boldsymbol{k}_2')_\beta\,\delta_{\boldsymbol{k}_1'+\boldsymbol{k}_2',\, \boldsymbol{k}_1+\boldsymbol{k}_2}, \tag{27.60}$$

where

$$Q^{\alpha\beta}_{\substack{m'n' \\ nm}} = \frac{\hbar^2}{m^2}\,\frac{\pi^\alpha_{m'n}\,\pi^\beta_{n'm}}{(E^0_m - E^0_n)^2}. \tag{27.60a}$$

To transform to the $\boldsymbol{r}$-representation, we multiply equation (27.57) with the Hamiltonian (27.58)–(27.60) by $\exp\left[i(\boldsymbol{k}_1'\boldsymbol{r}_1 + \boldsymbol{k}_2'\boldsymbol{r}_2)\right]$ and sum over $\boldsymbol{k}_1'$ and $\boldsymbol{k}_2'$. In so doing, since $\boldsymbol{k}$ and $\boldsymbol{k}'$ are small we may proceed as in (22.14) to replace summation over $\boldsymbol{k}_1$, $\boldsymbol{k}_1'$ and $\boldsymbol{k}_2$, $\boldsymbol{k}_2'$ by summation over the differences $\boldsymbol{q} = \boldsymbol{k}_1' - \boldsymbol{k}_1 = \boldsymbol{k}_2 - \boldsymbol{k}_2'$, and in the last term, which is (27.60), over $\boldsymbol{q} = \boldsymbol{k}_1' - \boldsymbol{k}_2 = \boldsymbol{k}_1 - \boldsymbol{k}_2'$, and then apply the transformation used to derive equation (27.34). The last term (27.60) will contain the sum

$$\sum_{\boldsymbol{k}_1\boldsymbol{k}_2} \bar{\bar{C}}_{m\boldsymbol{k}_1,\, n\boldsymbol{k}_2} \exp\left[i(\boldsymbol{k}_1\boldsymbol{r}_2 + \boldsymbol{k}_2\boldsymbol{r}_1)\right] = F_{mn}(\boldsymbol{r}_2, \boldsymbol{r}_1)$$

as opposed to the other terms, which contain $F_{mn}(\boldsymbol{r}_1, \boldsymbol{r}_2)$. Therefore, in the $\boldsymbol{r}$-representation the operator $\overline{\mathscr{H}}$ including (27.60) will be an integral operator, and the system of equations for the functions $F_{mn}(\boldsymbol{r}_1, \boldsymbol{r}_2)$ will be

$$\sum_{mn} \int d\boldsymbol{r}_1\, d\boldsymbol{r}_2 \Bigl\{\left[\mathscr{H}_{m'm}(\boldsymbol{k}_1)\delta_{n'n} + \mathscr{H}_{n'n}(\boldsymbol{k}_2)\delta_{m'm}\right]\delta(\boldsymbol{r}_1 - \boldsymbol{r}_1')\delta(\boldsymbol{r}_2 - \boldsymbol{r}_2') + \\ + U_{\substack{m'n' \\ m\,n}}\begin{pmatrix}\boldsymbol{r}_1'\boldsymbol{r}_2' \\ \boldsymbol{r}_1\boldsymbol{r}_2\end{pmatrix} - \mathscr{H}^{\text{exch}}_{\substack{m'n' \\ n\,m}}\begin{pmatrix}\boldsymbol{r}_1'\boldsymbol{r}_2' \\ \boldsymbol{r}_1\boldsymbol{r}_2\end{pmatrix}\Bigr\} F_{mn}(\boldsymbol{r}_1, \boldsymbol{r}_2) = EF_{m'n'}(\boldsymbol{r}_1', \boldsymbol{r}_2'). \tag{27.61}$$

Here $\mathscr{H}_{m'm}(\boldsymbol{k})$ and $\mathscr{H}_{n'n}(\boldsymbol{k})$ are the matrices (21.19) defining the spectrum near the extremum of each band, and

$$U_{\substack{m'n' \\ m\,n}}\begin{pmatrix}\boldsymbol{r}_1'\boldsymbol{r}_2' \\ \boldsymbol{r}_1\boldsymbol{r}_2\end{pmatrix} = \left[U(\boldsymbol{r}_1 - \boldsymbol{r}_2)\delta_{m'm}\delta_{n'n} + \Delta U_{m'm}(\boldsymbol{r}_1, \boldsymbol{r}_2)\delta_{n'n} + \right. \\ \left. + \Delta U_{n'n}(\boldsymbol{r}_1, \boldsymbol{r}_2)\delta_{m'm}\right]\delta(\boldsymbol{r}_1 - \boldsymbol{r}_1')\delta(\boldsymbol{r}_2 - \boldsymbol{r}_2'), \tag{27.61a}$$

$$\mathscr{H}^{\text{exch}}_{\substack{m'n' \\ n\,m}}\begin{pmatrix}\boldsymbol{r}_1'\boldsymbol{r}_2' \\ \boldsymbol{r}_1\boldsymbol{r}_2\end{pmatrix} = -\sum_{\alpha\beta} Q^{\alpha\beta}_{\substack{m'n' \\ n\,m}}\,\frac{\partial^2 U(\boldsymbol{r}_1 - \boldsymbol{r}_2)}{\partial r_{1\alpha}\,\partial r_{1\beta}}\,\delta(\boldsymbol{r}_1 - \boldsymbol{r}_2')\delta(\boldsymbol{r}_2 - \boldsymbol{r}_1'). \tag{27.61b}$$

The interaction operator (27.61a) includes, besides the Coulomb potential $U(\boldsymbol{r}_1 - \boldsymbol{r}_2)$, corrections

$$\Delta U_{m'm}(\boldsymbol{r}) = \Delta U^1_{m'm}(\boldsymbol{r}) + \Delta U^2_{m'm}(\boldsymbol{r}) \text{ and } \Delta U_{n'n}(\boldsymbol{r}).$$

defined by (27.34) and (27.35) for each band, and the last term describes the exchange interaction; $Q^{\alpha\beta}_{m'n',\,nm}$ is given in (27.60a).

The operator (27.61) describes the interaction of two electrons. In order to obtain the Hamiltonian for the interaction of an electron m and hole n, we must first reverse the sign of the term $\mathcal{H}_{n'n}(\boldsymbol{k}_2)$ defining the energy of an electron in the valence band, and the signs of the operators defining the interaction of electrons in different bands, since the energy of a hole is minus the energy of the missing electron and the electron-hole interaction is minus the interaction of the electron with the missing electron.

Furthermore, we must remember that if the wave function of the crystal with one electron belongs in state ψ_n to a representation $\mathcal{D}$, then the wave function of the crystal with one hole n, i.e., in the state in which all levels of the valence band except n are filled, belongs to the representation $\mathcal{D}^*$, so that it transforms like $K\psi_n = \psi_{Kn}$. This is because a state which has both an electron and a hole at level n (the whole band is filled) belongs to the identity representation, which occurs only in the product $\mathcal{D}\times\mathcal{D}^*$. Therefore, if the missing electron has, say, wave vector $\boldsymbol{k}$ and spin $+1/2$, the hole must have wave vector $\boldsymbol{k}^h = -\boldsymbol{k}^e$ and spin $-1/2$. Since the creation of an electron implies the destruction of a hole and vice versa, transition from electron to hole dictates that we must also reverse the order of the indices of the valence band. This gives the following rules, governing transition from the electron-electron interaction to the electron-hole interaction:*

$$\begin{gathered}\mathcal{H}^h_{n'n}(\boldsymbol{r}_2,\boldsymbol{r}_2') = -\mathcal{H}^e_{KnKn'}(\boldsymbol{r}_2',\boldsymbol{r}_2),\\ \mathcal{H}^{eh}_{\substack{m'n'\\ m\,n}}\begin{pmatrix}\boldsymbol{r}_1'\,\boldsymbol{r}_2'\\ \boldsymbol{r}_1\,\boldsymbol{r}_2\end{pmatrix} = -\mathcal{H}^{ee}_{\substack{m'Kn\\ m\,Kn'}}\begin{pmatrix}\boldsymbol{r}_1'\,\boldsymbol{r}_2\\ \boldsymbol{r}_1\,\boldsymbol{r}_2'\end{pmatrix}.\end{gathered} \tag{27.62}$$

Note that, by (18.46), $p^{\alpha}_{ns} = -p^{\alpha}_{KsKn}$, and so the elements of the matrices $\mathcal{H}$ (21.29) and $\boldsymbol{G}^{\alpha\beta}$ (27.32) satisfy the relations $\mathcal{H}_{KnKn'} = \mathcal{H}_{n'n}$, $G^{\alpha\beta}_{KnKn'} = G^{\alpha\beta}_{n'n}$, and for those of the matrix $\boldsymbol{g}$ (27.33), $g^{\gamma}_{KnKn'} = -g^{\gamma}_{n'n}$.** Here the term $\mathcal{H}^h_{n'n}(\boldsymbol{k}_2) = -\mathcal{H}^e_{KnKn'}(-\boldsymbol{k}_2) = -\mathcal{H}^e_{n'n}(\boldsymbol{k}_2)$ in (27.62) is the usual Hamiltonian, defining the hole spectrum in the effective mass approximation.

Finally, then, the electron-hole interaction operator is

$$\begin{aligned}\mathcal{H}^{eh}_{\substack{m'n'\\ m\,n}}\begin{pmatrix}\boldsymbol{r}_1'\,\boldsymbol{r}_2'\\ \boldsymbol{r}_1\,\boldsymbol{r}_2\end{pmatrix} = \Big[&\mathcal{H}^e_{m'm}(\boldsymbol{k}_1)\,\delta_{n'n} + \mathcal{H}^h_{n'n}(\boldsymbol{k}_2)\,\delta_{m'm} + U^{eh}(\boldsymbol{r})\,\delta_{m'm}\delta_{n'n} + \Delta U^1_{\substack{m'n'\\ m\,n}}(\boldsymbol{r}) + \Delta U^2_{\substack{m'n'\\ m\,n}}(\boldsymbol{r}_1,\boldsymbol{r}_2)\Big]\times\\ &\times\delta(\boldsymbol{r}_1-\boldsymbol{r}_1')\,\delta(\boldsymbol{r}_2-\boldsymbol{r}_2') + \Delta U_{\substack{m'n'\\ m\,n}}\begin{pmatrix}\boldsymbol{r}_1'\,\boldsymbol{r}_2'\\ \boldsymbol{r}_1\,\boldsymbol{r}_2\end{pmatrix},\end{aligned} \tag{27.63}$$

with

$$\begin{gathered}\mathcal{H}^e_{m'm}(\boldsymbol{k}) = \sum_{\alpha\beta}\mathcal{H}^{\alpha\beta}_{m'm}\,k_\alpha k_\beta,\\ \mathcal{H}^h_{n'n}(\boldsymbol{k}) = \sum_{\alpha\beta} -\mathcal{H}^{\alpha\beta}_{n'n}\,k_\alpha k_\beta,\end{gathered} \tag{27.63a}$$

---

* These rules are easily obtained if we write the Hamiltonian (27.61) in the second-quantized formalism. The derivation is given in the supplement to this section.

** The symbol $\mathcal{H}_{KnKn'}$ means that if the wave functions of the holes ψ^h_n are taken as the basis, the matrix elements of $\mathcal{H}$ must be calculated for the functions $K\psi^h_n$.

$$U^{eh}(\boldsymbol{r}) = -\frac{e^2}{\varkappa r}, \quad \text{where} \quad \boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2, \tag{27.63b}$$

$$\Delta U^1_{\substack{m'n'\\ m\,n}}(\boldsymbol{r}) = -\frac{e^2}{\varkappa}\left\{\sum_{\alpha\beta}\left(\frac{3r_\alpha r_\beta}{r^5} - \frac{1}{r^3}\delta_{\alpha\beta}\right)\left(G^{\alpha\beta}_{m'm}\delta_{n'n} + \right.\right.$$
$$\left.\left. + G^{\alpha\beta}_{n'n}\delta_{m'm}\right) - \frac{4\pi}{3}\delta(\boldsymbol{r})\left[\operatorname{Tr} G_{m'm}\delta_{n'n} + \operatorname{Tr} G_{n'n}\delta_{m'm}\right]\right\}, \tag{27.63c}$$

$$\Delta U^2_{\substack{m'n'\\ m\,n}}(\boldsymbol{r}_1, \boldsymbol{r}_2) = -\frac{e^2}{\varkappa r^3}\left[\left(\left[\boldsymbol{r}\nabla_{\boldsymbol{r}_1}\right]\boldsymbol{g}_{m'm}\right)\delta_{n'n} - \left(\left[\boldsymbol{r}\nabla_{\boldsymbol{r}_2}\right]\boldsymbol{g}_{n'n}\right)\delta_{m'm}\right].$$

The last operator ΔU in (27.63) is nonlocalized.

By (27.61b) and (27.62),

$$\Delta U_{\substack{m'n'\\ m\,n}}\begin{pmatrix}\boldsymbol{r}_1'\boldsymbol{r}_2'\\ \boldsymbol{r}_1\boldsymbol{r}_2\end{pmatrix} = \mathcal{H}^{\text{exch}}_{\substack{m'n\\ n'm}}\begin{pmatrix}\boldsymbol{r}_1'\boldsymbol{r}_2\\ \boldsymbol{r}_1\boldsymbol{r}_2'\end{pmatrix} = -\sum_{\alpha\beta} Q^{\alpha\beta}_{\substack{m'Kn\\ Kn'm}}\frac{\partial^2 U(\boldsymbol{r}_1 - \boldsymbol{r}_2')}{\partial r_{1\alpha}\,\partial r_{1\beta}}\delta(\boldsymbol{r}_1 - \boldsymbol{r}_2)\,\delta(\boldsymbol{r}_1' - \boldsymbol{r}_2'), \tag{27.64}$$

where

$$Q^{\alpha\beta}_{\substack{m'Kn\\ Kn'm}} = \frac{\hbar^2}{m^2}\frac{\pi^\alpha_{m'Kn'}\pi^\beta_{Knm}}{\left(E^0_m - E^0_n\right)^2}. \tag{27.65}$$

Note that if the spatial dispersion of the dielectric constant $\varkappa$ is introduced, we get corrections to the energy of the exciton (and to that of an impurity center), of the same order as $\Delta U^{1,2}$.

It is convenient to transform from the variables $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ to variables

$$\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2, \quad \mathscr{R} = \boldsymbol{a}\boldsymbol{r}_1 + \boldsymbol{b}\boldsymbol{r}_2. \tag{27.66}$$

The matrices $\boldsymbol{a}$ and $\boldsymbol{b}$ may be chosen arbitrarily, provided they satisfy the condition $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{I}$ to ensure that the Jacobian is such that $\left|\frac{\partial(\boldsymbol{r}_1, \boldsymbol{r}_2)}{\partial(\boldsymbol{r}, \mathscr{R})}\right| = 1$. In the new variables,

$$\boldsymbol{k}_1 = \boldsymbol{k} + \tilde{\boldsymbol{a}}\mathscr{K}, \quad \boldsymbol{k}_2 = -\boldsymbol{k} + \tilde{\boldsymbol{b}}\mathscr{K}, \tag{27.66a}$$

where

$$\boldsymbol{k} = -i\nabla_{\boldsymbol{r}}, \qquad \mathscr{K} = -i\nabla_{\mathscr{R}}.$$

We then have

$$\mathscr{K} = \boldsymbol{k}_1 + \boldsymbol{k}_2, \quad \boldsymbol{k} = \tilde{\boldsymbol{b}}\boldsymbol{k}_1 - \tilde{\boldsymbol{a}}\boldsymbol{k}_2. \tag{27.66b}$$

Therefore,

$$\boldsymbol{k}_1\boldsymbol{r}_1 + \boldsymbol{k}_2\boldsymbol{r}_2 = \boldsymbol{k}\boldsymbol{r} + \mathscr{K}\mathscr{R}. \tag{27.66c}$$

In the new variables (27.66), the operator (27.64) becomes

$$\Delta U_{\substack{m'n'\\ m\,n}}\begin{pmatrix}\boldsymbol{r}' & \mathscr{R}'\\ \boldsymbol{r} & \mathscr{R}\end{pmatrix} = -\sum_{\alpha\beta} Q_{\substack{m'Kn\\ Kn'm}}\frac{\partial^2 U(\mathscr{R} - \mathscr{R}')}{\partial\mathscr{R}_\alpha\,\partial\mathscr{R}_\beta}\delta(\boldsymbol{r})\,\delta(\boldsymbol{r}'). \tag{27.67}$$

Since the operator (27.63) is translation-invariant with respect to $\mathscr{R}$, the effective mass wave function $F^l_{\mathscr{K}}(\boldsymbol{r}, \mathscr{R})$ of the exciton may be written as a column matrix:

$$F^l_{\mathscr{K}}(\boldsymbol{r}, \mathscr{R}) = \frac{1}{\sqrt{\mathscr{V}}}e^{i\mathscr{K}\mathscr{R}}\left\|\begin{matrix} f^l_{\mathscr{K}m_1n_1}\\ f^l_{\mathscr{K}m_1n_2}\\ \cdots\\ f^l_{\mathscr{K}m_2n_1}\\ \cdots\end{matrix}\right\|. \tag{27.68}$$

The matrix element of the operator (27.67) between the functions (27.68) is

$$\Delta U_{l'l\boldsymbol{\mathscr{K}}} = \frac{4\pi e^2}{\varkappa} \sum_{\alpha\beta m n m' n'} Q^{\alpha\beta}_{\substack{m'Kn\\ Kn'm}} \frac{\mathscr{K}_\alpha \mathscr{K}_\beta}{\mathscr{K}^2} f^{*l'}_{\boldsymbol{\mathscr{K}} m'n'}(0) f^{l}_{\boldsymbol{\mathscr{K}} mn}(0). \tag{27.69}$$

Note that although the last term originates in the exchange interaction of electrons in the valence band and the conduction band, it describes the long-range electron-hole interaction, which may be considered as the result of virtual recombination and generation of an exciton /22.6, 22.8/. In the literature it is therefore known variously as the exchange, resonance, or annihilation interaction.

It is clear from (27.69) that ΔU depends essentially on the direction of the vector $\boldsymbol{\mathscr{K}}$. If we introduce corrections due to the short-range part of the potential $U(\mathbf{r})$, the Hamiltonian $\mathscr{H}$ will include exchange terms independent of the direction of $\boldsymbol{\mathscr{K}}$.

Inclusion of short-range potential. To allow for these terms we must retain the terms with $M \neq 0$ and $L \neq 0$ in (27.55). We first consider the non-exchange terms. Disregarding the dependence of $U_{\mathbf{q}_1+\mathbf{b}_M}$ on the small quantity $\mathbf{q}_1 = \mathbf{k}_1' - \mathbf{k}_1$ in (27.55) and noting that when $\mathbf{q}_1 \ll \mathbf{b}_M$ and $\mathbf{q}_2 = \mathbf{k}_2' - \mathbf{k}_2 \ll \mathbf{b}_L$

$$\delta_{\mathbf{b}_M+\mathbf{q}_1,\, -\mathbf{b}_L-\mathbf{q}_2} = \delta_{\mathbf{b}_M,\, -\mathbf{b}_L} \delta_{\mathbf{q}_1,\, -\mathbf{q}_2},$$

we obtain

$$\begin{aligned} \delta U_{\substack{m'\mathbf{k}_1',\, n'\mathbf{k}_2'\\ m\mathbf{k}_1,\, n\mathbf{k}_2}} &= \mathscr{V}^2 \sum_{\substack{M\neq 0\\ L\neq 0}} B_M^{m'm} B_L^{n'n} U_{\mathbf{b}_M} \delta_{\mathbf{b}_M,\, -\mathbf{b}_L} \delta_{\mathbf{q}_1,\, -\mathbf{q}_2} = \\ &= \Bigl(\mathscr{V}^2 \sum_{ML} B_M^{m'm} B_L^{n'n} U_{\mathbf{b}_M} \delta_{\mathbf{b}_M,\, -\mathbf{b}_L} - U_0 \delta_{m'm} \delta_{n'n}\Bigr) \delta_{\mathbf{q}_1,\, -\mathbf{q}_2}. \end{aligned} \tag{27.70}$$

It follows from (27.55) that the first term in this equation is equal to $U_{\substack{m'0,\, n'0\\ m0,\, n0}} = U_{\substack{m'n'\\ mn}}$, and so

$$\delta U_{\substack{m'\mathbf{k}_1',\, n'\mathbf{k}_2'\\ m\mathbf{k}_1,\, n\mathbf{k}_2}} = \Bigl(U_{\substack{m'n'\\ mn}} - U_0 \delta_{m'm} \delta_{n'n}\Bigr) \delta_{\mathbf{q}_1,\, -\mathbf{q}_2}, \tag{27.71}$$

where

$$U_{\substack{m'n'\\ mn}} = \frac{1}{\mathscr{V}^2} \int\!\!\int d\mathbf{x}_1\, d\mathbf{x}_2\, U(\mathbf{r}_1 - \mathbf{r}_2)\, u^*_{m'}(\mathbf{x}_1)\, u_m(\mathbf{x}_1)\, u^*_{n'}(\mathbf{x}_2)\, u_n(\mathbf{x}_2). \tag{27.71a}$$

Going over to the $\mathbf{r}$-representation, i.e., multiplying by $\exp\,[i(\mathbf{q}_1\mathbf{r}_1 + \mathbf{q}_2\mathbf{r}_2)]$ and summing over $\mathbf{q}_1$ and $\mathbf{q}_2$, we find

$$\delta U_{\substack{m'n'\\ mn}}(\mathbf{r}_1, \mathbf{r}_2) = \mathscr{V}\Bigl(U_{\substack{m'n'\\ mn}} - U_0 \delta_{m'm} \delta_{n'n}\Bigr) \delta(\mathbf{r}_1 - \mathbf{r}_2). \tag{27.72}$$

Now, using (27.62) to go over from the electron-electron to the electron-hole interaction, we have

$$\delta U^{eh}_{\substack{m'n'\\ mn}}(\mathbf{r}_1, \mathbf{r}_2) = -\mathscr{V}\Bigl(U_{\substack{m'n'\\ mn}} - U_0 \delta_{m'm} \delta_{n'n}\Bigr) \delta(\mathbf{r}_1 - \mathbf{r}_2), \tag{27.73}$$

since $U_{m'Kn,\, mKn'} = U_{m'n',\, mn}$.

Now for the exchange terms:

$$\delta\mathscr{H}^{\text{exch}}_{\substack{m'\mathbf{k}_1',\, n'\mathbf{k}_2'\\ n\mathbf{k}_2,\, m\mathbf{k}_1}} = \int\!\!\int d\mathbf{x}_1\, d\mathbf{x}_2\, \varphi^*_{m'\mathbf{k}_1'}(\mathbf{x}_1)\, \varphi^*_{n'\mathbf{k}_2'}(\mathbf{x}_2)\, \varphi_{n\mathbf{k}_2}(\mathbf{x}_1)\, \varphi_{m\mathbf{k}_1}(\mathbf{x}_2)\, U(\mathbf{r}_1 - \mathbf{r}_2).$$

The first term $B_0^{m'n}$ in the expansion

$$\sum_{a} \psi^*_{m'}(\boldsymbol{x})\,\psi_n(\boldsymbol{x}) = \mathscr{V} \sum_{M} B_M^{m'n}\, e^{-i\boldsymbol{b}_M \boldsymbol{r}}$$

vanishes, since the functions u_m and u_n are orthogonal. If $\boldsymbol{b}_M \neq 0$, as above, we may disregard the dependence of the matrix element on $\boldsymbol{k}$ and $\boldsymbol{k}'$. Then we obtain an expression similar to (27.71):

$$\delta\mathscr{H}^{\text{exch}}_{\substack{m'\boldsymbol{k}'_1,\, n'\boldsymbol{k}'_2 \\ n\ \boldsymbol{k}_2,\, m\boldsymbol{k}_1}} = U_{\substack{m'n' \\ n\ m}}\, \delta_{\boldsymbol{k}'_1+\boldsymbol{k}'_2,\ \boldsymbol{k}_1+\boldsymbol{k}_2}, \tag{27.74}$$

where

$$U_{\substack{m'n' \\ nm}} = \frac{1}{\mathscr{V}^2} \int\!\!\int U(\boldsymbol{r}_1 - \boldsymbol{r}_2)\, \psi^*_{m'}(\boldsymbol{x}_1)\, \psi_m(\boldsymbol{x}_2)\, \psi_n(\boldsymbol{x}_1)\, \psi^*_{n'}(\boldsymbol{x}_2)\, d\boldsymbol{x}_1\, d\boldsymbol{x}_2. \tag{27.75}$$

Hence

$$\delta\mathscr{H}^{\text{exch}}_{\substack{m'n' \\ m\ n}} \begin{pmatrix} \boldsymbol{r}'_1\, \boldsymbol{r}'_2 \\ \boldsymbol{r}_1\, \boldsymbol{r}_2 \end{pmatrix} = \mathscr{V} U_{\substack{m'n' \\ n\ m}}\, \delta(\boldsymbol{r}_1 - \boldsymbol{r}_2)\, \delta(\boldsymbol{r}_1 - \boldsymbol{r}'_1)\, \delta(\boldsymbol{r}_2 - \boldsymbol{r}'_2). \tag{27.76}$$

It is evident from (27.76) and (27.62) that in transition to the electron-hole interaction the operator $\delta\mathscr{H}^{\text{exch}}$, like δU, retains its localized nature:

$$\delta\mathscr{H}^{\text{exch}\ eh}_{\substack{m'n' \\ m\ n}} (\boldsymbol{r}_1, \boldsymbol{r}_2) = -\mathscr{V} U_{\substack{m'Kn \\ Kn'm}}\, \delta(\boldsymbol{r}_1 - \boldsymbol{r}_2). \tag{27.77}$$

The integral (27.75), which defines the order of magnitude of this exchange splitting, is close to the exchange interaction value in the corresponding atoms /22.5/. Thus the exchange splitting of the exciton is a factor of approximately $(a_0/a_B)^3$ lower than the atomic splitting, where a_B is the Bohr radius of the exciton and a_0 lattice constant.

Indirect excitons. For an indirect exciton, $\boldsymbol{k}_c \neq \boldsymbol{k}_v$, the wave function $\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ must be expressed as a linear combination of products of the functions

$$\begin{aligned} \varphi_{s\boldsymbol{k}} &= \frac{1}{\sqrt{\mathscr{V}}}\, \psi_{s\boldsymbol{k}_c} e^{i\boldsymbol{k}\boldsymbol{r}} = \frac{1}{\sqrt{\mathscr{V}}}\, u_{s\boldsymbol{k}_c} e^{i(\boldsymbol{k}_c+\boldsymbol{k})\boldsymbol{r}}, \\ \varphi_{t\boldsymbol{k}} &= \frac{1}{\sqrt{\mathscr{V}}}\, \psi_{t\boldsymbol{k}_v} e^{i\boldsymbol{k}\boldsymbol{r}} = \frac{1}{\sqrt{\mathscr{V}}}\, u_{t\boldsymbol{k}_v} e^{i(\boldsymbol{k}_v+\boldsymbol{k})\boldsymbol{r}}, \end{aligned} \tag{27.78}$$

where $\psi_{s\boldsymbol{k}_c}$ and $\psi_{t\boldsymbol{k}_v}$ are eigenfunctions of $\mathscr{H}_0$ at the points $\boldsymbol{k}_c$ and $\boldsymbol{k}_v$, respectively, for all bands. The function $\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$, which satisfies the antisymmetry condition, has the form

$$\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{1}{\sqrt{2}} \sum_{st} F_{st}(\boldsymbol{r}_1, \boldsymbol{r}_2)\, \psi_{s\boldsymbol{k}_c}(\boldsymbol{x}_1)\, \psi_{t\boldsymbol{k}_v}(\boldsymbol{x}_2) - F_{st}(\boldsymbol{r}_2, \boldsymbol{r}_1)\, \psi_{s\boldsymbol{k}_c}(\boldsymbol{x}_2)\, \psi_{t\boldsymbol{k}_v}(\boldsymbol{x}_1). \tag{27.79}$$

If we now write $F_{st}(\boldsymbol{r}_1, \boldsymbol{r}_2)$ in the form (27.51a), we obtain, instead of (27.51),

$$\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{s\boldsymbol{k}_1 t\boldsymbol{k}_2} C_{s\boldsymbol{k}_1,\, t\boldsymbol{k}_2}\, \varphi^a_{s\boldsymbol{k}_1,\, t\boldsymbol{k}_2}(\boldsymbol{x}_1, \boldsymbol{x}_2), \tag{27.80}$$

where

$$\varphi^a_{s\boldsymbol{k}_1,\, t\boldsymbol{k}_2} = \frac{1}{\sqrt{2}} \left[\varphi_{s\boldsymbol{k}_1}(\boldsymbol{x}_1)\, \varphi_{t\boldsymbol{k}_2}(\boldsymbol{x}_2) - \varphi_{s\boldsymbol{k}_1}(\boldsymbol{x}_2)\, \varphi_{t\boldsymbol{k}_2}(\boldsymbol{x}_1)\right]. \tag{27.81}$$

Substituting (27.81) into (27.50), multiplying on the left by $\varphi^{*a}_{s'\boldsymbol{k}'_1,\, t'\boldsymbol{k}'_2}(\boldsymbol{x}_1, \boldsymbol{x}_2)$ and integrating over $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, we obtain a system of equations for $C_{s\boldsymbol{k}_1,\, t\boldsymbol{k}_2}$

similar to (27.53), in which the diagonal terms $\mathcal{H}_0$ and the interband terms $\mathcal{H}_2$ are defined, as for direct excitons, by (27.54) and (27.56), respectively.

If we disregard the short-range forces, i.e., neglect all components $U_{\boldsymbol{k}}$ with $\boldsymbol{k}$ containing either reciprocal lattice vectors $\boldsymbol{b}_M$ or the difference $\boldsymbol{k}_c - \boldsymbol{k}_v$, the expression for $\mathcal{H}_1$ will coincide with (27.55). However, as opposed to the case of direct excitons, where we must keep the coefficients C_{st} with $s = m$, $t = n$ and $s = n$, $t = m$ in (27.57), here we have conduction band functions $\psi_{m\boldsymbol{k}_c}$ only among the function $\psi_{s\boldsymbol{k}_c}$ and valence band functions $\psi_{m\boldsymbol{k}_v}$ only among the functions $\psi_{t\boldsymbol{k}_v}$. Therefore, if only the long-range potential is included, there are no exchange terms for indirect excitons, while the non-exchange terms are defined by the same equations (27.63a)–(27.63c) as for the direct exciton.

If we include the short-range forces, i.e. retain in U all the components $U_{\boldsymbol{k}}$ with large $\boldsymbol{k}$ and disregard their dependence on the small quantities $\boldsymbol{k}_1$, $\boldsymbol{k}_1'$, $\boldsymbol{k}_2$ and $\boldsymbol{k}_2'$ in (27.74), as in (27.70), the result is the same equation (27.73) for the non-exchange terms. The expression for $U_{m'n',\,mn}$ is similar to (27.71a)

Since the expansion functions (27.81) represent an antisymmetrized product, it follows that if the short-range effect is included $\mathcal{H}^{eh}$ will also include terms as defined by (27.77), where, for indirect excitons,

$$U_{\substack{m\,Kn\\ Kn'm}} = \frac{1}{\mathcal{V}^2}\int d\boldsymbol{x}_1\, d\boldsymbol{x}_2\, u^*_{m'\boldsymbol{k}_c}(\boldsymbol{x}_2)\, K u_{n'\boldsymbol{k}_v}(\boldsymbol{x}_1) \times$$
$$\times\, U(\boldsymbol{r}_1 - \boldsymbol{r}_2)\, u_{m\boldsymbol{k}_c}(\boldsymbol{x}_2)\left(K u_{n\boldsymbol{k}_v}(\boldsymbol{x}_2)\right)^* e^{i(\boldsymbol{k}_v - \boldsymbol{k}_c)(\boldsymbol{r}_1 - \boldsymbol{r}_2)}. \tag{27.82}$$

Excitons in different band structures. We consider the main types of band structures, without introducing the correction terms (27.63).

If both bands are degenerate when spin is neglected, the matrices $\mathcal{H}^c$ and $\mathcal{H}^h$ are diagonal and the system of equations (27.63) reduces to the equation

$$\left(\sum_{\alpha\beta} \frac{\hbar^2}{2m^e_{\alpha\beta}} k_{1\alpha} k_{1\beta} + \frac{\hbar^2}{2m^h_{\alpha\beta}} k_{2\alpha} k_{2\beta} - \frac{e^2}{\varkappa r} - E\right) F(\boldsymbol{r}_1, \boldsymbol{r}_2) = 0. \tag{27.83}$$

If the constant energy ellipsoids of the two bands are coaxial, then, setting $a^{ll} = m^e_{ll}/(m^e_{ll} + m^h_{ll})$ in (27.66), we obtain an equation for $f_{\mathcal{K}}(\boldsymbol{r})$ in terms of the principal axes of the ellipsoids:

$$\left(\sum_{\alpha} \frac{\hbar^2}{2m_\alpha} k_\alpha^2 - \frac{e^2}{\varkappa r} + \frac{\hbar^2 \mathcal{K}_\alpha^2}{2M_\alpha} - E\right) f_{\mathcal{K}}(\boldsymbol{r}) = 0. \tag{27.84}$$

Here the orbital kinetic energy does not depend on the translatory motion in this approximation, and the equation for $f_{\mathcal{K}}(\boldsymbol{r})$ is similar to the corresponding equation (27.15) for an impurity center, except that $1/m^e_\alpha$ is replaced by

$$\frac{1}{m_\alpha} = \frac{1}{m^e_\alpha} + \frac{1}{m^h_\alpha}, \tag{27.85}$$

and the energy of the translatory motion is defined by the overall mass of the electron and the hole:

$$M_\alpha = m^e_\alpha + m^h_\alpha. \tag{27.86}$$

For spherical bands, the binding energy of the exciton has a hydrogen-like spectrum

$$E_n = -\frac{me^4}{2\hbar^2\varkappa^2 n^2} + \frac{\hbar^2\mathcal{K}^2}{2M}. \tag{27.87}$$

If the constant energy surfaces are ellipsoids of revolution with the same principal axes, the trial function $f_{\mathscr{K}}(\boldsymbol{r})$ for the ground state may be taken in the form (27.17) and its energy can be determined from Figure 28. In this case, the representation according to which the total wave function of a direct exciton transforms at the point $\mathscr{K}=0$ is the product of three representations: $\mathscr{D}_f$, according to which the functions $f_0(\boldsymbol{r})$ transform, and representations $\mathscr{D}_m$ and $\mathscr{D}_n^*$ of the groups $G_{\boldsymbol{k}}$, according to which the basis functions $\psi_{m\boldsymbol{k}_0}$ and $K\psi_{n\boldsymbol{k}_0}$ transform:

$$\mathscr{D}^{ex}=\mathscr{D}_f\times\mathscr{D}_m\times\mathscr{D}_n^*. \tag{27.88}$$

This representation is usually reducible, and then a perturbation due to the exchange interaction splits it into irreducible representations.

If only one of the bands, say the m band, is nondegenerate when spin is neglected, while the n band is degenerate, equations (27.63) split into two identical systems of N_n equations. For degenerate bands, the binding energy itself depends on $\mathscr{K}$, i.e., the orbital motion may cause splitting of the exciton band, as in degenerate electron bands.

For $\mathscr{K}=0$, the corresponding equations for a direct exciton in Ge coincide in form with the analogous equations (27.1)–(27.2) for an impurity center except that $|A|$ is replaced by $|A|+(\hbar^2/2m^e)$, where m^e is the effective mass of an electron in the conduction band at the point $\boldsymbol{k}=0$. The variational functions for the ground state, which, as in the case of an impurity center, is fourfold degenerate, may be chosen in the form (27.23). The calculation /22.4/ yields a value 0.00138 ev for the binding energy, in excellent agreement with the experimental value; $r_1=320\,\text{Å}$, $r_2=220\,\text{Å}$.

References /22.12, 21.14/ propose a different approximation method to solve equation (27.1) for a shallow impurity center or an exciton in the case of a degenerate valence band of the Ge type. The gist of the method is to consider a Hamiltonian (27.2) with operator $\mathscr{H}(\boldsymbol{k})$ defined by an equation similar to (26.12), isolate the spherical part

$$\overline{\mathscr{H}}(\boldsymbol{k})=\left(\frac{\hbar^2}{2m_e}+\left(\gamma_1+\frac{5}{2}\gamma_2\right)\right)k^2-2\bar{\gamma}(\boldsymbol{J}\boldsymbol{k})^2+V(\boldsymbol{r}), \tag{27.88a}$$

and treat the remaining nonspherical part

$$\mathscr{H}'(\boldsymbol{k})=-2(\gamma_2-\bar{\gamma})\sum_i J_i^2k_i^2-4(\gamma_3-\bar{\gamma})\sum_{i>j}[J_iJ_j]\,k_ik_j \tag{27.88b}$$

as a perturbation. The eigenfunctions of the Hamiltonian (27.88a) are products of the eigenfunctions of the operator $\boldsymbol{F}=\boldsymbol{L}+\boldsymbol{J}$ and radial functions defined by a system of two differential equations, evaluated by numerical methods. Ref. /22.12/ takes $\bar{\gamma}=\frac{1}{5}(2\gamma_2+3\gamma_3)$, and then $\gamma_2-\bar{\gamma}=\frac{3}{5}(\gamma_2-\gamma_3)$, $\gamma_3-\bar{\gamma}=-\frac{2}{5}(\gamma_2-\gamma_3)$. In calculating the spectrum, it is apparently more convenient to choose $\bar{\gamma}$ in such a way that the corrections to the energy of a given term from $\mathscr{H}'$ vanish to a first approximation.

Since the ground state functions transform according to the representation Γ_8^+, the total exciton function transforms according to the representation $\mathscr{D}^{ex}=\Gamma_8^+\times\Gamma_7^-=\Gamma_{12}^-+\Gamma_{15}^-+\Gamma_{25}^-$. Thus the exchange interaction splits this eightfold degenerate state into two threefold and one twofold degenerate states.

The total wave function for indirect excitons transforms according to a representation of the intersection G' of the little groups G_{k_c} and G_{k_v}. Representations $\mathscr{D}_m$ and $\mathscr{D}_n^*$ of the groups G_{k_c} and G_{k_v} become respectively representations $\mathscr{D}_m'$ and $\mathscr{D}_n'$ of G', which in principle may be reducible.

If the envelope function f transforms according to a representation $\mathscr{D}_f$ of the group G' in the case of simple bands, the total function transforms according to the representation $\mathscr{D}_f \times \mathscr{D}_m' \times \mathscr{D}_n^*$, and this state splits only as a result of the exchange interaction.

If one of the bands, say the valence band, is degenerate and the representation $\mathscr{D}_n^*$ is reducible in the group G', splitting of the exciton state will also occur when the exchange interaction is neglected. For example, for indirect excitons in Ge and Si, the eightfold degenerate ground state splits into two fourfold degenerate state, since the representation Γ_8^+ splits in the group $C_{3v} \times C_i$ (point L) into $L_4 + L_5$ and L_6, and in the group C_{4v} (point Δ) into Δ_6 and Δ_7.* The exchange interaction will cause additional splitting of these representations.

Transition from electron-electron to electron-hole representation (Supplement to §27)

In the second-quantized formalism, the operator (27.61) may be written

$$\begin{aligned}\mathscr{H}^{ee} = \sum_{\substack{mm' \\ nn'}} \mathscr{H}_{m'm}(\mathbf{r}_1', \mathbf{r}_1)\, \psi_{m'}^{+}(\mathbf{r}_1')\, \psi_m(\mathbf{r}_1)\, \delta(\mathbf{r}_2' - \mathbf{r}_2)\, \delta_{n'n} + \\ + \mathscr{H}_{n'n}(\mathbf{r}_2', \mathbf{r}_2)\, \psi_{n'}^{+}(\mathbf{r}_2')\, \psi_n(\mathbf{r}_2)\, \delta(\mathbf{r}_1' - \mathbf{r}_1)\, \delta_{m'm} + \\ + \mathscr{H}_{\substack{m'n' \\ m\,n}}\begin{pmatrix} \mathbf{r}_1' \, \mathbf{r}_2' \\ \mathbf{r}_1 \, \mathbf{r}_2 \end{pmatrix} \psi_{m'}^{+}(\mathbf{r}_1')\, \psi_{n'}^{+}(\mathbf{r}_2')\, \psi_m(\mathbf{r}_1)\, \psi_n(\mathbf{r}_2).\end{aligned} \qquad (27.89)$$

The commutators of the electron creation operator $\psi_m^{+}(\mathbf{r})$ and destruction operator $\psi_m(\mathbf{r})$ satisfy the relations

$$\begin{gathered}\{\psi_m^{+}(\mathbf{r})\, \psi_{m'}(\mathbf{r}')\} = \delta_{mm'}\delta(\mathbf{r} - \mathbf{r}'), \\ \{\psi_m^{+}(\mathbf{r})\, \psi_{m'}^{+}(\mathbf{r}')\} = \{\psi_m(\mathbf{r})\, \psi_{m'}(\mathbf{r}')\} = 0.\end{gathered} \qquad (27.90)$$

The average over the vacuum is

$$\langle 0 | \psi_m(\mathbf{r})\, \psi_{m'}^{+}(\mathbf{r}') | 0 \rangle = \delta(\mathbf{r} - \mathbf{r}')\, \delta_{mm'}. \qquad (27.91)$$

To derive equation (27.61) in the $\mathbf{r}$-representation, we must apply the operator $\mathscr{H}$ to the function $F_{m''n''}(\mathbf{r}_1'', \mathbf{r}_2'')\, \psi_{m''}^{+}(\mathbf{r}_1'')\, \psi_{n''}^{+}(\mathbf{r}_2'') | 0 \rangle$, multiply on the left by $\langle 0 | \psi_{m'''}(\mathbf{r}_1''')\, \psi_{n'''}(\mathbf{r}_2''')$, and average, integrating over $\mathbf{r}_{1,2}$, $\mathbf{r}_{1,2}'$ and $\mathbf{r}_{1,2}''$.

We define hole creation and destruction operators $\varphi_n^{+}(\mathbf{r})$ and $\varphi_n(\mathbf{r})$ by

$$\varphi_n(\mathbf{r}) = \psi_{Kn}^{+}(\mathbf{r}), \quad \varphi_n^{+}(\mathbf{r}) = \psi_{Kn}(\mathbf{r}). \qquad (27.92)$$

We express ψ_n in terms of φ_n, substitute into (27.89) and, using (27.90), bring $\mathscr{H}$ to a form which differs from (27.89) only in that ψ_n is replaced by φ_n. For this purpose we must change the notation, replacing $K^{-1}n$ by n' and

---

* According to data for Ge in Kleiner, W. H. and B. Lax, Progress in Semicond., **5**, this splitting is $1 \cdot 10^{-3}$ ev for binding energy $3.5 \cdot 10^{-3}$ ev.

$K^{-1}n'$ by n, $\mathbf{r}_2$ by $\mathbf{r}_2'$ and by $\mathbf{r}_2$. Omitting terms that describe the interaction of a conduction electron with all the electrons in a filled band (i.e., those of these terms actually included in $\mathcal{H}_0$), we obtain

$$\begin{aligned}\mathcal{H}^{eh} = \sum_{mm'nn'} \mathcal{H}_{m'm}(\mathbf{r}_1', \mathbf{r}_1)\, \psi_{m'}^{+}(\mathbf{r}_1')\, \psi_m(\mathbf{r}_1)\, \delta(\mathbf{r}_2' - \mathbf{r}_2)\, \delta_{n'n} - \\ - \mathcal{H}_{Kn,\, Kn'}(\mathbf{r}_2, \mathbf{r}_2')\, \varphi_{n'}^{+}(\mathbf{r}_2')\, \varphi_n(\mathbf{r}_2)\, \delta(\mathbf{r}_1' - \mathbf{r}_1)\, \delta_{m'm} - \\ - \mathcal{H}_{\substack{m'Kn \\ m\,Kn'}} \begin{pmatrix} \mathbf{r}_1' & \mathbf{r}_2 \\ \mathbf{r}_1 & \mathbf{r}_2' \end{pmatrix} \psi_{m'}^{+}(\mathbf{r}_1')\, \varphi_{n'}^{+}(\mathbf{r}_2')\, \psi_m(\mathbf{r}_1)\, \varphi_n(\mathbf{r}_2). \qquad (27.93)\end{aligned}$$

Applying these operators to the function $F_{m''n''}(\mathbf{r}_1'', \mathbf{r}_2'')\, \psi_{m''}^{+}(\mathbf{r}_1'')\, \varphi_{n''}^{+}(\mathbf{r}_2'')\,|\,0\rangle$, multiplying on the left by $\langle 0\,|\,\psi_{m'''}(\mathbf{r}_1''')\, \varphi_{n'''}(\mathbf{r}_2''')$, and performing the necessary integration, we obtain the equations in the $\mathbf{r}$-representation, which differ from (27.61) by the substitution:

$$\begin{aligned}\mathcal{H}_{n'n}(\mathbf{r}_2', \mathbf{r}_2) &\longrightarrow \mathcal{H}_{KnKn'}(\mathbf{r}_2, \mathbf{r}_2'), \\ \mathcal{H}_{\substack{m'n' \\ m\,n}} \begin{pmatrix} \mathbf{r}_1' & \mathbf{r}_2' \\ \mathbf{r}_1 & \mathbf{r}_2 \end{pmatrix} &\longrightarrow \mathcal{H}_{\substack{m'Kn \\ m\,Kn'}} \begin{pmatrix} \mathbf{r}_1' & \mathbf{r}_2 \\ \mathbf{r}_1 & \mathbf{r}_2' \end{pmatrix},\end{aligned}$$

which implies the rules (27.62).

In similar fashion, one can show that transition from the electron-electron to the electron-hole interaction in the $\mathbf{k}$-representation obeys the rules

$$\mathcal{H}^{h}_{\substack{n'\mathbf{k}_2' \\ n\,\mathbf{k}_2}} = -\,\mathcal{H}^{e}_{\substack{Kn,\ -\mathbf{k}_2' \\ Kn',\ -\mathbf{k}_2'}}, \qquad \mathcal{H}^{eh}_{\substack{m'\mathbf{k}_1'n'\mathbf{k}_2' \\ m\,\mathbf{k}_1 n\,\mathbf{k}_2}} = -\,\mathcal{H}^{ee}_{\substack{m'\mathbf{k}_1'Kn,\ -\mathbf{k}_2' \\ m\,\mathbf{k}_1 Kn',\ -\mathbf{k}_2}}$$

since $K\mathbf{k} = -\mathbf{k}$.