Quantum Processes in Semiconductors

SECOND EDITION

B. K. RIDLEY

Quantum Processes in Semiconductors

Second Edition

B. K. Ridley
Professor of Physics
University of Essex

Preface to the Second Edition

This section edition contains three new chapters—'Quantum processes in a magnetic field, Scattering in a degenerate gas, and Dynamic screening'—which I hope will enhance the usefulness of the book. Following the ethos of the first edition I have tried to make the rather heavy mathematical content of these new topics as straightforward and accessible as possible. I have also taken the opportunity to make some corrections and additions to the original material—a brief account of alloy scattering is now included—and I have completely rewritten the section on impact ionization. An appendix on the average separation of impurities has been added, and the term 'third-body exclusion' has become 'statistical screening', but otherwise the material in the first edition remains substantially unchanged.

Thorpe-le-Soken 1988   B. K. R.

Preface to the First Edition

It is a curious fact that in spite of, or perhaps because of, their overwhelming technological significance, semiconductors receive comparatively modest attention in books devoted to solid state physics. A student of semiconductor physics will find the background theory common to all solids well described, but somehow all the details, the applications, and the examples—just those minutiae which reveal so vividly the conceptual cast of mind which clarifies a problem—are all devoted to metals or insulators or, more recently, to amorphous or even liquid matter. Nor have texts devoted exclusively to semiconductors, excellent though they are, fully solved the student's problem, for they have either attempted global coverage of all aspects of semiconductor physics or concentrated on the description of the inhomogeneous semiconducting structures which are used in devices, and in both cases they have tended to confine their discussion of basic physical processes to bare essentials in order to accommodate breadth of coverage in the one and emphasis on application in the other. Of course, there are distinguished exceptions to these generalizations, texts which have specialized on topics within semiconductor physics, such as statistics and band structure to take two examples, but anyone who has attempted to teach the subject to postgraduates will, I believe, agree that something of a vacuum exists, and that filling it means resorting to research monographs and specialist review articles, many of which presuppose a certain familiarity with the field.

Another facet to this complex and fascinating structure of creating, assimilating, and transmitting knowledge is that theory, understandably enough, tends to be written by theoreticians. Such is today's specialization that the latter tend to become removed from direct involvement in the empirical basis of their subject to a degree that makes communication with the experimentalist fraught with mutual incomprehension. Sometimes the difficulty is founded on a simple confusion between the disparate aims of mathematics and physics—an axiomatic formulation of a theory may make good mathematical sense but poor physical sense—or it may be founded on a real subtlety of physical behaviour perceived by one and incomprehended by the other, or more usually it may be founded on ignorance of each other's techniques, of the detailed analytic and numerical approximations propping up a theory on the one hand, and of the detailed method and machinery propping up an experimental result on the other. Certainly, experimentalists cannot avoid being theoreticians from time to time, and they have to be aware of the basic theoretical

structure of their subject. As students of physics operating in an area where physical intuition is more important than logical deduction they are not likely to appreciate a formalistic account of that basic structure even though it may possess elegance. Intuition functions on rough approximation rather than rigour, but too few accounts of theory take that as a guide.

This book, then, is written primarily for the post-graduate student and the experimentalist. It attempts to set out the theory of those basic quantum-mechanical processes in homogeneous semiconductors which are most relevant to applied semiconductor physics. Therefore the subject matter is concentrated almost exclusively on electronic processes. Thus no mention is made of phonon–phonon interactions, nor is the optical absorption by lattice modes discussed. Also, because I had mainstream semiconductors like silicon and gallium arsenide in mind, the emphasis is on crystalline materials in which the electrons and holes in the bands obey non-degenerate statistics, and little mention is made of amorphous and narrow-gap semiconductors. Only the basic quantum-mechanics is discussed; no attempt is made to follow detailed applications of the basic theory in fields such as hot electrons, negative-differential resistance, acousto-electric effects, etc. To do that would more than triple the size of the book. The theoretical level is at elementary first- and second-order perturbation theory, with not a Green's function in sight; this is inevitable, given that the author is an experimentalist with a taste for doing his own theoretical work. Nevertheless, those elementary conceptions which appear in the book are, I believe, the basic ones in the field which most of us employ in everyday discussions, and since there is no existing book to my knowledge which contains a description of all these basic processes I hope that this one will make a useful reference source for anyone engaged in semiconductor research and device development.

Finally, a word of caution for the reader. A number of treatments in the book are my own and are not line-by-line reproductions of standard theory. Principally, this came about because the latter did not exist in a form consistent with the approach of the book. One or two new expressions have emerged as a by-product, although most of the final results are the accepted ones. Where the treatment is mine, the text makes this explicit.

Colchester 1981   B.K.R.

Contents

1. Band structure of semiconductors

1.1. The crystal Hamiltonian

For an assembly of atoms the classical energy is the sum of the following:
- (a) the kinetic energy of the nuclei;
- (b) the potential energy of the nuclei in one another's electrostatic field;
- (c) the kinetic energy of the electrons;
- (d) the potential energy of the electrons in the field of the nuclei;
- (e) the potential energy of the electrons in one another's field;
- (f) the magnetic energy associated with the spin and the orbit.

Dividing the electrons into core and valence electrons and leaving out magnetic effects leads to the following expression for the crystal Hamiltonian:

$$H = \sum_l \frac{\mathbf{p}_l^2}{2M_l} + \sum_{l,m} U(\mathbf{R}_l - \mathbf{R}_m) + \sum_i \frac{\mathbf{p}_i^2}{2m} + \sum_{i,l} V(\mathbf{r}_i - \mathbf{R}_l) + \sum_{i,j} \frac{e^2/4\pi\epsilon_0}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{1.1}$$

where l and m label the ions, i and j label the electrons, $\mathbf{p}$ is the momentum, M is the ionic mass, m is the mass of the electron, $U(\mathbf{R}_l - \mathbf{R}_m)$ is the interionic potential, and $V(\mathbf{r}_i - \mathbf{R}_l)$ is the valence–electron–ion potential.

The Schrödinger equation determines the time-independent energies of the system:

$$H\Xi = E\Xi \tag{1.2}$$

where H is now the Hamiltonian operator.

1.2. Adiabatic approximation

The mass of an ion is at least a factor of $1\cdot8 \times 10^3$ greater than that of an electron, and for most semiconductors the factor is well over 10^4. For comparable energies and perturbations ions therefore move some 10^2 times slower than do electrons, and the latter can be regarded as instantaneously adjusting their motion to that of the ions. Therefore the total wavefunction is approximately of the form

$$\Xi = \Psi(\mathbf{r}, \mathbf{R})\Phi(\mathbf{R}) \tag{1.3}$$

where $\Phi(\mathbf{R})$ is the wavefunction for all the ions and $\Psi(\mathbf{r}, \mathbf{R})$ is the wavefunction for all the electrons instantaneously dependent on the ionic position.

The Schrödinger equation can be written

$$\Psi(\mathbf{r}, \mathbf{R})H_L\Phi(\mathbf{R}) + \Phi(\mathbf{R})H_e\Psi(\mathbf{r}, \mathbf{R}) + H'\Psi(\mathbf{r}, \mathbf{R})\Phi(\mathbf{R}) = E\Psi(\mathbf{r}, \mathbf{R})\Phi(\mathbf{R}) \quad (1.4)$$

where

$$H'\Psi(\mathbf{r}, \mathbf{R})\Phi(\mathbf{R}) = H_L\Psi(\mathbf{r}, \mathbf{R})\Phi(\mathbf{R}) - \Psi(\mathbf{r}, \mathbf{R})H_L\Phi(\mathbf{R}) \quad (1.5)$$

$$H_L = \sum_l \frac{\mathbf{p}_l^2}{2M_l} + \sum_{l,m} U(\mathbf{R}_l - \mathbf{R}_m) \quad (1.6)$$

$$H_e = \sum_i \frac{\mathbf{p}_i^2}{2m} + \sum_{i,l} V(\mathbf{r}_i - \mathbf{R}_l) + \sum_{i,j} \frac{e^2/4\pi\epsilon_0}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (1.7)$$

The relative contribution of H' is of the order m/M_l. The adiabatic approximation consists of neglecting this term. In this case eqn (1.4) splits into a purely ionic equation

$$H_L\Phi(\mathbf{R}) = E_L\Phi(\mathbf{R}) \quad (1.8)$$

and a purely electronic equation

$$H_e\Psi(\mathbf{r}, \mathbf{R}) = E_e\Psi(\mathbf{r}, \mathbf{R}). \quad (1.9)$$

1.3. Phonons

Provided that the ions do not move far from their equilibrium positions in the solid their motion can be regarded as simple harmonic. If the equilibrium position of an ion is denoted by the vector $\mathbf{R}_{l0}$ and its displacement by $\mathbf{u}_l$, the Hamiltonian can be written

$$H_L = \sum_l \frac{\mathbf{p}_l^2}{2M_l} + \sum_{l,m} D_{l,m}(\mathbf{R}_l - \mathbf{R}_m)\mathbf{u}_l\mathbf{u}_m + H_{L0}(\mathbf{R}_{l0}) + H'_L \quad (1.10)$$

where $D_{lm}(\mathbf{R}_l - \mathbf{R}_m)$ is the restoring force per unit displacement, $H_{L0}(\mathbf{R}_{l0})$ is an additive constant dependent only on the equilibrium separation of the ions, and H'_L represents the contribution of anharmonic forces. The displacements can be expanded in terms of the normal modes of vibration of the solid. The latter take the form of longitudinally polarized and transversely polarized acoustic waves plus, in the case of lattices with a basis, i.e. more than one atom per primitive unit cell, longitudinally and transversely polarized 'optical' modes. (See Section 3.9 for an account of the theory for long-wavelength acoustic modes.) Ionic motion therefore manifests itself in terms of travelling plane waves

$$\mathbf{u}(\omega, \mathbf{q}) = \mathbf{u}_0 \exp\{i(\mathbf{q}.\mathbf{r} - \omega t)\} \quad (1.11)$$

which interact weakly with one another through the anharmonic term H'_L. Figure 1.1 shows the typical dispersion relation between ω and $\mathbf{q}$.
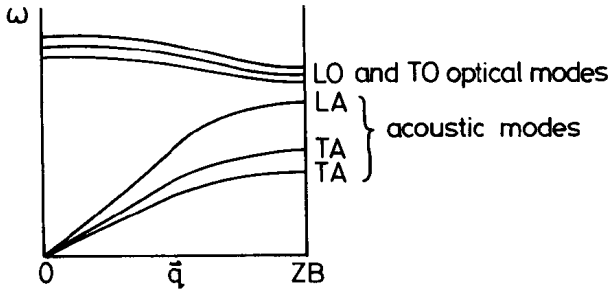
FIG. 1.1. Dispersion of lattice waves.

The energy in a mode is given by

$$E(\omega, \mathbf{q}) = \{n(\omega, \mathbf{q}) + \tfrac{1}{2}\}\hbar\omega \qquad (1.12)$$

where $n(\omega, \mathbf{q})$ is the statistical average number of phonons, i.e. vibrational quanta, excited. At thermodynamic equilibrium $n(\omega, q) = n(\omega)$ is given by the Bose–Einstein function for a massless particle

$$n(\omega) = \frac{1}{\exp(\hbar\omega/k_B T) - 1}. \qquad (1.13)$$

The following points should be noted.

(1) The limits of $\mathbf{q}$ according to periodic boundary conditions are $2\pi/Na$ and the Brillouin zone boundary, where N is the number of unit cells of length a along the cavity.

(2) The magnitude of a wavevector component is $2\pi l/Na$, where l is an integer. The curves in Fig. 1.1 are really closely spaced points.

(3) An impurity or other defect may introduce localized modes of vibration in its neighbourhood if its mass and binding energy are different enough from those of its host.

(4) For long-wavelength acoustic modes $\omega = v_s q$. For others it is often useful to approximate their dispersion by $\omega = constant$.

1.4. The one-electron approximation

If the electron–electron interaction is averaged we can regard any deviation from this average as a small perturbation. Thus we replace the repulsion term as follows:

$$\sum_{i,j} \frac{(e^2/4\pi\epsilon_0)}{|\mathbf{r}_i - \mathbf{r}_j|} = H_{e0} + H_{ee} \qquad (1.14)$$

where H_{e0} contributes a constant repulsive component to the electronic energy and H_{ee} is a fluctuating electron–electron interaction which can be

regarded as small. If H_{ee} is disregarded each electron reacts independently with the lattice of ions. Consequently we can take

$$\Psi(\mathbf{r}, \mathbf{R}) = \prod_i \psi_i(\mathbf{r}_i, \mathbf{R}) \tag{1.15}$$

with the proviso that the occupation of the one-electron states is in accordance with the Pauli exclusion principle. We obtain the one-electron Schrödinger equation

$$H_{ei}\psi_i(\mathbf{r}_i\mathbf{R}) = E_{ei}\psi_i(\mathbf{r}_i, \mathbf{R}) \tag{1.16}$$

where

$$H_{ei} = \frac{\mathbf{p}_i^2}{2m} + \sum_l V(\mathbf{r}_i - \mathbf{R}_l). \tag{1.17}$$

This Hamiltonian still depends on the fluctuating position of ions, and it is useful to reduce the Hamiltonian into one that depends on the interaction with the ions in their equilibrium positions with the effect of ionic vibrations taken as a perturbation. Thus we take

$$H_{ei} = H_{e0i} + H_{ep} \tag{1.18}$$

$$H_{e0i} = \frac{\mathbf{p}_i^2}{2m} + \sum_l V(\mathbf{r}_i - \mathbf{R}_{l0}) \tag{1.19}$$

where the H_{ep} is the electron–phonon interaction. The electronic band structure is obtained from (dropping the subscripts i and e)

$$\left\{ \frac{\mathbf{p}^2}{2m} + \sum_l V(\mathbf{r} - \mathbf{R}_{l0}) \right\} \psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{1.20}$$

1.5. Bloch functions

In the case of a perfectly periodic potential the eigenfunction is a Bloch function:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r})\exp(i\mathbf{k}.\mathbf{r}) \tag{1.21}$$

$$u_{n\mathbf{k}}(\mathbf{r}+\mathbf{R}) = u_{n\mathbf{k}}(\mathbf{r}) \tag{1.22}$$

where $\mathbf{R}$ is a vector of the Bravais lattice, n labels the band and $\mathbf{k}$ is the wavevector of the electron in the first Brillouin zone (Fig. 1.2). From eqns (1.21) and (1.22) if follows that

$$\psi_{n\mathbf{k}}(\mathbf{r}+\mathbf{R}) = \psi_{n\mathbf{k}}(\mathbf{r})\exp(i\mathbf{k}.\mathbf{R}). \tag{1.23}$$

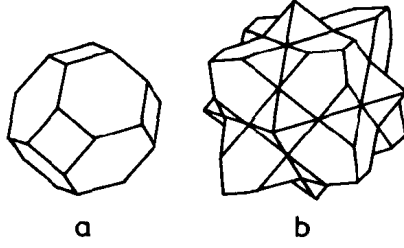If a macroscopic volume V is chosen whose shape is a magnified version

FIG. 1.2. The first and second zones for a face-centred cubic lattice. The first has half the volume of the cube that is determined by extending the six square faces. The second has the same volume as this cube.

of the primitive cell, then we can apply the periodic boundary condition

$$\psi_{n\mathbf{k}}(\mathbf{r}+N\mathbf{a}) = \psi_{n\mathbf{k}}(\mathbf{r}) \tag{1.24}$$

where $\mathbf{a}$ is a vector of the unit cell and N is the number of unit cells along the side of V in the direction of $\mathbf{a}$. This decouples the properties of the wavefunction from the size of a crystal, provided the crystal is macroscopic. Equations (1.23) and (1.24) constrain $\mathbf{k}$ such that

$$\exp(i\mathbf{k}.N\mathbf{a}) = 1.$$

Therefore

$$\mathbf{k}.N\mathbf{a} = 2\pi n \tag{1.25}$$

where n is an integer. In terms of reciprocal lattice vectors $\mathbf{K}$, defined by

$$\mathbf{K}_i.\mathbf{a}_j = 2\pi \, \delta_{ij} \tag{1.26}$$

the electronic wavevector assumes the values

$$\mathbf{k} = \frac{n_1}{N_1}\mathbf{K}_1 + \frac{n_2}{N_2}\mathbf{K}_2 + \frac{n_3}{N_3}\mathbf{K}_3. \tag{1.27}$$

Thus the volume of an electronic state in $\mathbf{k}$-space is given by

$$\Delta k_1 \, \Delta k_2 \, \Delta k_3 = (2\pi)^3/V. \tag{1.28}$$

If $\mathbf{q}$ is any vector that satisfies the periodic boundary conditions then the wavefunction can be written generally as an expansion in plane waves:

$$\psi(\mathbf{r}) = \sum_{\mathbf{q}} c_{\mathbf{q}} \exp(i\mathbf{q}.\mathbf{r}). \tag{1.29}$$

This general expansion can be related to the Bloch form by putting $\mathbf{q} = \mathbf{k} - \mathbf{K}$ where $\mathbf{k}$ is not necessarily confined to the first Brillouin zone:

$$\psi_{\mathbf{k}}(\mathbf{r}) = \exp(i\mathbf{k}.\mathbf{r}) \sum_{\mathbf{K}} c_{\mathbf{k}-\mathbf{K}} \exp(-i\mathbf{K}.\mathbf{r}) \tag{1.30}$$

and thus

$$u_\mathbf{k}(\mathbf{r}) = \sum_\mathbf{K} c_{\mathbf{k}-\mathbf{K}} \exp(-i\mathbf{K}.\mathbf{r}).$$ (1.31)

Yet another form for a Bloch function can be formed out of functions $\phi_n(\mathbf{r}-\mathbf{R})$ which are centred at the lattice points $\mathbf{R}$. These are known as Wannier functions. The relation between Bloch and Wannier functions is

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_\mathbf{R} \phi_n(\mathbf{r}-\mathbf{R})\exp(i\mathbf{k}.\mathbf{R}).$$ (1.32)

This is a useful formulation for describing narrow energy bands when the Wannier function can be approximated by atomic orbitals in the tight-binding approximation.

Since the Bloch functions are eigenfunctions of the one-electron Schrödinger equation they are orthogonal to one another, *viz.*

$$\int \psi^*_{n'\mathbf{k}'}\psi_{n\mathbf{k}}\, d\mathbf{r} = \delta_{n'n}\,\delta_{\mathbf{k}'\mathbf{k}}$$ (1.33)

with

$$\psi_{n\mathbf{k}} = \frac{1}{V^{1/2}}\, u_{n\mathbf{k}}(\mathbf{r})\exp(i\mathbf{k}.\mathbf{r}).$$ (1.34)

1.6. Nearly-free-electron model

When the periodic potential is very weak the valence electron is almost free, and hence

$$E_\mathbf{k} \approx \hbar^2 k^2/2m.$$ (1.35)

In the cases of semiconductors with diamond and sphalerite structure there are two atoms in each primitive cell and eight valence electrons. Therefore there have to be four valence bands with two electrons of opposing spin in each state. By allowing $\mathbf{k}$ to extend beyond the first zone, we can work out the total width of the four valence bands by equating it with the Fermi energy E_F for a free-electron gas of the same density as the valence electrons. Observations of soft X-ray emission confirm that the width of the valence band in these semiconductors is indeed close to E_F. Thus it is reasonable to assume that the valence electrons are almost free, and eqn (1.35) is a good approximation to the energy provided we take into account the effect of the lattice.

Restricting $\mathbf{k}$ to the first Brillouin zone (Fig. 1.2) we obtain

$$E_\mathbf{k} \approx \hbar^2 q^2/2m$$ (1.36)

$$\mathbf{q} = \mathbf{k} + \mathbf{K}.$$ (1.37)

The first band is obtained for $\mathbf{K} = (0, 0, 0)2\pi/a$, and is obviously parabolic. At the zone boundary there is an energy gap in general. The second band is obtained from the smallest non-zero reciprocal lattice vectors, which are $\mathbf{K} = (1, 1, 1)2\pi/a$ and its cubic fellows (e.g. $(-1, 1, 1)2\pi/a$) and $\mathbf{K}_2 = (2, 0, 0)2\pi/a$ and its cubic fellows (e.g. $(0, 2, 0)2\pi/a$)). At the zone boundary along the $\langle 100 \rangle$ direction $q = K_2/2 = 2\pi/a$ and $k = -K_2/2 = -2\pi/a$. As q increases k moves towards zero, reaching it when $q = K_2 = 4\pi/a$. At the zone boundary along the $\langle 111 \rangle$ direction $q = K_1/2 = \sqrt{3}\pi/a$ and $k = -K_1/2 = -\sqrt{3}\pi/a$. As q increases, k moves to zero, reaching it when $q = K_1 = 2\sqrt{3}\pi/a$. The band continues to be parabolic in both directions, except close to the zone boundaries.

The first and second bands are parabolic directions because the appropriate reciprocal lattice vector simply subtracts from q. Bands 3 and 4 are not that simple because $\mathbf{K}_1$ and $\mathbf{K}_2$ are neither parallel nor anti-parallel in this case. The region in reciprocal lattice space which contains the first four Brillouin zones is the Jones zone (Fig. 1.3).

Bands 1 and 2 reach the surface of the Jones zone at the points $(2, 0, 0)$ and $(1, 1, 1)$. Bands 3 and 4 are associated with combinations of $\mathbf{k}$, $\mathbf{K}_1$, and $\mathbf{K}_2$ which keep $\mathbf{q}$ close to the zone boundary for all $\mathbf{k}$. The smallest $\mathbf{q}$ corresponds to the centre of a face $\mathbf{q} = \mathbf{K}_1 - \mathbf{K}_2/2$ ($q = 2\sqrt{2}\pi/a$). With $\mathbf{k}$ along the $\langle 100 \rangle$ direction the band is described by $\mathbf{q} \doteq \mathbf{K}_1 - \mathbf{k}$. When $\mathbf{k} = 0$, $\mathbf{q} = \mathbf{K}_1$ ($q = 2\sqrt{3}\pi/a$). Thus q changes by an amount $\sqrt{3} - \sqrt{2}$ in units of $2\pi/a$ as $\mathbf{k}$ sweeps through the zone in the $\langle 100 \rangle$ direction, and hence the energy changes very little with $\mathbf{k}$. This band is far from being free-electron-like. The other band is also flat, for again q changes comparatively little with $\mathbf{k}$ because $\mathbf{k}$ is more or less perpendicular to the reciprocal vector.
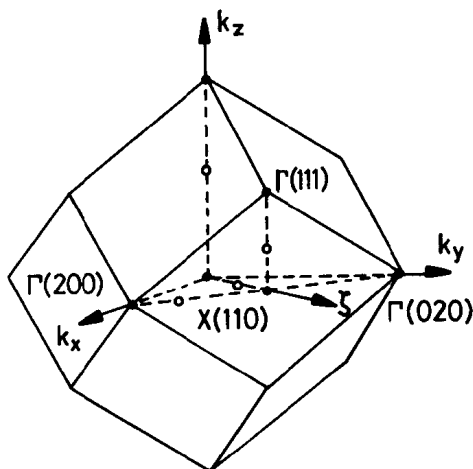

FIG. 1.3. The Jones zone for face-centred cubic crystals containing eight electrons per cell includes the first four Brillouin zones.