13

Energy Band Models Based on Symmetry

Chapter 12 addressed the general application of space groups to the one-electron energy bands in a periodic solid in the limit of vanishing periodic potential [$V(\boldsymbol{r}) \to 0$]. This chapter deals with a model for which $V(\boldsymbol{r}) \neq 0$ is present and where extensive use is made of crystal symmetry, namely $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory. The Slater–Koster model, which also has a basic symmetry formalism, is discussed in Chap. 15, ater the spin–orbit interaction is considered in Chap. 14.

13.1 Introduction

Just from the symmetry properties of a particular crystal, a good deal can be deduced concerning the form of the energy bands of that crystal. Our study of the group of the wave vector illustrates that some of the basic questions, such as band degeneracy and connectivity, are answered by group theory alone. It is not necessary to solve Schrödinger's equations explicitly to find the *degeneracies* and the *connectivity* relations for $E_n(\boldsymbol{k})$. An interpolation or extrapolation technique for determining energy band dispersion relations based on symmetry often provides the functional form of $E_n(\boldsymbol{k})$ without actual solution of Schrödinger's equation. Such an approach is useful as an interpolation scheme for experimental data or also for band calculations that are carried out with great care at a few high symmetry points in the Brillouin zone.

The interpolation/extrapolation method considered in this chapter is called $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory (extrapolation or a Taylor's series expansion of $E(\boldsymbol{k})$). A related method called the Slater–Koster Fourier expansion [29] (an interpolation or Fourier series expansion of $E(\boldsymbol{k})$) is the basis for symmetry formalism in the tight-binding method, and it will be discussed in Chap. 15, after spin–orbit interaction is considered in Chap. 14. If the available experimental data are limited to one small region in the Brillouin zone

Table 13.1. Irrreducible representations (IRs) of the cubic group O_h

even		odd	
Γ_1^+	Γ_1	Γ_1^-	$\Gamma_{1'}$
Γ_2^+	Γ_2	Γ_2^-	$\Gamma_{2'}$
Γ_{12}^+	Γ_{12}	Γ_{12}^-	$\Gamma_{12'}$
Γ_{15}^+	$\Gamma_{15'}$	Γ_{15}^-	Γ_{15}
Γ_{25}^+	$\Gamma_{25'}$	Γ_{25}^-	Γ_{25}

and that is all that is known and under consideration, then $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory is the appropriate method to use for describing $E(\boldsymbol{k})$. This is often the case in practice for semiconductors. If, however, the available experimental data relate to several points or regions in the Brillouin zone, then the Slater–Koster approach is more appropriate. Although such experiments might seem to yield unrelated information about the energy bands, the Slater–Koster approach is useful for interrelating the results of such experiments.

The particular example used here to illustrate $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory is the electronic structure for a material with simple cubic symmetry. This discussion is readily extended to the electronic structure of semiconductors that crystallize in the diamond structure (e.g., silicon) or the zincblende structure (e.g., GaAs). The valence and conduction bands for these semiconductors are formed from hybridized s- and p-bands.

We first consider cubic electronic energy band structures with inversion symmetry. To emphasize inversion symmetry we will here use the notation $\Gamma_i^{\pm}$ to denote irreducible representations that are even and odd under the inversion operator, when we write the irreducible representations of the cubic O_h group, see Table 13.1 For the nonsymmorphic diamond structure, the s- and p-functions at $k = 0$ in the O_h point group (at $\boldsymbol{k} = 0$) transform as the Γ_1^+ and Γ_{15}^- irreducible representations, respectively (see Sect. 10.8). In the diamond structure there are 2 atoms per unit cell and Γ^{equiv} at $\boldsymbol{k} = 0$ transforms as $\Gamma_1^+ + \Gamma_2^-$ (see Table 10.8). Thus we must consider eight bands in discussing the valence and conduction bands formed by s- and p-bands for the diamond structure. These bands have symmetries

$$\begin{aligned}
&\Gamma^{equiv} \otimes \Gamma_{s\text{-functions}}(\Gamma_1^+ + \Gamma_2^-) \otimes \Gamma_1^+ = \Gamma_1^+ + \Gamma_2^- \\
&\Gamma^{equiv} \otimes \Gamma_{p\text{-functions}}(\Gamma_1^+ + \Gamma_2^-) \otimes \Gamma_{15}^- = \Gamma_{15}^- + \Gamma_{25}^+ \,. \qquad (13.1)
\end{aligned}$$

We identify the Γ_1^+ and Γ_{25}^+ bands as the bonding s- and p-bands and the Γ_2^- and Γ_{15}^- bands as antibonding s- and p-bands. The reason why the bonding p-band has Γ_{25}^+ symmetry follows from the direct product $\Gamma_2^- \otimes \Gamma_{15}^- = \Gamma_{25}^+$ in (13.1). So long as the discussion of $E_n(\boldsymbol{k})$ remains close to $\boldsymbol{k} = 0$, the nonsymmorphic nature of the energy bands is not important and the simple discussion presented here remains valid.

Our discussion starts with a brief review of $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory in general (Sect. 13.2). An example of $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory for a nondegenerate level is then given in Sect. 13.3. This is followed by an example of degenerate first-order perturbation theory and a two-band model (Sect. 13.4) which is then followed by degenerate second-order $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory which is appropriate for describing the p-bonding and antibonding levels in the diamond structure (Sect. 13.5). In all of these cases, group theory tells us which are the nonvanishing matrix elements, which bands couple to one another and which matrix elements are equal to each other. The application of $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory to the electronic energy bands at a Δ point is discussed in Sect. 13.6, and to the valley-orbit interaction in semiconductors is given in Sect. 13.8.

13.2 $\boldsymbol{k}\cdot\boldsymbol{p}$ Perturbation Theory

An electron in a periodic potential obeys the one-electron Hamiltonian:

$$\left[\frac{p^2}{2m}+V(\boldsymbol{r})\right]\psi_{n,\boldsymbol{k}}(\boldsymbol{r})=E_n(\boldsymbol{k})\psi_{n,\boldsymbol{k}}(\boldsymbol{r})\,, \tag{13.2}$$

where the eigenfunctions of the Hamiltonian are the Bloch functions

$$\psi_{n,\boldsymbol{k}}(\boldsymbol{r})=\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}u_{n,\boldsymbol{k}}(\boldsymbol{r}) \tag{13.3}$$

and n is the band index. Substitution of $\psi_{n,\boldsymbol{k}}(\boldsymbol{r})$ into Schrödinger's equation gives an equation for the periodic function $u_{n,\boldsymbol{k}}(\boldsymbol{r})$

$$\left[\frac{p^2}{2m}+V(\boldsymbol{r})+\frac{\hbar\boldsymbol{k}\cdot\boldsymbol{p}}{m}+\frac{\hbar^2k^2}{2m}\right]u_{n,\boldsymbol{k}}(\boldsymbol{r})=E_n(\boldsymbol{k})\;u_{n,\boldsymbol{k}}(\boldsymbol{r})\,. \tag{13.4}$$

In the spirit of the $(\boldsymbol{k}\cdot\boldsymbol{p})$ method, we assume that $E_n(\boldsymbol{k})$ is known at point $\boldsymbol{k}=\boldsymbol{k}_0$ either from experimental information or from direct solution of Schrödinger's equation for some model potential $V(\boldsymbol{r})$. Assume the band in question has symmetry Γ_i so that the function $u_{n,\boldsymbol{k}_0}(\boldsymbol{r})$ transforms as the irreducible representation Γ_i. Then we have

$$\mathcal{H}_{\boldsymbol{k}_0}u^{(\Gamma_i)}_{n,\boldsymbol{k}_0}=\varepsilon_n(\boldsymbol{k}_0)\;u^{(\Gamma_i)}_{n,\boldsymbol{k}_0}\,, \tag{13.5}$$

where

$$\mathcal{H}_{\boldsymbol{k}_0}=\frac{p^2}{2m}+V(\boldsymbol{r})+\frac{\hbar\boldsymbol{k}_0\cdot\boldsymbol{p}}{m} \tag{13.6}$$

and

$$\varepsilon_n(\boldsymbol{k}_0)=E_n(\boldsymbol{k}_0)-\frac{\hbar^2\boldsymbol{k}_0^2}{2m}\,. \tag{13.7}$$

If $\varepsilon_n(\boldsymbol{k}_0)$ and $u_{n,\boldsymbol{k}_0}(\boldsymbol{r})$ are specified at $\boldsymbol{k}_0$, the $\boldsymbol{k}\cdot\boldsymbol{p}$ method prescribes the development of the periodic $u_{n,\boldsymbol{k}_0}(\boldsymbol{r})$ functions under variation of $\boldsymbol{k}$. At point $\boldsymbol{k}=\boldsymbol{k}_0+\boldsymbol{\kappa}$, the eigenvalue problem becomes

$$\begin{aligned}\mathcal{H}_{\boldsymbol{k}_0+\boldsymbol{\kappa}}u_{n,\boldsymbol{k}_0+\boldsymbol{\kappa}}(\boldsymbol{r}) &= \left(\mathcal{H}_{\boldsymbol{k}_0}+\frac{\hbar\boldsymbol{\kappa}\cdot\boldsymbol{p}}{m}\right)u_{n,\boldsymbol{k}_0+\boldsymbol{\kappa}}(\boldsymbol{r})\\ &= \varepsilon_n\left(\boldsymbol{k}_0+\boldsymbol{\kappa}\right)u_{n,\boldsymbol{k}_0+\boldsymbol{\kappa}}(\boldsymbol{r})\,. \end{aligned}\tag{13.8}$$

In the spirit of the usual $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory, $\boldsymbol{\kappa}$ is small so that the perturbation Hamiltonian is taken as $\mathcal{H}'=\hbar\boldsymbol{\kappa}\cdot\boldsymbol{p}/m$ and the energy eigenvalue at the displaced $\boldsymbol{k}$ vector $\varepsilon_n(\boldsymbol{k}_0+\boldsymbol{\kappa})$ is given by (13.7), and $E_n(\boldsymbol{k}_0)$ is given by (13.2). We will illustrate this method first for a nondegenerate band (a $\Gamma_1^{\pm}$ band for the simple cubic lattice) and then in Sect. 13.5 for a degenerate band (a $\Gamma_{15}^{\pm}$ band for the simple cubic lattice).

13.3 Example of $\boldsymbol{k}\cdot\boldsymbol{p}$ Perturbation Theory for a Nondegenerate $\boldsymbol{\Gamma_1^+}$ Band

Suppose the energy of the $\Gamma_1^{\pm}$ band at $\boldsymbol{k}=0$ in a crystal with O_h point symmetry is established by the identification of an optical transition and measurement of its resonant photon energy. The unperturbed wave function at $\boldsymbol{k}=0$ is $u_{n,0}^{\Gamma_1^+}(\boldsymbol{r})$ and its eigenvalue from (13.7) is $\varepsilon_n^{(\Gamma_1^+)}(0)=E_n^{(\Gamma_1^+)}(0)$ since $\boldsymbol{k}_0=0$. Away from $\boldsymbol{k}_0=0$, we use $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory [31, 45]:

$$\begin{aligned}\varepsilon_n^{(\Gamma_1^+)}(\boldsymbol{\kappa}) &= E_n^{(\Gamma_1^+)}(0)+\left(u_{n,0}^{\Gamma_1^+}|\mathcal{H}'|u_{n,0}^{\Gamma_1^+}\right)\\ &\quad+\sum_{n'\neq n}\frac{\left(u_{n,0}^{\Gamma_1^+}|\mathcal{H}'|u_{n',0}^{\Gamma_i}\right)\left(u_{n',0}^{\Gamma_i}|\mathcal{H}'|u_{n,0}^{\Gamma_1^+}\right)}{E_n^{\Gamma_1^+}(0)-E_{n'}^{\Gamma_i}(0)}\,,\end{aligned}\tag{13.9}$$

where the sum is over states n' which have symmetries Γ_i.

Now $\mathcal{H}'=\hbar\boldsymbol{\kappa}\cdot\boldsymbol{p}/m$ transforms like a vector, since $\mathcal{H}'$ is proportional to the vector $\boldsymbol{p}$, which pertains to the electronic system and $\boldsymbol{\kappa}$ is considered as an external variable not connected to the electronic system. If we expand the eigenfunctions and eigenvalues of (13.9) about the Γ point ($\boldsymbol{k}=0$), then $\mathcal{H}'$ which transforms according to the vector, will transform as the irreducible representation Γ_{15}^- in O_h symmetry. In the spirit of $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory, the vector $\boldsymbol{k}_0$ determines the point symmetry group that is used to classify the wave functions and eigenvalues for $\mathcal{H}'$.

For the $\boldsymbol{k}\cdot\boldsymbol{p}$ expansion about the Γ point, the linear term in k which arises in first order perturbation theory *vanishes* when $\boldsymbol{k}_0=0$ since $(u_{n,0}^{\Gamma_1^+}|\mathcal{H}'|u_{n,0}^{\Gamma_1^+})$ transforms according to the direct product $\Gamma_1^+\otimes\Gamma_{15}^-\otimes\Gamma_1^+=\Gamma_{15}^-$ which does not contain Γ_1^+ (see Sect. 6.7). The same result is obtained using arguments

relevant to the oddness and evenness of the functions which enter the matrix elements of (13.9). At other k points in the Brillouin zone, the $\boldsymbol{k}\cdot\boldsymbol{p}$ expansion may contain linear k terms since the group of the wave vector for that $\boldsymbol{k}\cdot\boldsymbol{p}$ expansion point may not contain the inversion operation.

Now let us look at the terms

$$\left(u_{n',0}^{\Gamma_i}|\mathcal{H}'|u_{n,0}^{\Gamma_1^+}\right)$$

that arise in second order perturbation theory. The product $\mathcal{H}'u_{n,0}^{\Gamma_1^+}$ transforms as $\Gamma_{15}^-\otimes\Gamma_1^+=\Gamma_{15}^-$ so that Γ_i must be of Γ_{15}^- symmetry if a nonvanishing matrix element is to result. We thus obtain

$$\varepsilon_n^{\Gamma_1^+}(\boldsymbol{\kappa})=E_n^{\Gamma_1^+}(0)+\Sigma_{n'\neq n\ (\Gamma_{15}^-)}\frac{\left(u_{n,0}^{\Gamma_1^+}|\mathcal{H}'|u_{n',0}^{\Gamma_{15}^-}\right)\left(u_{n',0}^{\Gamma_{15}^-}|\mathcal{H}'|u_{n,0}^{\Gamma_1^+}\right)}{E_n^{\Gamma_1^+}(0)-E_n^{\Gamma_{15}^-}(0)}+\cdots \quad (13.10)$$

and a corresponding relation is obtained for the nondegenerate Γ_1^- and $\Gamma_2^\pm$ levels. For a semiconductor that crystallizes in the diamond structure, the symmetry Γ_1^+ describes the valence band s-band bonding state, while symmetry Γ_2^- describes the conduction band s-band antibonding state (see Problem 13.1).

Thus we see that by using group theory, our $\boldsymbol{k}\cdot\boldsymbol{p}$ expansion is greatly simplified, since it is only the Γ_{15}^- levels that couple to the Γ_1^+ level by $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory in (13.10). These statements are completely independent of the explicit wave functions which enter the problem, but depend only on their *symmetry*. Further simplifications result from the observation that for cubic symmetry the matrix elements

$$\left(u_{n,0}^{\Gamma_1^+}|\mathcal{H}'|u_{n',0}^{\Gamma_{15}^-}\right)$$

can all be expressed in terms of a single matrix element, if $u_{n',0}^{\Gamma_{15}^-}$ is identified with specific basis functions, such as p-functions (denoted by x, y, z for brevity) and $u_{n,0}^{\Gamma_1^+}$ with an s-function (denoted by 1 for brevity). Thus for the O_h group, the selection rules (see Sect. 6.6) give

$$(1|p_x|x)=(1|p_y|y)=(1|p_z|z)\,, \quad (13.11)$$

and all other cross terms of the form $(1|p_x|y)$ vanish. This result, that the matrix elements of $\boldsymbol{p}$ in O_h symmetry have only one independent matrix element, also follows from the theory of permutation groups (see Chap. 17). Combining these results with

$$\varepsilon_n^{\Gamma_1^+}(\boldsymbol{\kappa})=E_n^{\Gamma_1^+}(\boldsymbol{\kappa})-\hbar^2\kappa^2/2m$$

we get

$$E_n^{\Gamma_1^+}(\boldsymbol{\kappa}) = E_n^{\Gamma_1^+}(0) + \frac{\hbar^2\kappa^2}{2m} + \frac{\hbar^2\kappa^2}{m^2}\sum_{n'\neq n}\frac{|(1|p_x|x)|^2}{E_n^{\Gamma_1^+}(0) - E_{n'}^{\Gamma_{15}^-}(0)}\,, \tag{13.12}$$

where the sum is over all states n' with Γ_{15}^- symmetry. A similar expansion formula is applicable to

$$E_n^{\Gamma_2^-}(\boldsymbol{k})\,,$$

which corresponds to the conduction antibonding s-band in the diamond structure. Equation (13.12) is sometimes written in the form

$$E_n^{\Gamma_1^+}(\boldsymbol{\kappa}) = E_n^{\Gamma_1^+}(0) + \frac{\hbar^2\kappa^2}{2m_n^*}\,, \tag{13.13}$$

where the effective mass parameter m_n^* is related to band couplings through the momentum matrix element:

$$\frac{m}{m_n^*} = 1 + \frac{2}{m}\sum_{n'\neq n}\frac{|(1|p_x|x)|^2}{E_n^{\Gamma_1^+}(0) - E_{n'}^{\Gamma_{15}^-}(0)}\,, \tag{13.14}$$

in which the sum over n' is restricted to states with Γ_{15}^- symmetry. Consistent with (13.12), the effective mass m_n^* is related to the band curvature by the relation

$$\frac{\partial^2 E_n^{\Gamma_1^+}(\kappa)}{\partial\kappa^2} = \frac{\hbar^2}{m_n^*}\,. \tag{13.15}$$

Thus m_n^* is proportional to the inverse of the band curvature. If the curvature is large, the effective mass is small and conversely, and if the bands are "flat" (essentially k-independent), the effective masses are large. Thus the $\boldsymbol{k}\cdot\boldsymbol{p}$ expansion for a nondegenerate band in a cubic crystal leads to an isotropic parabolic dependence of $E_n(\boldsymbol{k})$ on $\boldsymbol{k}$ which looks just like the free electron dispersion relation except that the free electron mass m is replaced by m^* which reflects the effect of the crystalline potential on the motion of the electron.

For the case that the nondegenerate level with Γ_1^+ symmetry is predominantly coupled to a single degenerate band (such as one degenerate band with Γ_{15}^- symmetry which in this case relates to the p bonding state in the conduction band), the effective mass formula (13.14) becomes

$$\frac{m}{m_n^*} = 1 + \frac{2}{m}\frac{|(1|p_x|x)|^2}{\varepsilon_g}\,, \tag{13.16}$$

which is useful for estimating effective masses, provided that we know the magnitude of the matrix element and the band gap ε_{g}. On the other hand,

if m^* and ε_g are known experimentally, then (13.16) is useful for evaluating $|(1|p_x|x)|^2$. This is, in fact, the most common use of (13.16). The words *matrix element* or *oscillator strength* typically refer to the momentum matrix element $(u_{n,\boldsymbol{k}}|p_x|u_{n',\boldsymbol{k}})$ when discussing the optical properties of solids.

The treatment given here for the nondegenerate bands is easily carried over to treating the $\boldsymbol{k}\cdot\boldsymbol{p}$ expansion about some other high symmetry point in the Brillouin zone for symmorphic structures. For arbitrary points in the Brillouin zone, the diagonal term arising from first order perturbation theory does not vanish. Also the matrix element

$$\left(u^{\Gamma_i^{\pm}}_{n,\boldsymbol{k}_0}|p_\alpha|u^{\Gamma_j^{\mp}}_{n,\boldsymbol{k}_0}\right)$$

need not be the same for each component $\alpha = x, y, z$, and for the most general case, six independent matrix elements would be expected. For example, along the Δ and Λ axes, the matrix element for momentum $\parallel$ to the high symmetry axis is not equal to the components $\perp$ to the axis, and there are two independent matrix elements along each of the Δ and Λ axes (see Sect. 13.6).

These two directions are called *longitudinal* ($\parallel$ to the axis) and *transverse* ($\perp$ to the axis), and lead to longitudinal and transverse effective mass components away from the Γ point. Furthermore, for the case of nonsymmorphic structures like the diamond structure, the nonsymmorphic symmetry elements involving translations must be considered in detail away from $k = 0$.

13.4 Two Band Model: Degenerate First-Order Perturbation Theory

One of the simplest applications of $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory is to two-band models for crystalline solids. These models are applicable to describe the energy dispersion $E(\boldsymbol{k})$ about a point $\boldsymbol{k}_0$ for one of two bands that are strongly coupled to each other and are weakly coupled to all other bands. The strongly coupled set is called the nearly degenerate set (NDS) and, if need be, the weakly coupled bands can always be treated in perturbation theory after the problem of the strongly interacting bands is solved. Simple extensions of the two-band model are made to handle three strongly coupled bands, such as the valence band of silicon, germanium and related semiconductors, or even to handle four strongly coupled bands as occur in graphite. We illustrate the procedure here for symmorphic systems, but for application to nonsymmorphic groups, care with handling phase factors becomes important (see Sect. 12.5).

The eigenvalue problem to be solved is

$$\left[\frac{p^2}{2m} + V(\boldsymbol{r}) + \frac{\hbar\boldsymbol{k}_0\cdot\boldsymbol{p}}{m} + \frac{\hbar\boldsymbol{\kappa}\cdot\boldsymbol{p}}{m}\right] u_{n,\boldsymbol{k}_0+\boldsymbol{\kappa}}(\boldsymbol{r}) = \varepsilon_n(\boldsymbol{k}_0+\boldsymbol{\kappa})u_{n,\boldsymbol{k}_0+\boldsymbol{\kappa}}(\boldsymbol{r})\,, \tag{13.17}$$

in which $\varepsilon_n(\boldsymbol{k}_0)$ is related to the solution of Schrödinger's equation $E_n(\boldsymbol{k}_0)$ by (13.7).

Let $n = i, j$ be the two bands that are nearly degenerate. Using first-order degenerate perturbation theory, the secular equation is written as

$$\begin{array}{c} \\ i \\ j \end{array} \begin{array}{c} \begin{array}{cc} i & j \end{array} \\ \begin{vmatrix} \langle i|\mathcal{H}_0 + \mathcal{H}'|i\rangle - \varepsilon & \langle i|\mathcal{H}_0 + \mathcal{H}'|j\rangle \\ \langle j|\mathcal{H}_0 + \mathcal{H}'|i\rangle & \langle j|\mathcal{H}_0 + \mathcal{H}'|j\rangle - \varepsilon \end{vmatrix} \end{array} = 0\,, \tag{13.18}$$

in which we have explicitly written i and j to label the rows and columns.

Equation (13.18) is exact within the two-band model, i.e., all the coupling occurs between the nearly degenerate set and no coupling is made to bands outside this set. For many cases where the two-band model is applied (e.g., PbTe), the unperturbed wave functions $u_{n,\boldsymbol{k}_0}(\boldsymbol{r})$ are invariant under inversion. Then because of the oddness of $\mathcal{H}' = \hbar\boldsymbol{\kappa}\cdot\boldsymbol{p}/m$, the matrix elements vanish

$$\langle i|\mathcal{H}'|i\rangle = \langle j|\mathcal{H}'|j\rangle = 0\,. \tag{13.19}$$

Also since the "band edge" wave functions $u_{n,\boldsymbol{k}_0}(\boldsymbol{r})$ are constructed to diagonalize the Hamiltonian

$$\mathcal{H}_0 u_{n,\boldsymbol{k}_0}(\boldsymbol{r}) = \varepsilon_n(\boldsymbol{k}_0) u_{n,\boldsymbol{k}_0}(\boldsymbol{r})\,, \tag{13.20}$$

there are no off-diagonal matrix elements of $\mathcal{H}_0$ or

$$\langle i|\mathcal{H}_0|j\rangle = 0\,, \quad \text{for}\, i \neq j\,. \tag{13.21}$$

We then write

$$\langle i|\mathcal{H}_0|i\rangle = E_i^0 \quad \text{and} \quad \langle j|\mathcal{H}_0|j\rangle = E_j^0\,, \tag{13.22}$$

where for $n = i, j$

$$E_n^0 = E_n(\boldsymbol{k}_0) - \frac{\hbar^2 k_0^2}{2m}\,. \tag{13.23}$$

In this notation the secular equation can be written as

$$\begin{vmatrix} E_i^0 - \varepsilon & (\hbar/m)\boldsymbol{\kappa}\cdot\langle i|\boldsymbol{p}|j\rangle \\ (\hbar/m)\boldsymbol{\kappa}\cdot\langle j|\boldsymbol{p}|i\rangle & E_j^0 - \varepsilon \end{vmatrix} = 0\,, \tag{13.24}$$

where $\langle i|\boldsymbol{p}|j\rangle \neq 0$ for the two-band model. The secular equation implied by (13.24) is equivalent to the quadratic equation

$$\varepsilon^2 - \varepsilon\left[E_i^0 + E_j^0\right] + E_i^0 E_j^0 - \frac{\hbar^2}{m^2}\boldsymbol{\kappa}\cdot\langle i|\boldsymbol{p}|j\rangle\langle j|\boldsymbol{p}|i\rangle\cdot\boldsymbol{\kappa} = 0\,. \tag{13.25}$$

We write the symmetric tensor $\overleftrightarrow{p^2_{ij}}$ coupling the two bands as

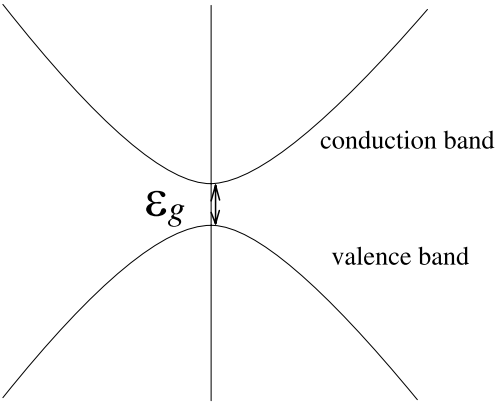

Fig. 13.1. Two strongly coupled mirror bands separated by an energy $\varepsilon_{\rm g}$ at the band extremum. This sketch is based on the concept that these two bands would be degenerate at the center of the band gap but a strong interaction splits this degeneracy at a high symmetry point and creates a band gap $\varepsilon_{\rm g}$

$$\overleftrightarrow{p^2_{ij}} = \langle i|\boldsymbol{p}|j\rangle\langle j|\boldsymbol{p}|i\rangle\,, \tag{13.26}$$

where i and j in the matrix elements refer to the band edge wave functions $u_{n,\boldsymbol{k}_0}(\boldsymbol{r})$ and $n = i, j$. The solution to the quadratic equation (13.25) yields

$$\varepsilon(\boldsymbol{\kappa}) = \frac{E_i^0 + E_j^0}{2} \pm \frac{1}{2}\sqrt{(E_i^0 - E_j^0)^2 + \frac{4\hbar^2}{m^2}\boldsymbol{\kappa}\cdot\overleftrightarrow{p^2_{ij}}\cdot\boldsymbol{\kappa}}\,. \tag{13.27}$$

We choose our zero of energy symmetrically such that

$$E_i^0 = \varepsilon_{\rm g}/2\,, \quad E_j^0 = -\varepsilon_{\rm g}/2 \tag{13.28}$$

to obtain the two-band model result (see Fig. 13.1):

$$\varepsilon(\boldsymbol{\kappa}) = \pm\frac{1}{2}\sqrt{\varepsilon_{\rm g}^2 + \frac{4\hbar^2}{m^2}\boldsymbol{\kappa}\cdot\overleftrightarrow{p^2_{ij}}\cdot\boldsymbol{\kappa}}\,, \tag{13.29}$$

which at $\boldsymbol{\kappa} = 0$ reduces properly to $\varepsilon(0) = \pm 1/2\varepsilon_{\rm g}$.

Equation (13.29) gives a *nonparabolic dependence of E upon* $\boldsymbol{\kappa}$. For strongly coupled bands, the two-band model is characterized by its nonparabolicity. In the approximation that there is no coupling to bands outside the nondegenerate set, these bands are strictly *mirror bands*, whereby one band is described by an $E(\boldsymbol{\kappa})$ relation given by the + sign; the other by the identical relation with the − sign. For cubic materials there is only one independent matrix element

$$\overleftrightarrow{p^2_{ij}} = \langle i|p_\alpha|j\rangle\langle j|p_\alpha|i\rangle \equiv p_{ij}^2\,, \qquad \alpha = x, y, z\,, \tag{13.30}$$

and the $\overleftrightarrow{p^2_{ij}}$ tensor assumes the form

$$\overleftrightarrow{p^2_{ij}} = \begin{pmatrix} p^2_{ij} & 0 & 0 \\ 0 & p^2_{ij} & 0 \\ 0 & 0 & p^2_{ij} \end{pmatrix} . \tag{13.31}$$

In applying the two-band model to cubic symmetry, the degeneracy of the Γ^+_{25} valence bands or the Γ^-_{15} conduction bands is often ignored. The two-band model formula then becomes

$$\varepsilon(\boldsymbol{\kappa}) = \pm\frac{1}{2}\sqrt{\varepsilon_{\mathrm{g}}^2 + \frac{4\hbar^2\kappa^2 p^2_{ij}}{m^2}}\,, \quad \text{where} \quad \kappa^2 = \kappa_x^2 + \kappa_y^2 + \kappa_z^2\,. \tag{13.32}$$

In this form, (13.32) is called the Kane two-band model. The generalization of (13.32) to noncubic materials is usually called the Lax two-band model, and in the case of bismuth the $\overleftrightarrow{p^2_{ij}}$ tensor has the following form

$$\overleftrightarrow{p^2_{ij}} = \begin{pmatrix} p^2_{xx} & 0 & 0 \\ 0 & p^2_{yy} & p^2_{yz} \\ 0 & p^2_{yz} & p^2_{zz} \end{pmatrix} , \tag{13.33}$$

where the x axis is a binary axis $\perp$ to the mirror plane in bismuth (space group $R\bar{3}m$, #166), and the matrix elements of (13.33) have four independent components.

We now show that for small κ we recover the parabolic $\varepsilon(\boldsymbol{\kappa})$ relations. For example, for the Kane two-band model, a Taylor's series expansion of (13.32) yields

$$\varepsilon(\boldsymbol{\kappa}) = \pm\frac{1}{2}\sqrt{\varepsilon_{\mathrm{g}}^2 + \frac{4\hbar^2\kappa^2 p^2_{ij}}{m^2}} = \pm\frac{\varepsilon_{\mathrm{g}}}{2}\left[1 + \frac{4\hbar^2\kappa^2 p^2_{ij}}{\varepsilon_{\mathrm{g}}^2 m^2}\right]^{1/2} , \tag{13.34}$$

which to order κ^4 becomes

$$\varepsilon(\boldsymbol{\kappa}) = \pm\left[\frac{\varepsilon_{\mathrm{g}}}{2} + \frac{\hbar^2\kappa^2 p^2_{ij}}{\varepsilon_{\mathrm{g}} m^2} - \frac{\hbar^4\kappa^4 p^4_{ij}}{\varepsilon_{\mathrm{g}}^3 m^4} + \cdots\right] , \tag{13.35}$$

where $\varepsilon(\boldsymbol{\kappa})$ is given by (13.7), the momentum matrix elements, which reflect group theoretical considerations, are given by

$$p^2_{ij} = |(1|p_x|x)|^2\,, \tag{13.36}$$

and the bandgap at the band extrema is given by $E_n(\boldsymbol{k}_0) - E_{n'}(\boldsymbol{k}_0) = \pm\varepsilon_{\mathrm{g}}$.

If the power series expansion in (13.35) is rapidly convergent (either because κ is small or the bands are not that strongly coupled – i.e., p^2_{ij} is not

too large), then the expansion through terms in κ^4 is useful. We note that, within the two-band model, the square root formula of (13.34) is exact and is the one that is not restricted to small κ or small p_{ij}^2. It is valid so long as the two-band model itself is valid.

Some interesting consequences arise from these nonparabolic features of the dispersion relations. For example, the effective mass (or band curvature) is energy or κ-dependent. Consider the expression which follows from (13.35):

$$E_n(\boldsymbol{k}_0+\boldsymbol{\kappa}) \simeq \frac{\hbar^2|\boldsymbol{k}_0+\boldsymbol{\kappa}|^2}{2m} \pm \left[\frac{\varepsilon_{\mathrm{g}}}{2} + \frac{\hbar^2\kappa^2 p_{ij}^2}{\varepsilon_{\mathrm{g}} m^2} - \frac{\hbar^4\kappa^4 p_{ij}^4}{\varepsilon_{\mathrm{g}}^3 m^4}\right]. \tag{13.37}$$

Take $\boldsymbol{k}_0 = 0$, so that

$$\frac{\partial^2 E}{\partial \kappa^2} = \frac{\hbar^2}{m} \pm \left[\frac{2\hbar^2 p_{ij}^2}{\varepsilon_{\mathrm{g}} m^2} - \frac{12\kappa^2\hbar^4 p_{ij}^4}{\varepsilon_{\mathrm{g}}^3 m^4}\right] \equiv \frac{\hbar^2}{m^*}. \tag{13.38}$$

From this equation we see that the curvature $\partial^2 E/\partial\kappa^2$ is κ dependent. In fact as we more further from the band extrema, the band curvature decreases, the bands become more flat and the effective mass increases. This result is also seen from the definition of m^* (13.38)

$$\frac{m}{m^*} = 1 \pm \left[\frac{2}{m}\frac{p_{ij}^2}{\varepsilon_{\mathrm{g}}} - \frac{12\hbar^2\kappa^2 p_{ij}^4}{\varepsilon_{\mathrm{g}}^3 m^3}\right]. \tag{13.39}$$

Another way to see that the masses become heavier as we move higher into the band (away from $\boldsymbol{k}_0$) is to work with the square root formula (13.34):

$$\varepsilon = \pm\frac{1}{2}\sqrt{\varepsilon_{\mathrm{g}}^2 + \frac{4\hbar^2\kappa^2 p_{ij}^2}{m^2}}. \tag{13.40}$$

Squaring (13.40) and rewriting this equation, we obtain

$$(2\varepsilon - \varepsilon_{\mathrm{g}})(2\varepsilon + \varepsilon_{\mathrm{g}}) = \frac{4\hbar^2\kappa^2 p_{ij}^2}{m^2}, \tag{13.41}$$

$$(2\varepsilon - \varepsilon_{\mathrm{g}}) = \frac{4\hbar^2\kappa^2 p_{ij}^2}{m^2(2\varepsilon + \varepsilon_{\mathrm{g}})}. \tag{13.42}$$

For $\kappa = 0$ we have $\varepsilon = \varepsilon_{\mathrm{g}}/2$, and we then write an expression for $\varepsilon(\kappa)$:

$$\varepsilon(\kappa) = \frac{\varepsilon_{\mathrm{g}}}{2} + \frac{2\hbar^2\kappa^2 p_{ij}^2}{m^2(2\varepsilon + \varepsilon_{\mathrm{g}})} = \frac{\varepsilon_{\mathrm{g}}}{2} + \frac{\hbar^2\kappa^2 p_{ij}^2}{m^2(\varepsilon + \frac{\varepsilon_{\mathrm{g}}}{2})}. \tag{13.43}$$

Therefore we obtain the nonparabolic two-band model relation

$$E(\boldsymbol{\kappa}) = \frac{\varepsilon_{\mathrm{g}}}{2} + \frac{\hbar^2\kappa^2}{2m}\left[1 + \frac{2p_{ij}^2}{m(\varepsilon + \frac{\varepsilon_{\mathrm{g}}}{2})}\right], \tag{13.44}$$

which is to be compared with the result for simple nondegenerate bands (13.12):

$$E_i(\boldsymbol{\kappa}) = E_i(0) + \frac{\hbar^2\kappa^2}{2m}\left[1 + \frac{2p_{ij}^2}{m\varepsilon_{\mathrm{g}}}\right]. \tag{13.45}$$

Equation (13.44) shows that for the nonparabolic two-band model, the effective mass at the band edge is given by

$$\frac{m}{m^*} = \left[1 + \frac{2p_{ij}^2}{m\varepsilon_{\mathrm{g}}}\right], \tag{13.46}$$

but the effective mass becomes heavier as we move away from $\boldsymbol{k}_0$ and as we move up into the band. The magnitude of the k or energy dependence of the effective mass is very important in narrow gap materials such as bismuth. At the band edge, the effective mass parameter for electrons in Bi is $\sim 0.001m_0$ whereas at the Fermi level $m^* \sim 0.008m_0$. The number of electron carriers in Bi is only $10^{17}\,\mathrm{cm}^{-3}$. Since the density of states for simple bands in a 3D crystal has a dependence $\sim m^{*3/2}E^{1/2}$, we can expect a large increase in the density of states with increasing energy in a nonparabolic band with a small effective mass at the band edge. Since bismuth has relatively low symmetry, the tensorial nature of the effective mass tensor must be considered and the dispersion relations for the coupled bands at the L point in bismuth are generally written as

$$\varepsilon(\boldsymbol{\kappa}) = \pm\frac{1}{2}\sqrt{\varepsilon_{\mathrm{g}}^2 + 2\hbar^2\varepsilon_{\mathrm{g}}\frac{\boldsymbol{\kappa}\cdot\overleftrightarrow{\alpha}\cdot\boldsymbol{\kappa}}{m}}, \tag{13.47}$$

in which $\overleftrightarrow{\alpha}$ is a reciprocal effective mass tensor.

13.5 Degenerate second-order $\boldsymbol{k}\cdot\boldsymbol{p}$ Perturbation Theory

For many cubic crystals it is common to have triply degenerate energy bands arising from degenerate p states, with extrema at $\boldsymbol{k} = 0$. Such bands are of great importance in the transport properties of semiconductors such as silicon, germanium, and III–V compounds. The analysis of experiments such as cyclotron resonance in the valence band of semiconductors depends upon degenerate second-order $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory which is discussed in this section.

Second-order degenerate $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory becomes much more complicated than the simpler applications of perturbation theory discussed in Sect. 13.2–13.4. Group theory thus provides a valuable tool for the solution of practical problems. For example, we consider here how the degeneracy is lifted as we move away from $\boldsymbol{k}=0$ for a Γ_{15}^{-} level for a crystal with O_h symmetry; a similar analysis applies for the Γ_{25}^{+} level, which pertains to the degenerate p-band bonding states in the valence band in the diamond structure.

Suppose that we set up the secular equation for a Γ_{15}^{-} level using degenerate perturbation theory

$$\begin{array}{c} \\ x \\ y \\ z \end{array}\!\!\begin{array}{c} \begin{array}{ccc} x & y & z \end{array} \\ \left|\begin{array}{ccc} (x|\mathcal{H}'|x)-\varepsilon & (x|\mathcal{H}'|y) & (x|\mathcal{H}'|z) \\ (y|\mathcal{H}'|x) & (y|\mathcal{H}'|y)-\varepsilon & (y|\mathcal{H}'|z) \\ (z|\mathcal{H}'|x) & (z|\mathcal{H}'|y) & (z|\mathcal{H}'|z)-\varepsilon \end{array}\right| \end{array} = 0\,, \tag{13.48}$$

where the x, y and z symbols denote the (x, y, z) partners of the basis functions in the Γ_{15}^{-} irreducible representation derived from atomic p-functions and the diagonal matrix elements for $\mathcal{H}'_0$ are set equal to zero at the band extremum, such as the top of the valence band. We notice that since $\mathcal{H}' = \hbar\boldsymbol{k}\cdot\boldsymbol{p}/m$, then $\mathcal{H}'$ transforms like the Γ_{15}^{-} irreducible representation. Therefore we get $(\Gamma_{15}^{-}|\mathcal{H}'|\Gamma_{15}^{-}) = 0$, since

$$\Gamma_{15}^{-}\otimes\Gamma_{15}^{-} = \Gamma_1^{+} + \Gamma_{12}^{+} + \Gamma_{15}^{+} + \Gamma_{25}^{+}\,, \tag{13.49}$$

or more simply, since $\mathcal{H}'$ is odd under inversion, each matrix element in (13.48) vanishes because of parity considerations. Since each of the matrix elements of (13.48) vanishes, the degeneracy of the Γ_{15}^{-} level is not lifted in first-order degenerate perturbation theory; thus we must use *second-order* degenerate perturbation theory to lift this level degeneracy. We show below the derivation of the form of the matrix elements for the off-diagonal matrix elements in (13.48) showing that the vanishing $\mathcal{H}'_{mn}$ is replaced by

$$\mathcal{H}'_{mn} \to \mathcal{H}'_{mn} + \sum_{\alpha}\frac{\mathcal{H}'_{m\alpha}\mathcal{H}'_{\alpha n}}{E_m^{(0)} - E_n^{(0)}}\,. \tag{13.50}$$

We will see below that the states with symmetries given in (13.49) will serve as the intermediate states α which arise in second-order perturbation theory. In applying second-order degenerate perturbation theory, we assume that we have a degenerate (or nearly degenerate) set of levels – abbreviated NDS. We assume that the states inside the NDS are strongly coupled and those outside the NDS are only weakly coupled to states within the NDS (see Fig. 13.2).

The wave function for a state is now written in terms of the unperturbed wave functions and the distinction is made as to whether we are dealing with

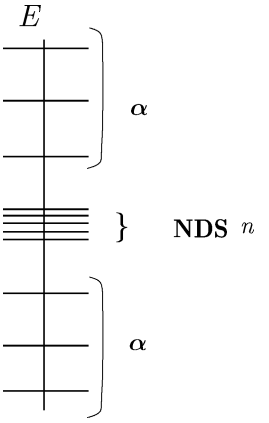

Fig. 13.2. NDS ≡ nearly degenerate set. We use Roman letter subscripts for levels within the NDS (such as n) and Greek indices for levels outside the NDS (such as α)

a state inside or outside of the NDS. If we now expand the wavefunction $\psi_{n'}$ in terms of the unperturbed band edge states, we obtain

$$\psi_{n'} = \sum_n a_n \psi_n^{(0)} + \sum_\alpha a_\alpha \psi_\alpha^{(0)} \,, \tag{13.51}$$

where $\psi_n^{(0)}$ and $\psi_\alpha^{(0)}$ are, respectively, the unperturbed wavefunctions inside (n) and outside (α) of the nearly degenerate set. Substitution into Schrödinger's equation yields

$$\mathcal{H}\psi_{n'} = E\psi_{n'} = \sum_n a_n (E_n^0 + \mathcal{H}')\psi_n^{(0)} + \sum_\alpha a_\alpha (E_\alpha^{(0)} + \mathcal{H}')\psi_\alpha^{(0)} \,. \tag{13.52}$$

We multiply the left-hand side of (13.52) by $\psi_{m_0}^{(0)*}$ and integrate over all space, making use of the orthogonality theorem $\int \psi_m^{(0)*} \psi_n^{(0)} \mathrm{d}\boldsymbol{r} = \delta_{mn}$ to obtain the iterative relation between the expansion coefficients (Brillouin–Wigner Perturbation Theory)

$$[E - E_m^{(0)}]a_m = a_m \mathcal{H}'_{mm} + \sum_{n' \neq m} a_{n'} \mathcal{H}'_{mn'} + \sum_\alpha a_\alpha \mathcal{H}'_{m\alpha} \,, \tag{13.53}$$

where the sum over n' denotes coupling to states in the NDS and the sum over α denotes coupling to states outside the NDS (see Fig. 13.2). A similar procedure also leads to a similar equation for levels outside the NDS:

$$[E - E_\alpha^{(0)}]a_\alpha = a_\alpha \mathcal{H}'_{\alpha\alpha} + \sum_n a_n \mathcal{H}'_{\alpha n} + \sum_{\beta \neq \alpha} a_\beta \mathcal{H}'_{\alpha\beta} \,. \tag{13.54}$$

We now substitute (13.54) for the coefficients a_α outside the NDS in (13.53) to obtain

$$[E - E_m^{(0)}]a_m = a_m\mathcal{H}'_{mm} + \sum_{n'\neq m} a_{n'}\mathcal{H}'_{mn'} \tag{13.55}$$

$$+ \sum_\alpha \frac{\mathcal{H}'_{m\alpha}}{E - E_\alpha^{(0)}} \left\{ \sum_n a_n\mathcal{H}'_{\alpha n} + a_\alpha\mathcal{H}'_{\alpha\alpha} + \sum_\beta a_\beta\mathcal{H}'_{\alpha\beta} \right\}.$$

If we neglect terms in (13.56) which couple states outside the NDS to other states outside the NDS, we obtain

$$a_m(E_m^{(0)} - E) + \sum_n a_n\mathcal{H}'_{mn} + \sum_n a_n \sum_\alpha \frac{\mathcal{H}'_{m\alpha}\mathcal{H}'_{\alpha n}}{E_m^{(0)} - E_\alpha^{(0)}} = 0\,, \tag{13.56}$$

in which the first sum is over all n without restriction, and for E in the denominator of the second-order perturbation term in (13.56) we replace E by $E_m^{(0)}$ in the spirit of perturbation theory. Equation (13.56) then implies the secular equation

$$\sum_{n=1}^{n} a_n \left[(E_m^{(0)} - E)\delta_{mn} + \mathcal{H}'_{mn} + \sum_\alpha \frac{\mathcal{H}'_{m\alpha}\mathcal{H}'_{\alpha n}}{E_m^{(0)} - E_\alpha^{(0)}} \right] = 0\,, \tag{13.57}$$

which yields an $n \times n$ secular equation with each matrix element given by

$$\mathcal{H}'_{mn} + \sum_\alpha \frac{\mathcal{H}'_{m\alpha}\mathcal{H}'_{\alpha n}}{E_m^{(0)} - E_\alpha^{(0)}}\,, \tag{13.58}$$

as indicated in (13.50).In degenerate $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory, we found that $\mathcal{H}'_{mn} = 0$ for a Γ_{15}^- level, and it was for this precise reason that we had to go to degenerate *second*-order perturbation theory. In this case, each state in the NDS couples to other states in the NDS only through an intermediate state outside of the NDS.

In second-order degenerate perturbation theory (13.49) shows us that for a threefold Γ_{15}^- level $\boldsymbol{k}\cdot\boldsymbol{p}$ degenerate perturbation theory will involve only states of $\Gamma_1^+, \Gamma_{12}^+, \Gamma_{15}^+$, or Γ_{25}^+ symmetry as intermediate states. In our discussion of nondegenerate $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory (see Sect. 13.3), we found that there was only one independent matrix element of $\boldsymbol{p}$ coupling a Γ_1^+ state to a Γ_{15}^- state. To facilitate the use of (13.48) and its more explicit form (13.58), we include in Table 13.2 a useful list of matrix elements of $\boldsymbol{p}$ between states of different symmetries for Γ point levels in cubic crystals. These matrix elements are found using the basis functions for each of the irreducible representations of O_h given in Table 10.2 and appearing also in Tables C.17 and 10.9 for the Γ point and Δ point of the diamond structure. Table 13.2 lists the nonvanishing matrix elements appearing in the $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory for electronic energy bands with cubic O_h symmetry.

For the matrix element A_2 in Table 13.2 we note with the help of Table 10.2 that the pertinent basis functions are $\Gamma_2^- = xyz$ and $\Gamma_{25,x}^+ = yz$. For

Table 13.2. Matrix elements for $\mathcal{H}' = \hbar \boldsymbol{k} \cdot \boldsymbol{p}/m$ in cubic O_h symmetry, where $\mathcal{H}'$ transforms as Γ_{15}^-

$(\Gamma_1^{\pm}\|\mathcal{H}'\|\Gamma_{15,\alpha}^{\mp}) = A_1 \frac{\hbar}{m} k_\alpha$	$A_1 = (\Gamma_1^{\pm}\|p_x\|\Gamma_{15,x}^{\mp})$
$(\Gamma_2^{\pm}\|\mathcal{H}'\|\Gamma_{25,\alpha}^{\mp}) = A_2 \frac{\hbar}{m} k_\alpha$	$A_2 = (\Gamma_2^{\pm}\|p_x\|\Gamma_{25,x}^{\mp})$
$(\Gamma_{12,1}^{\pm}\|\mathcal{H}'\|\Gamma_{15,x}^{\mp}) = A_3 \frac{\hbar}{m} k_x$ $(\Gamma_{12,1}^{\pm}\|\mathcal{H}'\|\Gamma_{15,y}^{\mp}) = A_3 \frac{\hbar}{m} k_y \omega^2$ $(\Gamma_{12,1}^{\pm}\|\mathcal{H}'\|\Gamma_{15,z}^{\mp}) = A_3 \frac{\hbar}{m} k_z \omega$ $(\Gamma_{12,2}^{\pm}\|\mathcal{H}'\|\Gamma_{15,x}^{\mp}) = A_3^* \frac{\hbar}{m} k_x$ $(\Gamma_{12,2}^{\pm}\|\mathcal{H}'\|\Gamma_{15,y}^{\mp}) = A_3^* \frac{\hbar}{m} k_y \omega$ $(\Gamma_{12,2}^{\pm}\|\mathcal{H}'\|\Gamma_{15,z}^{\mp}) = A_3^* \frac{\hbar}{m} k_z \omega^2$	$A_3 = (\Gamma_{12}^{\pm}\|p_x\|\Gamma_{15,x}^{\mp})$ $f_1 = f_2^* = x^2 + \omega y^2 + \omega^2 z^2$ $\omega = \exp(2\pi \mathrm{i}/3)$
$(\Gamma_{12,1}^{\pm}\|\mathcal{H}'\|\Gamma_{25,x}^{\mp}) = A_4 \frac{\hbar}{m} k_x$ $(\Gamma_{12,1}^{\pm}\|\mathcal{H}'\|\Gamma_{25,y}^{\mp}) = A_4 \frac{\hbar}{m} k_y \omega^2$ $(\Gamma_{12,1}^{\pm}\|\mathcal{H}'\|\Gamma_{25,z}^{\mp}) = A_4 \frac{\hbar}{m} k_z \omega$ $(\Gamma_{12,2}^{\pm}\|\mathcal{H}'\|\Gamma_{25,x}^{\mp}) = A_4^* \frac{\hbar}{m} k_x$ $(\Gamma_{12,2}^{\pm}\|\mathcal{H}'\|\Gamma_{25,y}^{\mp}) = A_4^* \frac{\hbar}{m} k_y \omega$ $(\Gamma_{12,2}^{\pm}\|\mathcal{H}'\|\Gamma_{25,z}^{\mp}) = A_4^* \frac{\hbar}{m} k_z \omega^2$	$A_4 = (\Gamma_{12}^{\pm}\|p_x\|\Gamma_{25,x}^{\mp})$ $f_1 = f_2^* = x^2 + \omega y^2 + \omega^2 z^2$
$(\Gamma_{15,x}^{\pm}\|\mathcal{H}'\|\Gamma_{15,x}^{\mp}) = 0$ $(\Gamma_{15,x}^{\pm}\|\mathcal{H}'\|\Gamma_{15,y}^{\mp}) = -A_5 \frac{\hbar}{m} k_z$ $(\Gamma_{15,x}^{\pm}\|\mathcal{H}'\|\Gamma_{15,z}^{\mp}) = A_5 \frac{\hbar}{m} k_y$	$A_5 = (\Gamma_{15,y}^{\pm}\|p_x\|\Gamma_{15,z}^{\mp})$
$(\Gamma_{15,y}^{\pm}\|\mathcal{H}'\|\Gamma_{15,x}^{\mp}) = A_5 \frac{\hbar}{m} k_z$ $(\Gamma_{15,y}^{\pm}\|\mathcal{H}'\|\Gamma_{15,y}^{\mp}) = 0$ $(\Gamma_{15,y}^{\pm}\|\mathcal{H}'\|\Gamma_{15,z}^{\mp}) = -A_5 \frac{\hbar}{m} k_x$	
$(\Gamma_{15,z}^{\pm}\|\mathcal{H}'\|\Gamma_{15,x}^{\mp}) = -A_5 \frac{\hbar}{m} k_y$ $(\Gamma_{15,z}^{\pm}\|\mathcal{H}'\|\Gamma_{15,y}^{\mp}) = A_5 \frac{\hbar}{m} k_x$ $(\Gamma_{15,z}^{\pm}\|\mathcal{H}'\|\Gamma_{15,z}^{\mp}) = 0$	
$(\Gamma_{15,x}^{\pm}\|\mathcal{H}'\|\Gamma_{25,x}^{\mp}) = 0$ $(\Gamma_{15,x}^{\pm}\|\mathcal{H}'\|\Gamma_{25,y}^{\mp}) = A_6 \frac{\hbar}{m} k_z$ $(\Gamma_{15,x}^{\pm}\|\mathcal{H}'\|\Gamma_{25,z}^{\mp}) = A_6 \frac{\hbar}{m} k_y$	$A_6 = (\Gamma_{15,x}^{\pm}\|p_y\|\Gamma_{25,z}^{\mp})$
$(\Gamma_{15,y}^{\pm}\|\mathcal{H}'\|\Gamma_{25,x}^{\mp}) = A_6 \frac{\hbar}{m} k_z$ $(\Gamma_{15,y}^{\pm}\|\mathcal{H}'\|\Gamma_{25,y}^{\mp}) = 0$ $(\Gamma_{15,y}^{\pm}\|\mathcal{H}'\|\Gamma_{25,z}^{\mp}) = A_6 \frac{\hbar}{m} k_x$	
$(\Gamma_{15,z}^{\pm}\|\mathcal{H}'\|\Gamma_{25,x}^{\mp}) = A_6 \frac{\hbar}{m} k_y$ $(\Gamma_{15,z}^{\pm}\|\mathcal{H}'\|\Gamma_{25,y}^{\mp}) = A_6 \frac{\hbar}{m} k_x$ $(\Gamma_{15,z}^{\pm}\|\mathcal{H}'\|\Gamma_{25,z}^{\mp}) = 0$	
$(\Gamma_{25,x}^{\pm}\|\mathcal{H}'\|\Gamma_{25,x}^{\mp}) = 0$ $(\Gamma_{25,x}^{\pm}\|\mathcal{H}'\|\Gamma_{25,y}^{\mp}) = -A_7 \frac{\hbar}{m} k_z$ $(\Gamma_{25,x}^{\pm}\|\mathcal{H}'\|\Gamma_{25,z}^{\mp}) = A_7 \frac{\hbar}{m} k_y$	$A_7 = (\Gamma_{25,x}^{\pm}\|p_y\|\Gamma_{25,z}^{\mp})$
$(\Gamma_{25,y}^{\pm}\|\mathcal{H}'\|\Gamma_{25,x}^{\mp}) = A_7 \frac{\hbar}{m} k_z$ $(\Gamma_{25,y}^{\pm}\|\mathcal{H}'\|\Gamma_{25,y}^{\mp}) = 0$ $(\Gamma_{25,y}^{\pm}\|\mathcal{H}'\|\Gamma_{25,z}^{\mp}) = -A_7 \frac{\hbar}{m} k_x$	
$(\Gamma_{25,z}^{\pm}\|\mathcal{H}'\|\Gamma_{25,x}^{\mp}) = -A_7 \frac{\hbar}{m} k_y$ $(\Gamma_{25,z}^{\pm}\|\mathcal{H}'\|\Gamma_{25,y}^{\mp}) = A_7 \frac{\hbar}{m} k_x$ $(\Gamma_{25,z}^{\pm}\|\mathcal{H}'\|\Gamma_{25,z}^{\mp}) = 0$	

$+$ denotes even and $-$ denotes odd states under inversion,
except for $f_1 \equiv f_1^+$ and $f_2 \equiv f_1^-$.
See Table 10.2 for explicit forms for the basis functions for the O_h group

A_4 we note that the basis function $\Gamma^-_{25,z} = z(x^2 - y^2)$ gives $C_2\Gamma^-_{25,z} = -\Gamma^-_{25,z}$ where C_2 denotes a rotation of π around the (011) axis. For A_5 we use as basis functions: $\Gamma^-_{15,x} = x$ and $\Gamma^+_{15,x} = yz(z^2 - y^2)$ which is odd under the interchange $y \leftrightarrow z$. For A_6 we use as basis functions: $\Gamma^+_{25,x} = yz$ and $\Gamma^-_{15,x} = x$, where $A_6 = (\Gamma^\pm_{15,x}|p_y|\Gamma^\mp_{25,z})$. For A_7 we use as basis functions: $\Gamma^+_{25,x} = yz$; $\Gamma^-_{25,x} = x(y^2 - z^2)$; $\Gamma^-_{25,z} = z(x^2 - y^2)$.

Let us make a few general comments about Table 13.2. Since $\mathcal{H}'$ is odd, only states of opposite parity are coupled. For each of the seven symmetry type couplings given in the table, there is only one independent matrix element. For example, the coupling between the Γ^+_{12} and Γ^-_{15} representations involve $2\times 3\times 3 = 18$ matrix elements, but there is only one *independent* matrix element:

$$(x|p_x|f_1) = (x|p_x|f_2) = \omega(y|p_y|f_1) = \omega^2(y|p_y|f_2) = \omega^2(z|p_z|f_1) = \omega(z|p_z|f_2)$$

and all others vanish. Here we write

$$\left.\begin{aligned} f_1 &= x^2 + \omega y^2 + \omega^2 z^2 \\ f_2 &= x^2 + \omega^2 y^2 + \omega z^2 \end{aligned}\right\} \tag{13.59}$$

as the basis functions for the Γ^+_{12} representation. For Γ^+_{25} symmetry we can take our basis functions as

$$\left\{\begin{aligned} &yz \\ &zx \\ &xy \end{aligned}\right. \quad \text{which in the table are denoted by} \quad \left\{\begin{aligned} &(\Gamma^+_{25,x}) \\ &(\Gamma^+_{25,y}) \\ &(\Gamma^+_{25,z})\,. \end{aligned}\right.$$

The three Γ^+_{25} basis functions are derived from three of the five atomic d functions, the other two being Γ^+_{12} functions. Using these results for the matrix elements, the secular equation (13.48) can be written as a function of k_x, k_y and k_z to yield the dispersion relations for the degenerate Γ^-_{15} bands as we move away from the Γ point $k = 0$ in the Brillouin zone.

Since $\Gamma^-_{15} \otimes \Gamma^-_{15} = \Gamma^+_1 + \Gamma^+_{12} + \Gamma^+_{15} + \Gamma^+_{25}$, and from (13.57), the secular equation (13.48) for the Γ^-_{15} levels involves the following sums:

$$\begin{aligned} F &= \frac{\hbar^2}{m^2}\sum_{\Gamma^+_1(n')} \frac{|(x|p_x|1)|^2}{E_n^{\Gamma^-_{15}}(0) - E_{n'}^{\Gamma^+_1}(0)}\,, \\ G &= \frac{\hbar^2}{m^2}\sum_{\Gamma^+_{12}(n')} \frac{|(x|p_x|f_1)|^2}{E_n^{\Gamma^-_{15}}(0) - E_{n'}^{\Gamma^+_{12}}(0)}\,, \\ H_1 &= \frac{\hbar^2}{m^2}\sum_{\Gamma^+_{25}} \frac{|(x|p_y|xy)|^2}{E_n^{\Gamma^-_{15}}(0) - E_n^{\Gamma^+_{25}}(0)}\,, \\ H_2 &= \frac{\hbar^2}{m^2}\sum_{\Gamma^+_{15}} \frac{|(x|p_y|xy(x^2 - y^2))|^2}{E_n^{\Gamma^-_{15}}(0) - E_n^{\Gamma^+_{15}}(0)}\,. \end{aligned} \tag{13.60}$$

We are now ready to solve the secular equation (13.48) using (13.57) to include the various terms which occur in second-order degenerate perturbation theory. Let us consider the diagonal entries first, as for example the xx entry. We can go from an initial $\Gamma_{15,x}^-$ state to the same final state through an intermediate Γ_1^+ state which brings down a k_x^2 term through the F term in (13.60). We can also couple the initial Γ_{15}^- state to itself through an intermediate $\Gamma_{12,1}^+$ or $\Gamma_{12,2}^+$ state, in either case bringing down a k_x^2 term through the G contribution – so far we have $Fk_x^2 + 2Gk_x^2$. We can also go from a $\Gamma_{15,x}^-$ state and back again through a $\Gamma_{25,y}^+$ or $\Gamma_{25,z}^+$ state to give a $(k_y^2 + k_z^2)H_1$ contribution and also through a $\Gamma_{15,y}^+$ or $\Gamma_{15,z}^+$ state to give a $(k_y^2 + k_z^2)H_2$ contribution. Therefore on the diagonal xx entry we get

$$Lk_x^2 + M(k_y^2 + k_z^2)\,, \quad \text{where} \quad L = F + 2G \quad \text{and} \quad M = H_1 + H_2\,. \tag{13.61}$$

Using similar arguments, we obtain the results for other diagonal entries yy and zz, using a cyclic permutation of indices.

Now let us consider an off-diagonal entry such as $(x|\mathcal{H}'|y)$, where we start with an initial $\Gamma_{15,x}^-$ state and go to a final $\Gamma_{15,y}^-$ state. This can be done through either of four intermediate states:

$$\begin{aligned}
&\text{Intermediate state} \quad \Gamma_1^+ \quad \text{gives} \quad k_x k_y F \\
&\text{Intermediate state} \quad \Gamma_{12}^+ \quad \text{gives} \quad (\omega^2 + \omega)k_x k_y G = -k_x k_y G \\
&\text{Intermediate state} \quad \Gamma_{15}^+ \quad \text{gives} \quad -k_x k_y H_2 \\
&\text{Intermediate state} \quad \Gamma_{25}^+ \quad \text{gives} \quad k_x k_y H_1\,.
\end{aligned}$$

Therefore we get $Nk_xk_y = (F - G + H_1 - H_2)k_xk_y$ for the total xy entry.

Using the same procedure we calculate the other four independent entries to the secular equation. Collecting terms we have the final result for the Taylor expansion of the secular equation for the Γ_{15}^- degenerate p-band:

$$\begin{vmatrix}
Lk_x^2 + M(k_y^2 + k_z^2) - \varepsilon(k) & Nk_xk_y & Nk_xk_z \\
Nk_xk_y & Lk_y^2 + M(k_z^2 + k_x^2) - \varepsilon(k) & Nk_yk_z \\
Nk_xk_z & Nk_yk_z & Lk_z^2 + M(k_x^2 + k_y^2) - \varepsilon(k)
\end{vmatrix} = 0\,. \tag{13.62}$$

The secular equation (13.62) is greatly simplified along the high symmetry directions. For a [100] axis, $k_y = k_z = 0$, and $k_x = \kappa$, then (13.62) reduces to

$$\begin{vmatrix}
L\kappa^2 - \varepsilon(\kappa) & 0 & 0 \\
0 & M\kappa^2 - \varepsilon(\kappa) & 0 \\
0 & 0 & M\kappa^2 - \varepsilon(\kappa)
\end{vmatrix} = 0\,, \tag{13.63}$$

which has the roots

$$\begin{aligned}\varepsilon(\kappa) &= L\kappa^2\\ \varepsilon(\kappa) &= M\kappa^2 \quad \text{twice}\,.\end{aligned} \tag{13.64}$$

The result in (13.64) must be consistent with the compatibility relations about the $\boldsymbol{k}=0$ (Γ-point) whereby

$$\Gamma_{15}^{+} \to \Delta_{1'} + \Delta_5\,, \tag{13.65}$$

in which the $\Delta_{1'}$ level is nondegenerate and the Δ_5 level is doubly degenerate.

Along a Λ [111] axis, $k_x = k_y = k_z = \kappa$ and the general secular equation of (13.62) simplifies into

$$\begin{vmatrix} (L+2M)\kappa^2-\varepsilon(\kappa) & N\kappa^2 & N\kappa^2 \\ N\kappa^2 & (L+2M)\kappa^2-\varepsilon(\kappa) & N\kappa^2 \\ N\kappa^2 & N\kappa^2 & (L+2M)\kappa^2-\varepsilon(\kappa) \end{vmatrix} = 0\,, \tag{13.66}$$

which can readily be diagonalized to give

$$\begin{aligned}\varepsilon(\kappa) &= \frac{L+2M+2N}{3}\kappa^2 \quad \text{once} \quad (\Lambda_2\,\text{level})\,,\\ \varepsilon(\kappa) &= \frac{L+2M-N}{3}\kappa^2 \quad \text{twice} \quad (\Lambda_3\,\text{level})\,,\end{aligned} \tag{13.67}$$

where the Λ_2 level is nondegenerate and the Λ_3 level is doubly degenerate.

The secular equation for a general κ point is more difficult to solve, but it can still be done in closed form by solving a cubic equation. In practice, the problem is actually simplified by including the effects of the *electron spin* (see Chap. 15). For each partner of the Γ_{15}^{-} levels we get a spin up state and a spin down state so that the secular equation is now a (6×6) equation. However, we will see that spin–orbit interaction simplifies the problem somewhat and the secular equation can be solved analytically.

The band parameters L, M, and N, which enter the secular equation (13.62), express the strength of the coupling of the Γ_{15}^{-} levels to the various other levels. In practice, these quantities are determined from experimental data. The cyclotron resonance experiment carried out along various high symmetry directions provides accurate values [31] for the band curvatures and hence for the quantities L, M and N. In the spirit of the $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory, solution of the secular equation provides the most general form allowed by symmetry for $E(\boldsymbol{k})$ about $\boldsymbol{k}=0$. The solution reduces to the proper form along the high symmetry directions, Δ, Λ and Σ. However, group theory cannot provide information about the magnitude of these coefficients. These magnitudes are most easily obtained from experimental data.

The $\boldsymbol{k}\cdot\boldsymbol{p}$ method has also been used to obtain the energy bands throughout the Brillouin zone for such semiconductors as silicon and germanium [17]. In the $\boldsymbol{k}\cdot\boldsymbol{p}$ approach of Cardona and Pollack, seven other bands outside this "nearly degenerate set" of eight ($\Gamma_1^+, \Gamma_2^-, \Gamma_{15}^-, \Gamma_{25}^+$) bands are allowed to couple to this nearly degenerate set of bands.

New features in the electronic energy band problem arise in going from points of lower symmetry to points of higher symmetry. For example, the $\boldsymbol{k}\cdot\boldsymbol{p}$ expansion can be used to connect a Λ point to an L point, along the Λ or (111) axis. The $\boldsymbol{k}\cdot\boldsymbol{p}$ method has been made to work well in this context, to parametrize theoretical calculations at high symmetry points and axes for use in regions of the Brillouin zone adjoining the locations for which the calculations were carried out. This use of $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory for a high symmetry point in the interior of the Brillouin zone is illustrated in the next section.

13.6 Nondegenerate $\boldsymbol{k}\cdot\boldsymbol{p}$ Perturbation Theory at a Δ Point

Figure 13.3 shows that important aspects of the electronic band structure for many cubic semiconductors occurs at $\boldsymbol{k}$ points away from $k = 0$ in the Brillouin zone, examples being the location of band extrema, of energy gaps and of carrier pockets for electrons and holes. In this section we illustrate how $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory is used both as an interpolation method and as an extrapolation method for the solution of the energy eigenvalues and eigenfunctions for an unperturbed crystal for $\boldsymbol{k}$ points of high symmetry away from $\boldsymbol{k} = 0$. In Sect. 13.7 we will show how $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory is used to interpret experiments where a probe is used to interact with a sample to study the electronic structure of the perturbed electronic system (from a group theory standpoint, the procedure is quite similar).

Let us consider the use of $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory for the group of the wave vector for a Δ point rather than about a Γ point, which was considered in Sects. 13.3–13.5. The momentum operator $\boldsymbol{p}$ in the $\boldsymbol{k}\cdot\boldsymbol{p}$ Hamiltonian transforms as a vector. For the group of the wave vector at a Δ point, the vector transforms as Δ_1 for the longitudinal component x and as Δ_5 for the transverse components y, z.

Typically for semiconductors the conduction bands are nondegenerate. In most cases the conduction band extrema are at $k = 0$ but for silicon the conduction band extrema are located at the six equivalent $(\Delta, 0, 0)$ locations, where Δ is 85% of the distance from Γ to X. The nondegenerate level in the conduction band at $k = 0$ has Γ_2^- symmetry, but has $\Delta_{2'}$ symmetry as we move away from $\boldsymbol{k} = 0$ in a (100) direction (see the compatibility relations for cubic groups in Sect. 10.7 and the character table for the group of the wave vector at a Δ point in Table 10.9 for the diamond structure).

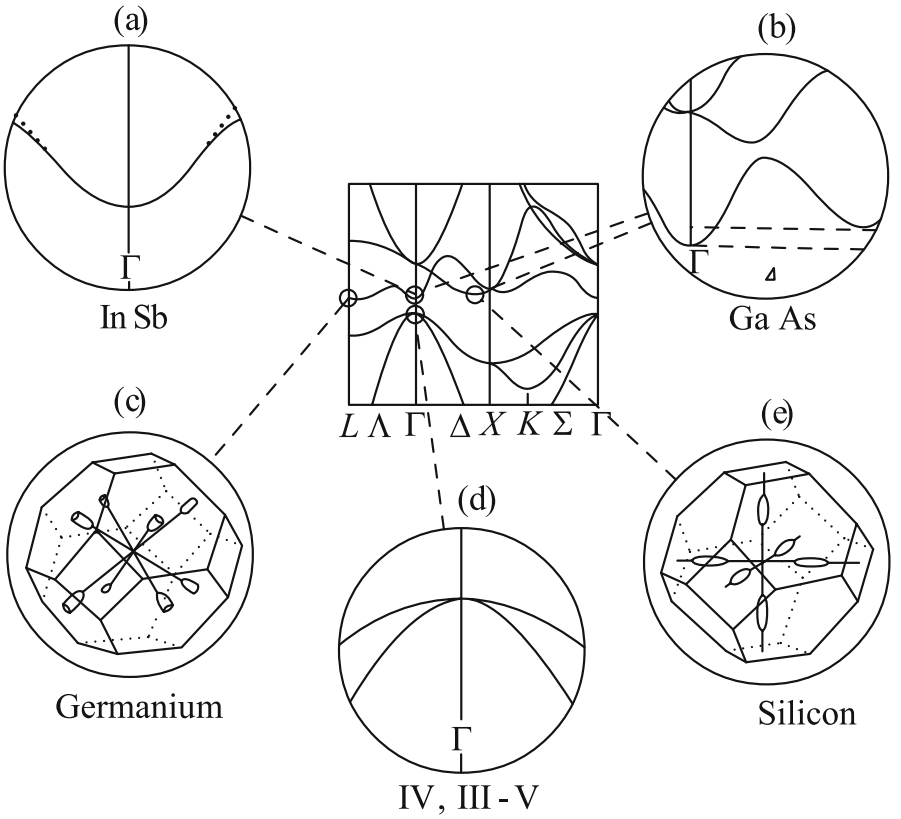

Fig. 13.3. Important details of the band structure of typical group IV and III–V semiconductors are found to occur both at $\boldsymbol{k}=0$ and for $\boldsymbol{k}$ points elsewhere in the Brillouin zone, including the location of conduction and valence band extrema and the location of carrier pockets

We now consider matrix elements of the form $(\Delta_{2'}|p_x|\Delta_{2'})$ which enter the expression for $E(\boldsymbol{k})$ about the Δ point. In first-order perturbation theory, we can have a nonvanishing contribution along k_x of the form $(\Delta_{2'}|p_x|\Delta_{2'})$ since $\Delta_1\otimes\Delta_{2'}=\Delta_{2'}$. Thus, there is in general a linear $\boldsymbol{k}$ term for $E(\boldsymbol{k})$ in the longitudinal direction. However, at the band extremum this matrix element vanishes (not by symmetry but because of the band extremum). We show below that the transverse matrix elements $(\Delta_{2'}|p_y|\Delta_{2'})$ and $(\Delta_{2'}|p_z|\Delta_{2'})$ vanish by symmetry along the Δ-axis. The second-order contributions to $E(\boldsymbol{k})$ are as follows:

$$E(\boldsymbol{k})=E(\boldsymbol{k}_0)+\frac{\hbar^2k_x^2}{2m_\ell^*}+\frac{\hbar^2(k_y^2+k_z^2)}{2m_t^*}\,. \tag{13.68}$$

The longitudinal terms $(\Delta_{2'}|\Delta_1|\Delta_j)$ require that the intermediate state Δ_j transforms as $\Delta_{2'}$ according to the compatibility relations, or else the matrix

element vanishes. States with $\Delta_{2'}$ symmetry at a Δ point arise by compatibility relations from Γ_{25}^{+}, Γ_{2}^{-}, and Γ_{12}^{-} states at $\boldsymbol{k} = 0$ and all of these intermediate states make contributions to a quadratic term in k_x^2 in the dispersion relation given by (13.68). For the transverse k_y and k_z terms, the matrix element $(\Delta_{2'}|\Delta_5|\Delta_j)$ requires the intermediate state Δ_j to transform as Δ_5. States with Δ_5 symmetry arise from $\Gamma_{25}^{\pm}$ levels at $\boldsymbol{k} = 0$.

Since the basis function for $\Delta_{2'}$ is yz (see Table 10.3), the vector component $\Delta_{5,y}$ couples to the z component of the intermediate state with symmetry $\Delta_{5,z}$ while the vector component $\Delta_{5,z}$ couples to the y component of the intermediate state with symmetry $\Delta_{5,y}$. Therefore there cannot be any nonvanishing matrix elements of the form $(\Delta_{2'}|\Delta_5|\Delta_{2'})$ for either a $\Delta_{5,y}$ or a $\Delta_{5,z}$ component of the vector.

However, in second-order we can have nonvanishing matrix elements about band extremum at $\boldsymbol{k}_0$ of the form $(\Delta_{2'}|\Delta_{5,y}|\Delta_{5,z})$ and $(\Delta_{2'}|\Delta_{5,z}|\Delta_{5,y})$ and therefore $E(\boldsymbol{k})$ about the Δ point extremum must be of the form of (13.68), in agreement with the expression used in solid state physics textbooks. As we move away from the Δ point extremum along the (100) axis, a linear term k_x in the $E(\boldsymbol{k})$ relation develops, but this term (allowed by group theory) is generally too small to be of significance to the constant energy contours applicable to practical situations, even for high doping levels and carrier pockets of larger volumes in k space.

The ellipsoidal form of $E(\boldsymbol{k})$ given by (13.68) is very common in semiconductor physics as we move away from $k = 0$. The case of the conduction band of silicon was shown here as an illustration, but similar ellipsoidal constant energy surfaces occur for germanium at the zone boundary L point and for other common III–V semiconductors at the X-point. These arguments given above can then be extended to other points in the Brillouin zone, and to two-band and three-band models for materials with cubic symmetry (see Problems 13.3 and 13.4). The $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory approach can of course also be extended to crystals described by other space groups.

13.7 Use of $\boldsymbol{k} \cdot \boldsymbol{p}$ Perturbation Theory to Interpret Optical Experiments

To carry out experiments in solid state crystalline physics, a probe is normally used to interrogate the materials system under investigation. These probes interact weakly with the system, causing perturbations that we measure in some way to provide information about the electronic structure of the solid state system. In this section we show how $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory is used to study the perturbation imposed on a material by an electromagnetic field and how information is provided by studying this perturbation with an electromagnetic (optical) probe.

The Hamiltonian in the presence of electromagnetic is discussed in Sect. 6.1, and the optical perturbation terms $\mathcal{H}'_{\text{opt}}$ are

$$\mathcal{H}'_{\text{opt}} = -\frac{e}{mc}\boldsymbol{A}\cdot\boldsymbol{p} + \frac{e^2A^2}{2mc^2}\,, \tag{13.69}$$

in which the lowest order term is

$$\mathcal{H}'_{\text{opt}} \cong -\frac{e}{mc}\boldsymbol{p}\cdot\boldsymbol{A}\,, \tag{13.70}$$

where the vector potential $\boldsymbol{A}$ relates to the optical fields and is not strongly affected by the crystal, while $\boldsymbol{p}$ relates directly to the momentum of electrons in the crystal and is strongly affected by the symmetry of the crystal. Therefore the momentum matrix elements $\langle v|\boldsymbol{p}|c\rangle$ coupling valence and conduction states mainly determine the strength of optical transitions in a low-loss (but finite loss) crystal. It is of interest that this same momentum matrix element governs $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation theory within a crystal and also governs the magnitudes of the effective mass components. With regard to the spatial dependence of the vector potential in (13.70) we can write

$$\boldsymbol{A} = \boldsymbol{A}_0\exp[i(\boldsymbol{k}_{n_i}\cdot\boldsymbol{r} - \omega t)]\,, \tag{13.71}$$

where for a loss-less medium described by a propagation constant $k_n = \tilde{n}\omega/c = 2\pi\tilde{n}/\lambda$ is a slowly varying function of $\boldsymbol{r}$, since $2\pi\tilde{n}/\lambda$ is much smaller than typical wave vectors in solids. Here $\tilde{n}$, ω, and λ are, respectively, the real part of the index of refraction, the optical frequency, and the wavelength of light. Thus, to the extent that we neglect the small spatial dependence of the optical propagation constant k_n, it is only the momentum matrix element $\langle v|\boldsymbol{p}|c\rangle$ coupling the valence and conduction bands that is important to lowest order perturbation theory. We note that electromagnetic interactions with a crystal involve the same matrix element that is connected with the effective mass components of the effective mass tensor for the unperturbed crystal. Group theory thus shows us that optical fields provide a very sensitive probe of the symmetry of a crystal by providing a way to measure this matrix element which is closely related to the effective mass tensor in the solid.

13.8 Application of Group Theory to Valley–Orbit Interactions in Semiconductors

In this section, we shall discuss the application of group theory to the impurity problem of a multivalley semiconductor, such as occurs in the donor carrier pockets in silicon and germanium. In the case of silicon, the lowest conduction bands occur at the six equivalent $(\Delta, 0, 0)$ points where $\Delta = 0.85$ on a scale where the Γ point is at the origin and the X point is at 1. In the

case of germanium, the conduction band minima occur at the L points so that the Fermi surface for electrons consists of eight equivalent half-ellipsoids of revolution (four full ellipsoids). Other cases where valley–orbit interactions are important are multivalley semiconductors, such as PbTe or Te, where the conduction and valence band extrema are both away from $\boldsymbol{k} = 0$.

Group theory tells us that the maximum degeneracy that energy levels or vibrational states can have with cubic symmetry is a threefold degeneracy. Cubic symmetry is imposed on the problem of donor doping of a semiconductor through the valley–orbit interaction which causes a partial lifting of the n-fold degeneracy of an n-valley semiconductor. In this section we show how group theory prescribes the partial lifting of this n-fold degeneracy. This effect is important in describing the ground state energy of a donor-doped n-valley semiconductor.

Our discussion of the application of group theory to the classification of the symmetries of the impurity levels in a degenerate semiconductor proceeds with the following outline:

(a) Review of the one-electron Hamiltonian and the effective mass Hamiltonian for a donor impurity in a semiconductor yielding hydrogenic impurity levels for a single-valley semiconductor.
(b) Discussion of the impurity states for multivalley semiconductors in the effective mass approximation.
(c) Discussion of the valley–orbit interaction. In this application we consider a situation where the lower symmetry group is not a subgroup of the higher symmetry group.

13.8.1 Background

In this section, we briefly review the one-electron Hamiltonian, effective mass approximation and the hydrogenic impurity problem for a single-valley semiconductor. We write the one-electron Hamiltonian for an electron in a crystal which experiences a perturbation potential $U(\boldsymbol{r})$ due to an impurity:

$$\left[\frac{p^2}{2m} + V(\boldsymbol{r}) + U(\boldsymbol{r})\right] \Psi(\boldsymbol{r}) = E\Psi(\boldsymbol{r})\,, \tag{13.72}$$

in which $V(\boldsymbol{r})$ is the periodic potential. In the effective mass approximation, the perturbing potential due to an impurity is taken as $U(\boldsymbol{r}) = -e^2/(\varepsilon r)$ where ε is the dielectric constant and the origin of the coordinate system is placed at the impurity sites. This problem is usually solved in terms of the effective mass theorem to obtain

$$\left[\frac{p^2}{2m^*_{\alpha\beta}} + U(\boldsymbol{r})\right] f_j(\boldsymbol{r}) = (E - E_j^0) f_j(\boldsymbol{r})\,, \tag{13.73}$$

where $m^*_{\alpha\beta}$ is the effective mass tensor for electrons in the conduction band about the band extremum at energy E_j^0, and $f_j(\boldsymbol{r})$ is the effective mass wave

function. We thus note that by replacing the periodic potential $V(\boldsymbol{r})$ by an effective mass tensor, we have *lost most of the symmetry information* contained in the original periodic potential. This symmetry information is restored by introducing the valley–orbit interaction, as in Sects. 13.8.2 and 13.8.3.

The simplest case for an impurity in a semiconductor is that for a shallow substitutional impurity level described by hydrogenic impurity states in a nondegenerate conduction band, as for example a Si atom substituted for a Ga atom in GaAs, a direct gap semiconductor with the conduction band extremum at the Γ point ($k = 0$). To satisfy the bonding requirements in this case, one electron becomes available for conduction and a donor state is formed. The effective mass equation in this case becomes

$$\left[\frac{p^2}{2m^*} - \frac{e^2}{\varepsilon r}\right] f(\boldsymbol{r}) = (E - E_j^0) f(\boldsymbol{r})\,, \tag{13.74}$$

where $U(\boldsymbol{r}) = -e^2/(\varepsilon r)$ is the screened Coulomb potential for the donor electron, ε is the low frequency dielectric constant, and the donor energies are measured from the band edge E_j^0. This screened Coulomb potential is expected to be a good approximation for $\boldsymbol{r}$ at a sufficiently *large distance from the impurity site*, so that ε is taken to be independent of r. The solutions to this hydrogenic problem are the hydrogenic levels

$$E_n - E_j^0 = -\frac{e^2}{2\varepsilon a_0^* n^2} \qquad\qquad n = 1, 2, \ldots\,, \tag{13.75}$$

where the effective Bohr radius is

$$a_0^* = \frac{\varepsilon \hbar^2}{m^* e^2}\,. \tag{13.76}$$

Since $(E_n - E_j^0) \sim m^*/\varepsilon^2$, we have *shallow* donor levels located below the band extrema, because of the large value of ε and the small value of m^* in many semiconductors of interest.

Group theoretical considerations enter in the following way. For many III–V compound semiconductors, the valence and conduction band extrema are at $\boldsymbol{k} = 0$ so that the effective mass Hamiltonian has full rotational symmetry. Since the hydrogenic impurity is embedded in a crystal with a periodic potential, the crystal symmetry (i.e., T_d point group symmetry) will perturb the hydrogenic levels and cause a splitting of various degenerate levels:

$$\begin{aligned}
&s \text{ levels} \rightarrow \Gamma_1 \quad (\text{no splitting})\,,\\
&p \text{ levels} \rightarrow \Gamma_{15} \quad (\text{no splitting})\,,\\
&d \text{ levels} \rightarrow \Gamma_{12} + \Gamma_{15} \quad (\text{splitting occurs})\,,\\
&f \text{ levels} \rightarrow \Gamma_2 + \Gamma_{15} + \Gamma_{25} \quad (\text{splitting occurs})\,.
\end{aligned}$$

In principle, if a multiplet has the same symmetry as an s or p level, then an interaction can occur giving rise to an admixture of states of similar symmetries. In practice, the splittings are very small in magnitude and the effects of

the crystal field are generally unimportant for shallow donor levels in single valley semiconductors.

13.8.2 Impurities in Multivalley Semiconductors

Group theory plays a more important role in the determination of impurity states in multivalley semiconductors than for the simple hydrogenic case described in Sect. 13.8.1. A common example of a multivalley impurity state is an As impurity in Si (or in Ge). In Si there are six equivalent valleys for the carrier pockets while for Ge there are four equivalent valleys. The multivalley aspect of the problem results in two departures from the simple hydrogenic series.

The first is associated with the fact that the constant energy surfaces are ellipsoids rather than spheres. We then write Schrödinger's equation for a single valley in the effective mass approximation as

$$\left[\frac{p_x^2+p_y^2}{2m_t}+\frac{p_z^2}{2m_l}-\frac{e^2}{\varepsilon \boldsymbol{r}}\right]=E\,f(\boldsymbol{r})\,, \tag{13.77}$$

in which m_t is the transverse mass component, m_l is the longitudinal mass component, and the energy E is measured from the energy band extremum. The appropriate symmetry group for the effective mass equation given by (13.77) is $D_{\infty h}$ rather than the full rotation group which applies to the hydrogenic impurity levels. This form for the effective mass Hamiltonian follows from the fact that the constant energy surfaces are ellipsoids of revolution, which in turn is a consequence of the selection rules for the $\boldsymbol{k}\cdot\boldsymbol{p}$ Hamiltonian at a Δ point (group of the wave vector C_{4v}) in the case of Si, and at an L point (group of the wave vector D_{3d}) in the case of Ge. The anisotropy of the kinetic energy terms corresponds to the anisotropy of the effective mass tensor. For example in the case of silicon $m_l/m_0=0.98$ (heavy mass), $m_t/m_0=0.19$ (light mass). This anisotropy in the kinetic energy terms results in a splitting of the impurity levels with angular momentum greater than 1, in accordance with the irreducible representations of $D_{\infty h}$. For example, in $D_{\infty h}$ symmetry we have the following correspondence with angular momentum states:

$$\begin{aligned}
&s \text{ states} \rightarrow \Sigma_g^+ = A_{1g}\,,\\
&p \text{ states} \rightarrow \Sigma_u^+ + \pi_u = A_{2u} + E_{1u}\,,\\
&d \text{ states} \rightarrow \Delta_g + \pi_g + \Sigma_g^+ = A_{1g} + E_{1g} + E_{2g}\,.
\end{aligned}$$

We note that s and d states are even (g) and p states are odd (u) under inversion in accordance with the character table for $D_{\infty h}$ (see Table A.34.). Thus a $2p$ level with an angular momentum of one splits into a twofold $2p^{\pm 1}$ level and a nondegenerate $2p^0$ level in which the superscripts denote

the n_l component of the angular momentum. Furthermore in $D_{\infty h}$ symmetry, the splitting of d-levels gives rise to the same irreducible representation (Σ_g^+) that describes the s-levels, and consequently a mixing of these levels occurs.

Referring back to (13.77), we note that the effective mass equation cannot be solved exactly if $m_l \neq m_t$. Thus, the donor impurity levels in these indirect gap semiconductors must be deduced from some approximate technique such as a variational calculation or using perturbation theory. The effective mass approximation itself works very well for these p-states because $|\psi_p|^2$ for p states vanishes for $\boldsymbol{r} = 0$; consequently, for $\boldsymbol{r}$ values small enough for central cell corrections to be significant, the wave function has a small amplitude and thus small $\boldsymbol{r}$ values do not contribute significantly to the expectation value of the energy for p-states.

13.8.3 The Valley–Orbit Interaction

The second departure from the hydrogenic series in a multivalley semiconductor is one that relates closely to group theory. This effect is most important for s-states, particularly for the $1s$ hydrogenic state.

For s-states, a sizable contribution to the expectation value for the energy is made by the perturbing potential for small $\boldsymbol{r}$. The physical picture of a spherically symmetric potential $U(\boldsymbol{r})$ for small $\boldsymbol{r}$ cannot fully apply because the tetrahedral bonding must become important for $|\boldsymbol{r}| \leq a$. This tetrahedral crystal field which is important within the central cell lifts the spherical symmetry of an isolated atom. Thus we need to consider corrections to the effective mass equation due to the tetrahedral crystal field. This tetrahedral crystal field term is called the valley–orbit effective Hamiltonian, $\mathcal{H}'_{\text{valley}-\text{orbit}}$, which couples equivalent conduction band extrema in the various conduction band valleys.

To find the wave functions for the donor states in a multivalley semiconductor, we must find linear combinations of wave functions from each of the conduction band valleys that transform as irreducible representations of the crystal field about the impurity ion. For example, in silicon, the symmetrized linear combination of valley wave functions is in the form

$$\psi^{\gamma}(\boldsymbol{r}) = \sum_{j=1}^{6} A_j^{\gamma} f_j(\boldsymbol{r}) u_{j,k_0^j}(\boldsymbol{r}) e^{i\boldsymbol{k_0^j}\cdot\boldsymbol{r}}\,, \tag{13.78}$$

in which $\psi^{\gamma}(\boldsymbol{r})$ denotes one of six possible linear combinations of the wave functions for the six carrier pockets denoted by γ. The index j is the valley index and $f_j(\boldsymbol{r})$ is the envelope effective mass wave function, while $u_{j,k_0^j}(\boldsymbol{r})$ is the periodic part of the Bloch function in which k_0^j is the wave vector to the band minimum of valley j. The six equivalent valleys along the (100) axes for the conduction band of silicon are shown in Fig. 13.4(a). The indices j which label the various ellipsoids or valleys in Fig. 13.4(a) correspond to the

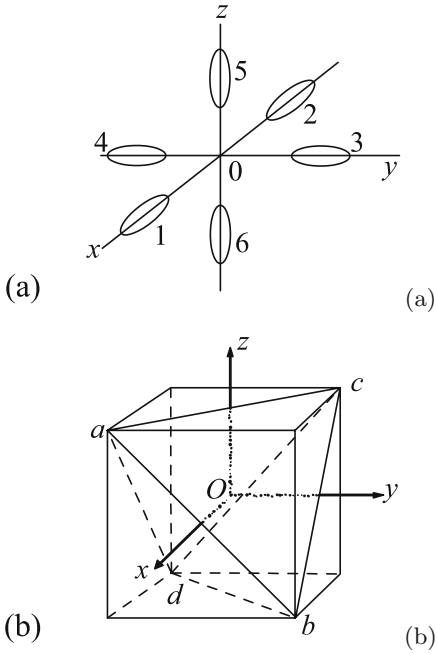

Fig. 13.4. (**a**) Constant energy ellipsoids of the conduction-band minima of silicon along {100} directions in reciprocal space. (**b**) The regular tetrahedron inscribed inside a cube, useful for seeing the symmetry operations of the six valleys in (**a**)

Table 13.3. Irreducible representations contained in Γ_{valley} sites of Si

	E	$8C_3$	$3C_2$	$6\sigma_d$	$6S_4$	
$\chi_{\text{valley sites}}$	6	0	2	2	0	$= \Gamma_1 + \Gamma_{12} + \Gamma_{15}$

indices j of (13.78). The local symmetry close to the impurity center is T_d, reflecting the tetrahedral bonding at the impurity site. The character table for the T_d point group is shown in Table A.32. The diagram which is useful for finding which valleys are invariant under the symmetry operations of T_d is given in Fig. 13.4(b). To get the equivalence transformation for the valley sites, we ask for the number of valleys which remain invariant under the various symmetry operations of T_d. This is equivalent to finding Γ^{equiv} or $\Gamma_{\text{valley sites}}$, which forms a *reducible* representation of point group T_d. From Fig. 13.4(b), we immediately see that the characters for the reducible representation $\Gamma_{\text{valley sites}}$ are (see Table 13.3), and that the irreducible representations contained in $\Gamma_{\text{valley sites}}$ are the $\Gamma_1 + \Gamma_{12} + \Gamma_{15}$ irreducible representations of the point group T_d. To find the splitting of a level we must take the direct product

of the symmetry of the level with $\Gamma_{\text{valley sites}}$, *provided that the level itself transforms as an irreducible representation of group* T_d:

$$\Gamma_{\text{level}} \otimes \Gamma_{\text{valley sites}} \,. \tag{13.79}$$

Since Γ_{level} for s-states transforms as Γ_1, the level splitting for s-states is just $\Gamma_{\text{valley sites}} = \Gamma_1 + \Gamma_{12} + \Gamma_{15}$:

$$\begin{aligned} &\text{—}\Gamma_{15} \\ &\text{—}\Gamma_{12} \\ &\text{—}\Gamma_1 \,. \end{aligned}$$

The appropriate linear combination of valley functions corresponding to each of these irreducible representations is (using the notation from (13.78)):

$$\left.\begin{aligned} A_j^{(\Gamma_1)} &= \tfrac{1}{\sqrt{6}}(1,1,1,1,1,1)\,, \\ A_j^{(\Gamma_{12,1})} &= \tfrac{1}{\sqrt{6}}(1,1,\omega,\omega,\omega^2,\omega^2) \\ A_j^{(\Gamma_{12,2})} &= \tfrac{1}{\sqrt{6}}(1,1,\omega^2,\omega^2,\omega,\omega) \end{aligned}\right\}, \qquad \left.\begin{aligned} A_j^{(\Gamma_{15,1})} &= \tfrac{1}{\sqrt{2}}(1,-1,0,0,0,0) \\ A_j^{(\Gamma_{15,2})} &= \tfrac{1}{\sqrt{2}}(0,0,1,-1,0,0) \\ A_j^{(\Gamma_{15,3})} &= \tfrac{1}{\sqrt{2}}(0,0,0,0,1,-1) \end{aligned}\right\}, \tag{13.80}$$

in which each of the six components of the coefficients A_j^{γ} refers to one of the valleys. The totally symmetric linear combination Γ_1 is a nondegenerate level, while the Γ_{12} basis functions have two partners which are given by $f_1 = x^2 + \omega y^2 + \omega^2 z^2$ and $f_2 = f_1^*$ and the Γ_{15} basis functions have three partners (x, y, z).

The analysis for the p-levels is more complicated because the p-levels in $D_{\infty h}$ do not transform as irreducible representations of group T_d. The p-level in group $D_{\infty h}$ transforms as a vector, with A_{2u} and E_{1u} symmetries for the longitudinal and transverse components, respectively. Since T_d does not form a subgroup of $D_{\infty h}$ we write the vector for group T_d as a sum over its longitudinal and transverse components

$$\Gamma_{\text{vec.}} = \Gamma_{\text{longitudinal}} + \Gamma_{\text{transverse}} \,, \tag{13.81}$$

where $\Gamma_{\text{vec.}} = \Gamma_{15}$. We treat the longitudinal component of the vector as forming a σ-bond and the transverse component as forming a π-bond so that $\Gamma_{\text{longitudinal}} = \Gamma_1$ and $\Gamma_{\text{transverse}} = \Gamma_{15} - \Gamma_1$, where we note that

$$\Gamma_{15} \otimes (\Gamma_1 + \Gamma_{12} + \Gamma_{15}) = \Gamma_{15} + (\Gamma_{15} + \Gamma_{25}) + (\Gamma_1 + \Gamma_{12} + \Gamma_{15} + \Gamma_{25})\,. \tag{13.82}$$

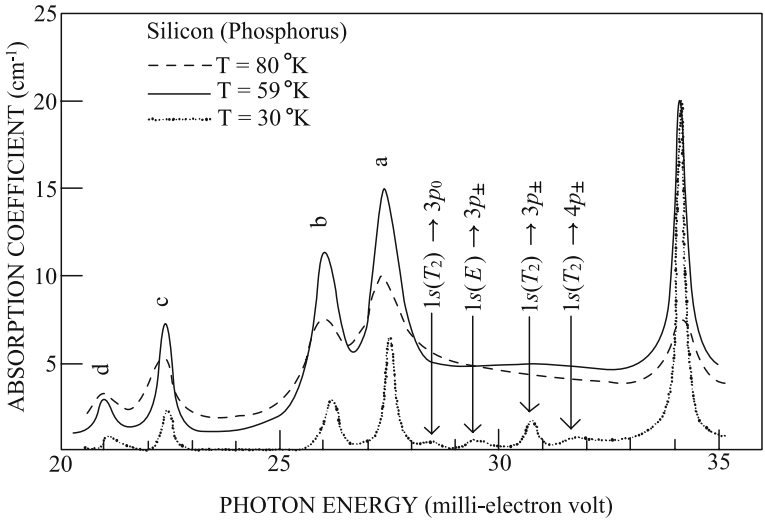

Fig. 13.5. Excitation spectrum of phosphorus donors in silicon. The donor concentration is $N_{\mathrm{D}} \sim 5 \times 10^{15}\,\mathrm{cm}^{-3}$. Various donor level transitions to valley–orbit split levels are indicated. The labels for the final state of the optical transitions are in accordance with the symmetries of point group T_d

We thus obtain for the longitudinal (Γ^{2p_0}) and transverse ($\Gamma^{2p_\pm}$) levels:

$$\Gamma^{2p_0} = \Gamma_{\text{valley sites}} \otimes \Gamma_1 = \Gamma_1 + \Gamma_{12} + \Gamma_{15} \quad \text{for} \quad m_\ell = 0 \tag{13.83}$$

$$\Gamma^{2p_\pm} = \Gamma_{\text{valley sites}} \otimes (\Gamma_{15} - \Gamma_1) = 2\Gamma_{15} + 2\Gamma_{25} \quad \text{for} \quad m_\ell = \pm 1 \tag{13.84}$$

for group T_d. If we perform high resolution spectroscopy experiments for the donor impurity levels, we would expect to observe transitions between the various $1s$ multiplets to the various $2p$-multiplets, as allowed by symmetry selection rules [46]. Experimental evidence for the splitting of the degeneracy of the $1s$ donor levels in silicon is provided by infrared absorption studies [4, 67]. An experimental trace for the excitation spectrum of phosphorus impurities in silicon is shown in Fig. 13.5 for several sample temperatures. The interpretation of this spectrum follows from the energy level diagram in Fig 13.6 [46].

It is of interest that the valley orbit splitting effect is only important for the $1s$ levels. For the higher levels, the tetrahedral site location of the impurity atom becomes less important since the Bohr orbit for the impurity level increases as n^2 which qualitatively follows from

$$a^*_{\mathrm{Bohr}} = \frac{\varepsilon \hbar^2}{m^* e^2} n^2 \tag{13.85}$$

where n is the principal quantum number for the donor impurity level.

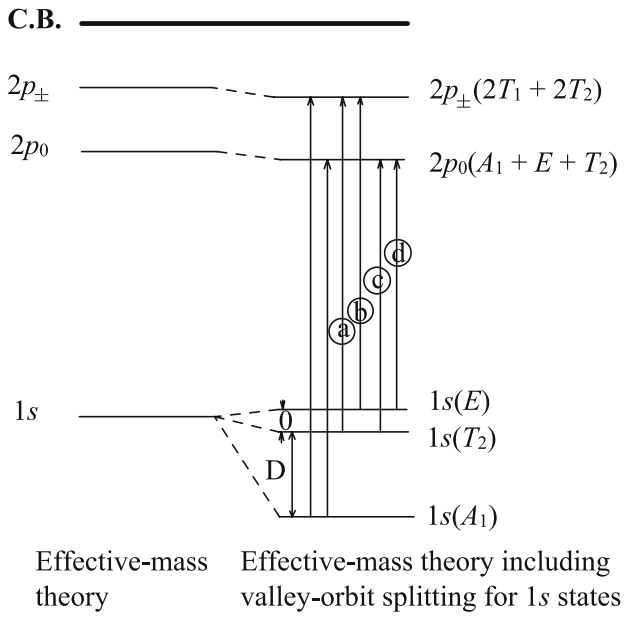

Fig. 13.6. Energy-level scheme for transitions from the valley–orbit split $1s$ multiplet of states to the $2p_0, 2p_\pm$ levels. The irreducible representations for the various valley–orbit split levels in T_d symmetry are indicated. The conduction band edge (C.B.) is also indicated schematically as are the splittings between the three constituents of the valley-orbit split $1s$ level, showing a separation of D between the A_1 and T_2 levels and a separation of O between the T_2 and E levels

In addition to spectroscopic studies of impurity states, these donor states for multivalley semiconductors have been studied by the ENDOR technique [35]. Here the nuclear resonance of the ^{29}Si atoms is observed. The random distribution of the ^{29}Si sites with respect to the donor impurity sites is used to study the spatial dependence of the donor wavefunction, and to determine the location in $\boldsymbol{k}$-space of the conduction band extrema.

Selected Problems

13.1. (a) Using $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory, find the dispersion relation $E(\boldsymbol{k})$ for the nondegenerate Γ_2^- (or $\Gamma_{2'}$) band around the conduction band extremum near $k = 0$ for a simple cubic solid.

(b) The conduction band for germanium which crystallizes in the diamond structure has Γ_2^- (or $\Gamma_{2'}$) symmetry. Explain how your result in (a) can be used to describe $E(\boldsymbol{k})$ about $\boldsymbol{k} = 0$ for the conduction band of germanium. What modifications occur to (13.12) and (13.14)?

13.2. In this problem, use $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory to find the form of the secular equation for the valence band of Si with Γ_{25}^{+} symmetry, neglecting the spin–orbit interaction

(a) Which intermediate states couple to the Γ_{25}^{+} valence band states in second-order $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory?
(b) Which matrix elements (listed in Table 13.2) enter the secular equation in (a)?
(c) Write the secular equation for the Γ_{25}^{+} valence bands that is analogous to (13.62) for the Γ_{15}^{-} band?
(d) Using the general result in (a), find the special form of the secular equation for the Γ_{25}^{+} valence band that is obtained along a Λ (111) axis?

13.3. (a) Using $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory, find the form of the $E(\boldsymbol{k})$ relation near the L-point in the Brillouin zone for a face centered cubic lattice arising from the lowest energy levels. In the free electron model these levels are doubly degenerate and have L_1 and L_2' symmetry. Which of the nonvanishing $\boldsymbol{k} \cdot \boldsymbol{p}$ matrix elements at the L-point are equal to each other by symmetry?
(b) Using $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory, find the form of $E(\boldsymbol{k})$ for a nondegenerate band with W_1 symmetry about the W point in the FCC lattice (see Table C.12).

13.4. The form of the $E(\boldsymbol{k})$ relation for the second level of the empty lattice for a BCC system was discussed in Problem 12.6 for both the empty lattice and in the presence of a small periodic potential

(a) Now consider the lowest energy levels at the H point where the Δ axis along (100) meets the Brillouin zone boundary (see Fig. 12.6 and Tables C.15 and C.8). Find the form of the dispersion relations near the H point using $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory and compare your results with the dispersion relations for Na shown in Fig. 12.6(b).
(b) Using symmetry arguments, why is the splitting between H_1 and H_{15} so much larger than between H_{12} and H_{15}?

13.5. Find the symmetries and appropriate linear combination of valley functions for the $1s$ and $2p$ donor levels for germanium (conduction band minima at the L-point in the Brillouin zone), including the effect of valley–orbit interaction. Indicate the transitions expected in the far infrared spectra for these low temperature donor level states.