

SEVENTH EDITION

Introduction to Solid State Physics

CHARLES KITTEL

CHARLES KITTEL

Introduction
to
Solid State
Physics

SEVENTH EDITION



*John Wiley & Sons, Inc., New York, Chichester,
Brisbane, Toronto, Singapore*

About the Author

Charles Kittel taught solid state physics at Berkeley from 1951 to 1978; earlier he was a member of the solid state group at the Bell Laboratories. His undergraduate work was at M.I.T. and at Cambridge University, followed by graduate work at the University of Wisconsin. He is a member of the National Academy of Science and of the American Academy of Arts and Sciences.

His research in solids began with studies of ferromagnetic, antiferromagnetic, and paramagnetic resonance, along with work on magnetic domains, spin waves, and domain boundaries in ferromagnets and ferroelectrics. His work on the single domain structure of fine particles has had broad application in magnetic recording, geomagnetism, and biomagnetism. Along with collaborators at Berkeley he did the first work on cyclotron resonance in semiconductors, which led to the understanding of the band structure of silicon, germanium, and indium antimonide, together with the theory of their impurity states. He also worked on the interpretation of magnetoplasma resonance in semiconductors and of Alfvén resonance in electron-hole drops in germanium.

The first edition of *ISSP* integrated the elementary aspects of solid state physics for study by seniors and beginning graduate students. Now in its seventh edition, *ISSP* plays the same part for the current generation of students.

Copyright © 1953, 1956, 1966, 1971, 1976, 1986, 1996 by John Wiley & Sons, Inc.

All rights reserved. Published simultaneously in Canada.

Reproduction or translation of any part of this work beyond that permitted by Sections 107 or 108 of the 1976 United States Copyright Act without the permission of the copyright owner is unlawful. Requests for permission or further information should be addressed to the Permissions Department, John Wiley & Sons, Inc.

Library of Congress Cataloging in Publication Data:

Charles Kittel.

Introduction to solid state physics / Charles Kittel. — 7th ed.

p. cm.

Includes index.

ISBN 0-471-11181-3 (cloth : alk. paper)

1. Solid state physics. I. Title.

QC176.K5 1996

530.4'1—dc20

95-18445

CIP

Printed in the United States of America

10987

Preface

This book is the seventh edition of an elementary text on solid state physics for senior and beginning graduate students of physical science and engineering. The book is an update of the sixth edition of 1986 and includes additions, improvements, and corrections made in that edition in 13 successive printings—which it was time to pull together—and a number of new topics besides. Significant advances in the field have been added or discussed more fully: thus high temperature superconductors are treated, and results of scanning tunneling microscopy are displayed; the treatment of fiber optics is expanded. There are discussions, among other topics, of nanostructures, superlattices, Bloch/Wannier levels, Zener tunneling, light-emitting diodes, and new magnetic materials. The additions have been made within a boundary condition intended to keep the text within one volume and at a reasonable price.

The theoretical level of the text itself has not been changed. There is more discussion of useful materials. The treatment of elastic constants and elastic waves which was dropped after the fourth edition has now been returned because, as many have pointed out, the matter is useful and not easily accessible elsewhere. The treatment of superconductors is much more extensive than is usual in a text at this level: either you do it or you don't.

Solid state physics is concerned with the properties, often astonishing and often of great utility, that result from the distribution of electrons in metals, semiconductors, and insulators. The book also tells how the excitations and imperfections of real solids can be understood with simple models whose power and scope are now firmly established. The subject matter supports a profitable interplay of experiment, application, and theory. The book, in English and in many translations, has helped give several generations of students a picture of the process. Students also find the field attractive because of the frequent possibility of working in small groups.

Instructors will use the book as the foundation of a course in their own way, yet there are two general patterns to the introduction, selection and order of the basic material. If students have a significant preparation in elementary quantum mechanics, they will like to begin with the quantum theory of elec-

trons in one-dimensional solids, starting with the free electron gas in Chapter 6 and energy bands in Chapter 7. One will need to treat the reciprocal lattice in three dimensions (Chapter 2) before plunging into semiconductors (Chapter 8) and Fermi surfaces (Chapter 9). Crystal structures, crystal binding, and phonons could be considered as recreational reading. In a more gradual approach, the first eight chapters through the physics of semiconductors are read consecutively as a one-semester introduction to the field.

What about the necessary statistical mechanics? A vague discomfort at the thought of the chemical potential is still characteristic of a physics education. This intellectual gap is due to the obscurity of the writings of J. Willard Gibbs, who discovered and understood the matter 100 years ago. Herbert Kroemer and I have clarified the physics of the chemical potential in the early chapters of our book on thermal physics.

Review series give excellent extended treatments of all the subjects treated in this book and many more besides; thus with good conscience I give few references to original papers. In these omissions no lack of honor is intended to those who first set sail on these seas.

The crystallographic notation conforms with current usage in physics. Important equations are repeated in SI and CGS-Gaussian units, where these differ. Exceptions are figure captions, chapter summaries, some problems, and any long section of text where a single indicated substitution will translate from CGS to SI. Chapter Contents pages discuss conventions adopted to make parallel usage simple. The dual usage in this book has been found useful and acceptable.

Tables are in conventional units. The symbol e denotes the charge on the proton and is positive. The notation (18) refers to Equation (18) of the current chapter, but (3.18) refers to Equation 18 of Chapter 3. A caret $\wedge$ over a vector refers to a unit vector. Few of the problems are exactly easy; most were devised to carry forward the subject of the chapter. With a few exceptions, the problems are those of the original sixth edition.

This edition owes much to the advice of Professor Steven G. Louie. For collected corrections, data, and illustrations I am grateful to P. Allen, M. Beasley, D. Chemla, T.-C. Chiang, M. L. Cohen, M. G. Craford, A. E. Curzon, D. Eigler, L. M. Falicov, R. B. Frankel, J. Friedel, T. H. Geballe, D. M. Ginsberg, C. Herring, H. F. Hess, N. Holonyak, Jr., M. Jacob, J. Mamin, P. McEuen, J. G. Mullen, J. C. Phillips, D. E. Prober, Marta Puebla, D. S. Rokhsar, L. Takacs, Tingye Li, M. A. Van Hove, E. R. Weber, R. M. White, J. P. Wolfe, and A. Zettl. Of the Wiley staff I have particularly great debts to Clifford Mills for publication supervision, to Cathy Donovan for her ingenuity in processing the additions between the thirteen successive printings, and to Suzanne Ingrao of Ingrao Associates for her skill and understanding during the editorial process.

Corrections and suggestions will be gratefully received and may be addressed to the author at the Department of Physics, University of California, Berkeley, CA 94720-7300; by email to kittel@uclink4.Berkeley.edu; and by fax to (510) 643-9473.

C. Kittel

An Instructor's Manual is available for this revision; several problems have been added (to Chapter 3 and Chapter 6); one dropped (from Chapter 4), and several corrections made. Instructors who have adopted the text for classroom use should direct a request on departmental letterhead to John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012. Limited requests for permission to copy figures or other material should be addressed to the Permissions Editor at this address.

8

Semiconductor Crystals

BAND GAP	199
EQUATIONS OF MOTION	203
Physical derivation of $\hbar\mathbf{k} = \mathbf{F}$	205
Holes	206
Effective mass	209
Physical interpretation of the effective mass	210
Effective masses in semiconductors	212
Silicon and germanium	214
INTRINSIC CARRIER CONCENTRATION	216
Intrinsic mobility	220
IMPURITY CONDUCTIVITY	221
Donor states	222
Acceptor states	224
Thermal ionization of donors and acceptors	226
THERMOELECTRIC EFFECTS	227
SEMIMETALS	228
SUPERLATTICES	228
SUMMARY	231
PROBLEMS	231
1. Impurity orbits	231
2. Ionization of donors	231
3. Hall effect with two carrier types	231
4. Cyclotron resonance	232
5. Magnetoresistance with two carrier types	232
REFERENCES	232

NOTE: The discussion of carrier orbits in applied fields is continued in Chapter 9. Amorphous semiconductors are treated in Chapter 17. Junctions and barriers are treated in Chapter 19.

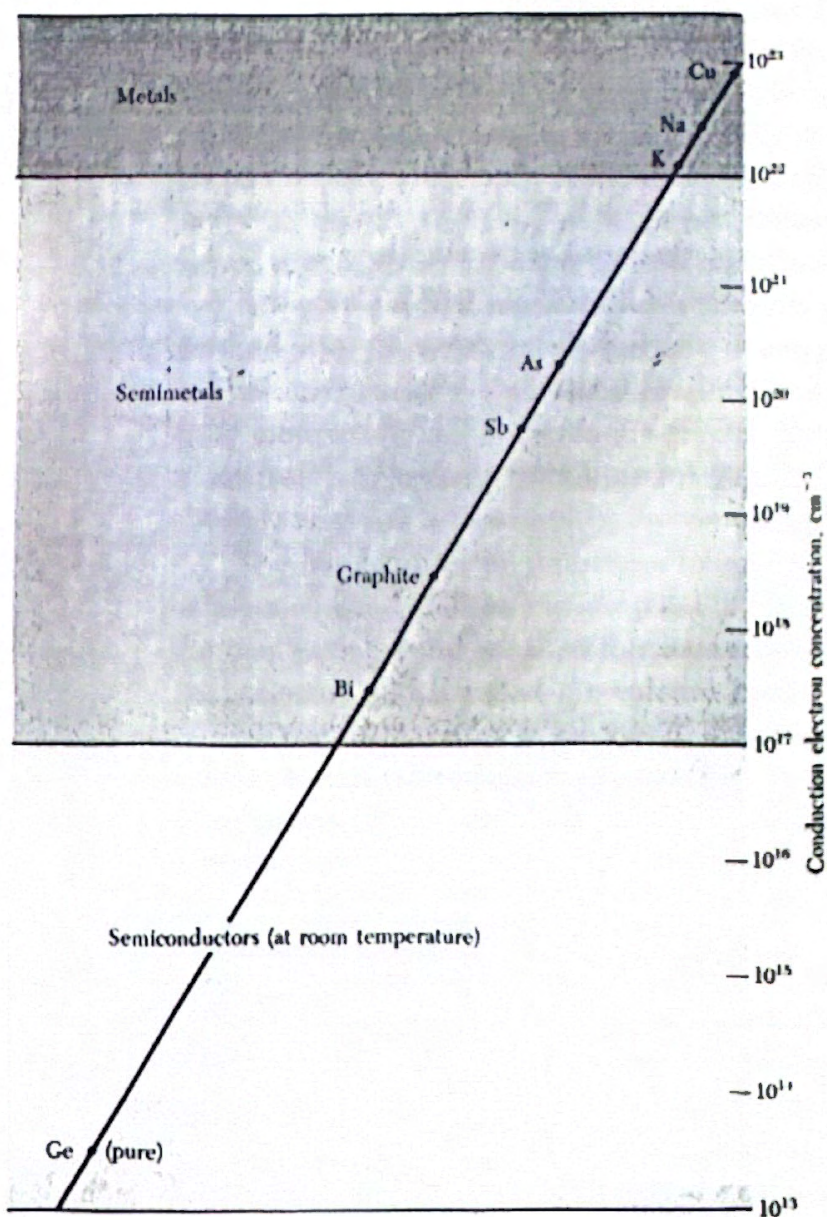


Figure 1 Carrier concentrations for metals, semimetals, and semiconductors. The semiconductor range may be extended upward by increasing the impurity concentration, and the range can be extended downward to merge eventually with the insulator range.

Carrier concentrations representative of metals, semimetals, and semiconductors are shown in Fig. 1. Semiconductors are generally classified by their electrical resistivity at room temperature, with values in the range of 10^{-2} to 10^9 ohm-cm, and strongly dependent on temperature. At absolute zero a pure, perfect crystal of most semiconductors will be an insulator, if we arbitrarily define an insulator as having a resistivity above 10^{14} ohm-cm.

Devices based on semiconductors include transistors, switches, diodes, photovoltaic cells, detectors, and thermistors. These may be used as single circuit elements or as components of integrated circuits. We discuss in this chapter the central physical features of the classical semiconductor crystals, particularly silicon, germanium, and gallium arsenide.

Some useful nomenclature: the semiconductor compounds of chemical formula AB , where A is a trivalent element and B is a pentavalent element, are called III-V (three-five) compounds. Examples are indium antimonide and gallium arsenide. Where A is divalent and B is hexavalent, the compound is called a II-VI compound; examples are zinc sulfide and cadmium sulfide. Silicon and germanium are sometimes called diamond-type semiconductors, because they have the crystal structure of diamond. Diamond itself is more an insulator rather than a semiconductor. Silicon carbide SiC is a IV-IV compound.

A highly purified semiconductor exhibits intrinsic conductivity, as distinguished from the impurity conductivity of less pure specimens. In the intrinsic temperature range the electrical properties of a semiconductor are not essentially modified by impurities in the crystal. An electronic band scheme leading to intrinsic conductivity is indicated in Fig. 2. The conduction band is vacant at absolute zero and is separated by an energy gap E_g from the filled valence band.

The **band gap** is the difference in energy between the lowest point of the conduction band and the highest point of the valence band. The lowest point in the conduction band is called the **conduction band edge**; the highest point in the valence band is called the **valence band edge**.

As the temperature is increased, electrons are thermally excited from the valence band to the conduction band (Fig. 3). Both the electrons in the conduction band and the vacant orbitals or holes left behind in the valence band contribute to the electrical conductivity.

BAND GAP

The intrinsic conductivity and intrinsic carrier concentrations are largely controlled by $E_g/k_B T$, the ratio of the band gap to the temperature. When this ratio is large, the concentration of intrinsic carriers will be low and the conductivity will be low. Band gaps of representative semiconductors are given in Table 1. The best values of the band gap are obtained by optical absorption.

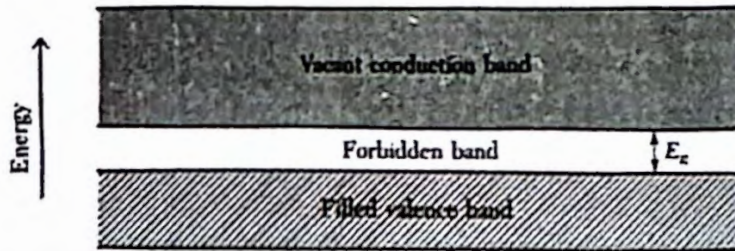
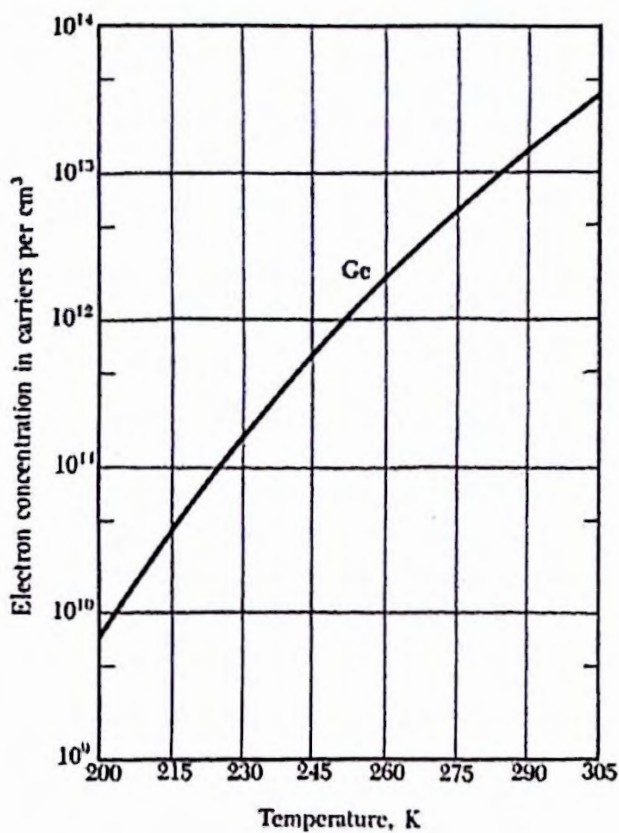
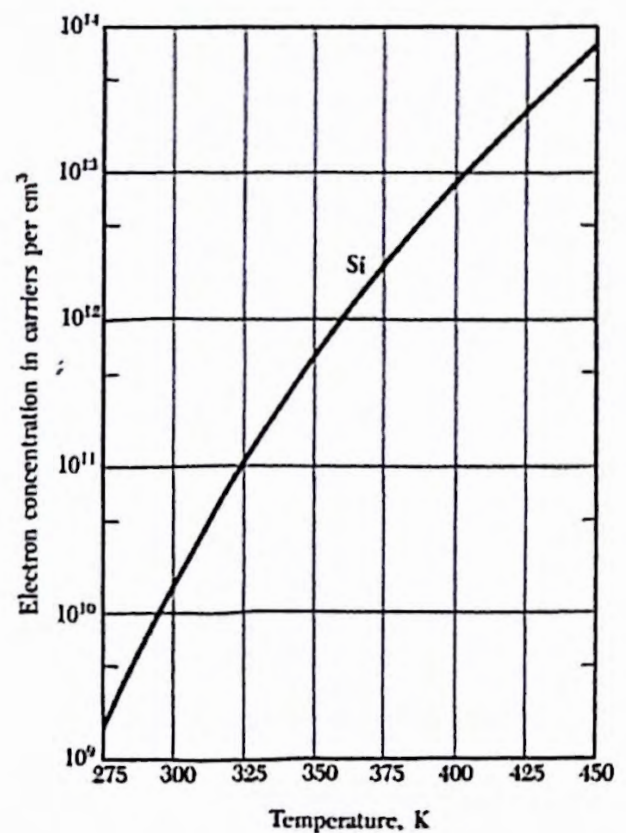


Figure 2 Band scheme for intrinsic conductivity in a semiconductor. At 0 K the conductivity is zero because all states in the valence band are filled and all states in the conduction band are vacant. As the temperature is increased, electrons are thermally excited from the valence band to the conduction band, where they become mobile.



(a)



(b)

Figure 3 Intrinsic electron concentration as a function of temperature for (a) germanium and (b) silicon. Under intrinsic conditions the hole concentration is equal to the electron concentration. The intrinsic concentration at a given temperature is higher in Ge than in Si because the energy gap is narrower in Ge (0.66 eV) than in Si (1.11 eV). (After W. C. Dunlap.)

Table 1 Energy gap between the valence and conduction bands
(*i* = indirect gap; *d* = direct gap)

Crystal	Gap	E_g , eV		Crystal	Gap	E_g , eV	
		0 K	300 K			0 K	300 K
Diamond	<i>i</i>	5.4		HgTe ^a	<i>d</i>	-0.30	
Si	<i>i</i>	1.17	1.11	PbS	<i>d</i>	0.286	0.34–0.37
Ge	<i>i</i>	0.744	0.66	PbSe	<i>i</i>	0.165	0.27
α Sn	<i>d</i>	0.00	0.00	PbTe	<i>i</i>	0.190	0.29
InSb	<i>d</i>	0.23	0.17	CdS	<i>d</i>	2.582	2.42
InAs	<i>d</i>	0.43	0.36	CdSe	<i>d</i>	1.840	1.74
InP	<i>d</i>	1.42	1.27	CdTe	<i>d</i>	1.607	1.44
GaP	<i>i</i>	2.32	2.25	ZnO		3.436	3.2
GaAs	<i>d</i>	1.52	1.43	ZnS		3.91	3.6
GaSb	<i>d</i>	0.81	0.68	SnTe	<i>d</i>	0.3	0.18
AlSb	<i>i</i>	1.65	1.6	AgCl		—	3.2
SiC(hex)	<i>i</i>	3.0	—	AgI		—	2.8
Te	<i>d</i>	0.33	—	Cu ₂ O	<i>d</i>	2.172	—
ZnSb		0.56	0.56	TiO ₂		3.03	—

^aHgTe is a semimetal, the bands overlap.

The threshold of continuous optical absorption at frequency ω_g determines the band gap $E_g = \hbar\omega_g$ in Figs. 4a and 5a. In the **direct absorption process** a photon is absorbed by the crystal with the creation of an electron and a hole.

In the **indirect absorption process** in Figs. 4b and 5b the minimum energy gap of the band structure involves electrons and holes separated by a substantial wavevector $\mathbf{k}_c$. Here a direct photon transition at the energy of the minimum gap cannot satisfy the requirement of conservation of wavevector, because photon wavevectors are negligible at the energy range of interest. But if a phonon of wavevector $\mathbf{K}$ and frequency Ω is created in the process, then we can have

$$\mathbf{k}(\text{photon}) = \mathbf{k}_c + \mathbf{K} \cong 0; \quad \hbar\omega = E_g + \hbar\Omega,$$

as required by the conservation laws. The phonon energy $\hbar\Omega$ will generally be much less than E_g ; a phonon even of high wavevector is an easily accessible source of crystal momentum because the phonon energies are characteristically small (~ 0.01 to 0.03 eV) in comparison with the energy gap. If the temperature is high enough that the necessary phonon is already thermally excited in the crystal, it is possible also to have a photon absorption process in which the phonon is absorbed.

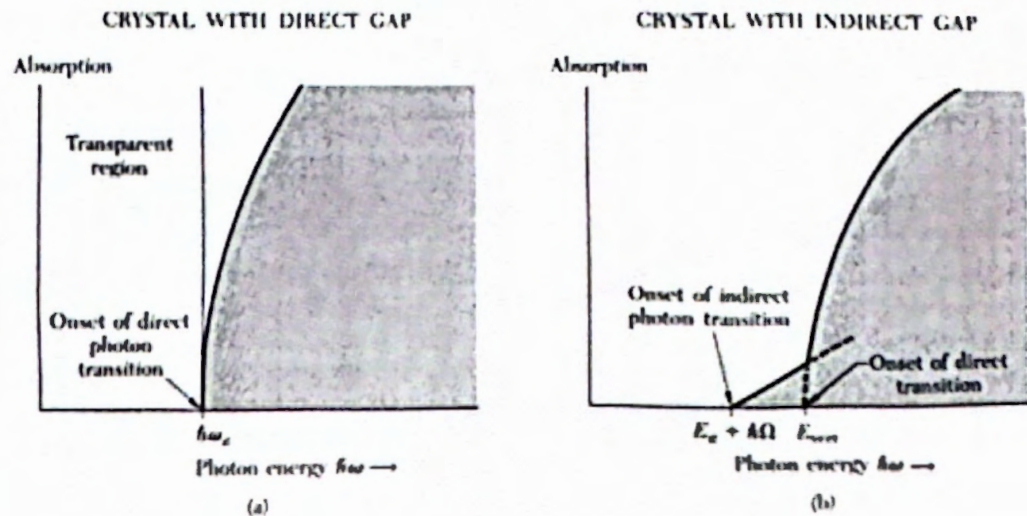


Figure 4 Optical absorption in pure insulators at absolute zero. In (a) the threshold determines the energy gap as $E_g = \hbar\omega_g$. In (b) the optical absorption is weaker near the threshold; at $\hbar\omega = E_g + \hbar\Omega$ a photon is absorbed with the creation of three particles: a free electron, a free hole, and a phonon of energy $\hbar\Omega$. In (b) the energy E_{vert} marks the threshold for the creation of a free electron and a free hole, with no phonon involved. Such a transition is called vertical; it is similar to the direct transition in (a). These plots do not show absorption lines that sometimes are seen lying just to the low energy side of the threshold. Such lines are due to the creation of a bound electron-hole pair, called an exciton.

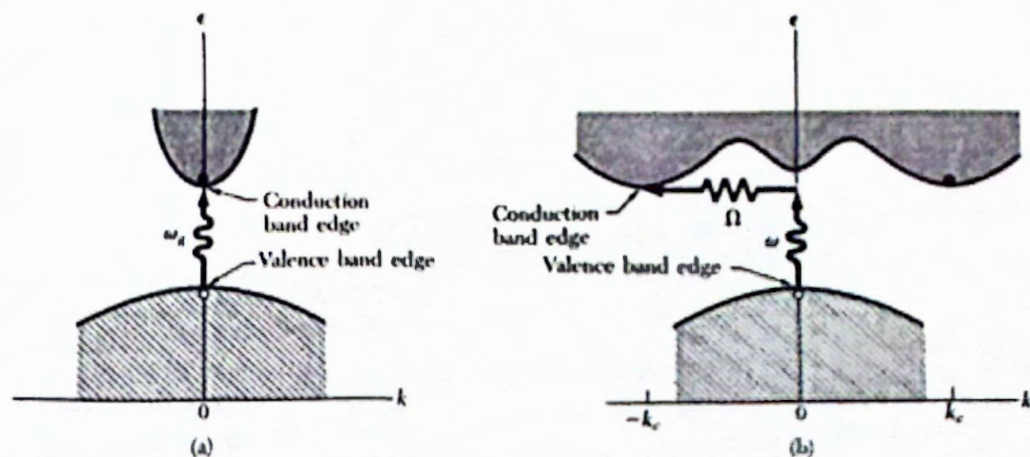


Figure 5 In (a) the lowest point of the conduction band occurs at the same value of k as the highest point of the valence band. A direct optical transition is drawn vertically with no significant change of k , because the absorbed photon has a very small wavevector. The threshold frequency ω_g for absorption by the direct transition determines the energy gap $E_g = \hbar\omega_g$. The indirect transition in (b) involves both a photon and a phonon because the band edges of the conduction and valence bands are widely separated in k space. The threshold energy for the indirect process in (b) is greater than the true band gap. The absorption threshold for the indirect transition between the band edges is at $\hbar\omega = E_g + \hbar\Omega$, where Ω is the frequency of an emitted phonon of wavevector $K \approx -k_c$. At higher temperatures phonons are already present; if a phonon is absorbed along with a photon, the threshold energy is $\hbar\omega = E_g - \hbar\Omega$. *Note:* The figure shows only the threshold transitions. Transitions occur generally between almost all points of the two bands for which the wavevectors and energy can be conserved.

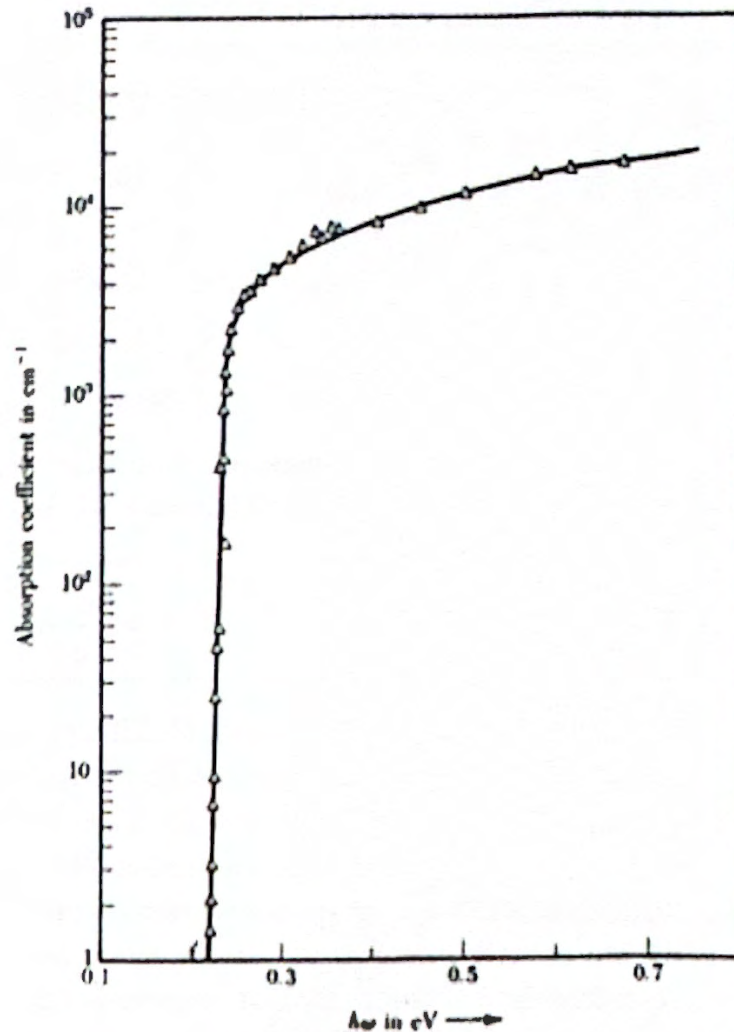


Figure 6 Optical absorption in pure indium antimonide, InSb. The transition is direct because both conduction and valence band edges are at the center of the Brillouin zone, $k = 0$. Notice the sharp threshold. (After G. W. Gobeli and H. Y. Fan.)

The band gap may also be deduced from the temperature dependence of the conductivity or of the carrier concentration in the intrinsic range. The carrier concentration is obtained from measurements of the Hall voltage (Chapter 6), sometimes supplemented by conductivity measurements. Optical measurements determine whether the gap is direct or indirect. The band edges in Ge and in Si are connected by indirect transitions; the band edges in InSb are connected by a direct transition (Fig. 6). The gap in α Sn is direct and is exactly zero; HgTe and HgSe are semimetals and have negative gaps—the bands overlap.

EQUATIONS OF MOTION

We derive the equation of motion of an electron in an energy band. We look at the motion of a wave packet in an applied electric field. Suppose that the wave packet is made up of wavefunctions near a particular wavevector k . The

group velocity is $v_g = d\omega/dk$. The frequency associated with a wavefunction of energy ϵ is $\omega = \epsilon/\hbar$, and so

$$v_g = \hbar^{-1} d\epsilon/dk \quad \text{or} \quad \mathbf{v} = \hbar^{-1} \nabla_{\mathbf{k}} \epsilon(\mathbf{k}) . \quad (1)$$

The effects of the crystal on the electron motion are contained in the dispersion relation $\epsilon(\mathbf{k})$.

The work $\delta\epsilon$ done on the electron by the electric field E in the interval δt is

$$\delta\epsilon = -eEv_g \delta t . \quad (2)$$

We observe that

$$\delta\epsilon = (d\epsilon/dk)\delta k = \hbar v_g \delta k , \quad (3)$$

using (1). On comparing (2) with (3) we have

$$\delta k = -(eE/\hbar)\delta t , \quad (4)$$

whence $\hbar dk/dt = -eE$, the same relation as for free electrons.

We may write (4) in terms of the external force F as

$$\boxed{\hbar \frac{d\mathbf{k}}{dt} = \mathbf{F} .} \quad (5)$$

This is an important relation: in a crystal $\hbar d\mathbf{k}/dt$ is equal to the external force on the electron. In free space $d(m\mathbf{v})/dt$ is equal to the force. We have not overthrown Newton's second law of motion: the electron in the crystal is subject to forces from the crystal lattice as well as from external sources.

The force term in (5) also includes the Lorentz force on an electron in a magnetic field, under ordinary conditions where the magnetic field is not so strong that it breaks down the band structure. Thus the equation of motion of an electron of group velocity $\mathbf{v}$ in a constant magnetic field $\mathbf{B}$ is

$$\text{(CGS)} \quad \hbar \frac{d\mathbf{k}}{dt} = -\frac{e}{c} \mathbf{v} \times \mathbf{B} ; \quad \text{(SI)} \quad \hbar \frac{d\mathbf{k}}{dt} = -e\mathbf{v} \times \mathbf{B} , \quad (6)$$

where the right-hand side is the Lorentz force on the electron. With the group velocity $\hbar\mathbf{v} = \nabla_{\mathbf{k}}\epsilon$, the rate of change of the wavevector is

$$\text{(CGS)} \quad \frac{d\mathbf{k}}{dt} = -\frac{e}{\hbar^2 c} \nabla_{\mathbf{k}} \epsilon \times \mathbf{B} ; \quad \text{(SI)} \quad \frac{d\mathbf{k}}{dt} = -\frac{e}{\hbar^2} \nabla_{\mathbf{k}} \epsilon \times \mathbf{B} , \quad (7)$$

where now both sides of the equation refer to the coordinates in $\mathbf{k}$ space.

We see from the vector cross-product in (7) that in a magnetic field an electron moves in $\mathbf{k}$ space in a direction normal to the direction of the gradient of the energy ϵ , so that the electron moves on a surface of constant energy.

The value of the projection k_B of $\mathbf{k}$ on $\mathbf{B}$ is constant during the motion. The motion in $\mathbf{k}$ space is on a plane normal to the direction of $\mathbf{B}$, and the orbit is defined by the intersection of this plane with a surface of constant energy.

Physical Derivation of $\hbar\mathbf{k} = \mathbf{F}$

We consider the Bloch eigenfunction $\psi_{\mathbf{k}}$ belonging to the energy eigenvalue $\epsilon_{\mathbf{k}}$ and wavevector $\mathbf{k}$:

$$\psi_{\mathbf{k}} = \sum_{\mathbf{G}} C(\mathbf{k} + \mathbf{G}) \exp[i(\mathbf{k} + \mathbf{G}) \cdot \mathbf{r}] . \quad (8)$$

The expectation value of the momentum of an electron in the state $\mathbf{k}$ is

$$\mathbf{p}_{el} = \langle \mathbf{k} | -i\hbar \nabla | \mathbf{k} \rangle = \sum_{\mathbf{G}} \hbar(\mathbf{k} + \mathbf{G}) |C(\mathbf{k} + \mathbf{G})|^2 = \hbar(\mathbf{k} + \sum_{\mathbf{G}} \mathbf{G} |C(\mathbf{k} + \mathbf{G})|^2) , \quad (9)$$

using $\sum_{\mathbf{G}} |C(\mathbf{k} + \mathbf{G})|^2 = 1$.

We examine the transfer of momentum between the electron and the lattice when the state $\mathbf{k}$ of the electron is changed to $\mathbf{k} + \Delta\mathbf{k}$ by the application of an external force. We imagine an insulating crystal electrostatically neutral except for a single electron in the state $\mathbf{k}$ of an otherwise empty band.

We suppose that a weak external force is applied for a time interval such that the total impulse given to the entire crystal system is $\mathbf{J} = \int \mathbf{F} dt$. If the conduction electron were free ($m^* = m$), the total momentum imparted to the crystal system by the impulse would appear in the change of momentum of the conduction electron:

$$\mathbf{J} = \Delta\mathbf{p}_{tot} = \Delta\mathbf{p}_{el} = \hbar\Delta\mathbf{k} . \quad (10)$$

The neutral crystal suffers no net interaction with the electric field, either directly or indirectly through the free electron.

If the conduction electron interacts with the periodic potential of the crystal lattice, we must have

$$\mathbf{J} = \Delta\mathbf{p}_{tot} = \Delta\mathbf{p}_{lat} + \Delta\mathbf{p}_{el} . \quad (11)$$

From the result (9) for $\mathbf{p}_{el}$ we have

$$\Delta\mathbf{p}_{el} = \hbar\Delta\mathbf{k} + \sum_{\mathbf{G}} \hbar\mathbf{G} [(\nabla_{\mathbf{k}} |C(\mathbf{k} + \mathbf{G})|^2) \cdot \Delta\mathbf{k}] . \quad (12)$$

The change $\Delta\mathbf{p}_{lat}$ in the lattice momentum resulting from the change of state of the electron may be derived by an elementary physical consideration. An electron reflected by the lattice transfers momentum to the lattice. If an incident electron with plane wave component of momentum $\hbar\mathbf{k}$ is reflected

with momentum $\hbar(\mathbf{k} + \mathbf{G})$, the lattice acquires the momentum $-\hbar\mathbf{G}$, as required by momentum conservation. The momentum transfer to the lattice when the state $\psi_{\mathbf{k}}$ goes over to $\psi_{\mathbf{k}+\Delta\mathbf{k}}$ is

$$\Delta p_{\text{lat}} = -\hbar \sum_{\mathbf{G}} \mathbf{G} [(\nabla_{\mathbf{k}} |C(\mathbf{k} + \mathbf{G})|^2 \cdot \Delta\mathbf{k}] , \quad (13)$$

as the portion

$$\nabla_{\mathbf{k}} |C(\mathbf{k} + \mathbf{G})|^2 \cdot \Delta\mathbf{k} \quad (14)$$

of each individual component of the initial state is reflected during the state change $\Delta\mathbf{k}$.

The total momentum change is therefore

$$\Delta p_{\text{el}} + \Delta p_{\text{lat}} = \mathbf{J} = \hbar \Delta\mathbf{k} , \quad (15)$$

exactly as for free electrons, Eq. (10). Thus from the definition of $\mathbf{J}$, we have

$$\hbar d\mathbf{k}/dt = \mathbf{F} , \quad (16)$$

derived in (5) by a different method. A rigorous derivation of (16) by an entirely different method is given in Appendix E.

Holes

The properties of vacant orbitals in an otherwise filled band are important in semiconductor physics and in solid state electronics. Vacant orbitals in a band are commonly called holes. A hole acts in applied electric and magnetic fields as if it has a positive charge $+e$. The reason is given in five steps in the boxes that follow.

$$1. \quad \mathbf{k}_h = -\mathbf{k}_e . \quad (17)$$

The total wavevector of the electrons in a filled band is zero: $\Sigma \mathbf{k} = 0$. This result follows from the geometrical symmetry of the Brillouin zone: every fundamental lattice type has symmetry under the inversion operation $\mathbf{r} \rightarrow -\mathbf{r}$ about any lattice point; it follows that the Brillouin zone of the lattice also has inversion symmetry. If the band is filled all pairs of orbitals $\mathbf{k}$ and $-\mathbf{k}$ are filled, and the total wavevector is zero.

If an electron is missing from an orbital of wavevector $\mathbf{k}_e$, the total wavevector of the system is $-\mathbf{k}_e$ and is attributed to the hole. This result is surprising: the electron is missing from $\mathbf{k}_e$ and the position of the hole is usually indicated graphically as situated at $\mathbf{k}_e$, as in Fig. 7. But the true wavevector $\mathbf{k}_h$ of the hole is $-\mathbf{k}_e$, which is the wavevector of the point $\mathbf{G}$ if the hole is at $\mathbf{E}$. The wavevector $-\mathbf{k}_e$ enters into selection rules for photon absorption.

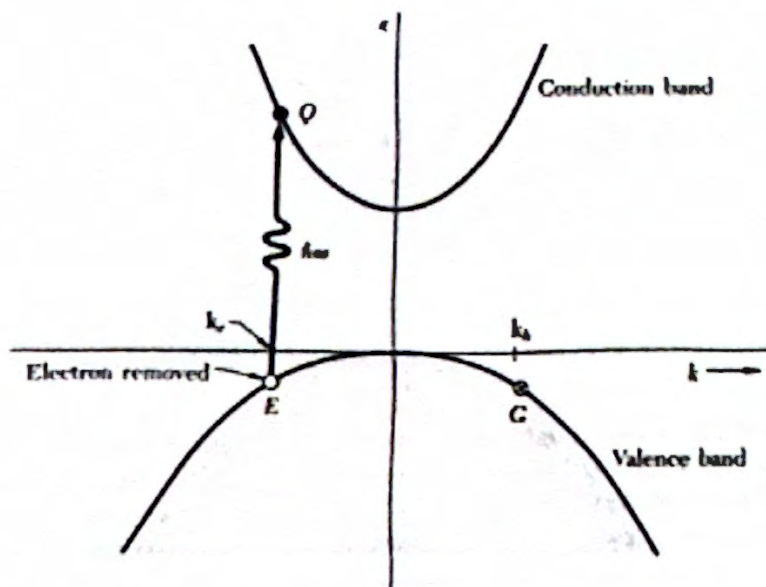


Figure 7 Absorption of a photon of energy $\hbar\omega$ and negligible wavevector takes an electron from E to the filled valence band to Q in the conduction band. If k_e was the wavevector of the electron at E , it becomes the wavevector of the electron at Q . The total wavevector of the valence band after the absorption is $-k_e$, and this is the wavevector we must ascribe to the hole if we describe the valence band as occupied by one hole. Thus $k_h = -k_e$; the wavevector of the hole is the same as the wavevector of the electron which remains at G . For the entire system the total wavevector after the absorption of the photon is $k_e + k_h = 0$, so that the total wavevector is unchanged by the absorption of the photon and the creation of a free electron and free hole.

The hole is an alternate description of a band with one missing electron, and we either say that the hole has wavevector $-k_e$ or that the band with one missing electron has total wavevector $-k_e$.

$$\epsilon_h(k_h) = -\epsilon_e(k_e) \quad (18)$$

Let the zero of energy of the valence band be at the top of the band. The lower in the band the missing electron lies, the higher the energy of the system. The energy of the hole is opposite in sign to the energy of the missing electron, because it takes more work to remove an electron from a low orbital than from a high orbital. Thus if the band is symmetric,¹ $\epsilon_e(k_e) = \epsilon_e(-k_e) = -\epsilon_h(-k_e) = -\epsilon_h(k_h)$. We construct in Fig. 8 a band scheme to represent the properties of a hole. This hole band is a helpful representation because it appears right side up.

¹Bands are always symmetric under the inversion $k \rightarrow -k$ if the spin-orbit interaction is neglected. Even with spin-orbit interaction, bands are always symmetric if the crystal structure permits the inversion operation. Without a center of symmetry, but with spin-orbit interaction, the bands are not symmetric if we compare subbands for which the spin direction is reversed: $\epsilon(\mathbf{k}, \uparrow) \neq \epsilon(-\mathbf{k}, \downarrow)$. See QTS, Chapter 9.

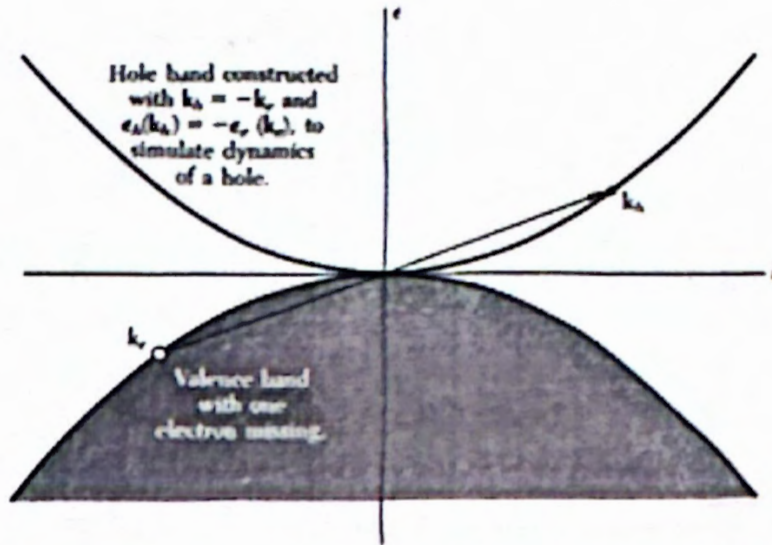


Figure 8 The upper half of the figure shows the hole band that simulates the dynamics of a hole, constructed by inversion of the valence band in the origin. The wavevector and energy of the hole are equal, but opposite in sign, to the wavevector and energy of the empty electron orbital in the valence band. We do not show the disposition of the electron removed from the valence band at k_e .

$$3. \quad \mathbf{v}_h = \mathbf{v}_e . \quad (19)$$

The velocity of the hole is equal to the velocity of the missing electron. From Fig. 8 we see that $\nabla \epsilon_h(k_h) = \nabla \epsilon_e(k_e)$, so that $\mathbf{v}_h(k_h) = \mathbf{v}_e(k_e)$.

$$4. \quad m_h = -m_e . \quad (20)$$

We show below that the effective mass is inversely proportional to the curvature $d^2\epsilon/dk^2$, and for the hole band this has the opposite sign to that for an electron in the valence band. Near the top of the valence band m_e is negative, so that m_h is positive.

$$5. \quad \hbar \frac{d\mathbf{k}_h}{dt} = e(\mathbf{E} + \frac{1}{c} \mathbf{v}_h \times \mathbf{B}) . \quad (21)$$

This comes from the equation of motion

$$(CGS) \quad \hbar \frac{d\mathbf{k}_e}{dt} = -e(\mathbf{E} + \frac{1}{c} \mathbf{v}_e \times \mathbf{B}) \quad (22)$$

that applies to the missing electron when we substitute $-\mathbf{k}_h$ for $\mathbf{k}_e$ and $\mathbf{v}_h$ for $\mathbf{v}_e$. The equation of motion for a hole is that of a particle of positive charge e . The positive charge is consistent with the electric current carried by the valence band of Fig. 9: the current is carried by the unpaired electron in the orbital G:

$$\mathbf{j} = (-e)\mathbf{v}(G) = (-e)[-v(E)] = e\mathbf{v}(E) , \quad (23)$$

which is just the current of a positive charge moving with the velocity ascribed to the missing electron at E . The current is shown in Fig. 10.

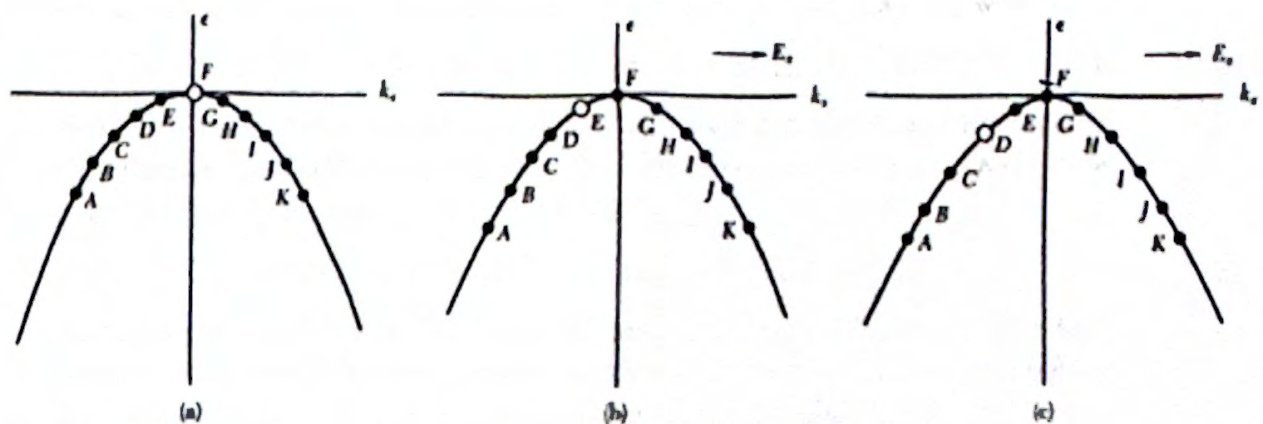


Figure 9 (a) At $t = 0$ all states are filled except F at the top of the band; the velocity v_x is zero at F because $d\epsilon/dk_x = 0$. (b) An electric field E_x is applied in the $+x$ direction. The force on the electrons is in the $-k_x$ direction and all electrons make transitions together in the $-k_x$ direction, moving the hole to the state E. (c) After a further interval the electrons move farther along in k space and the hole is now at D.

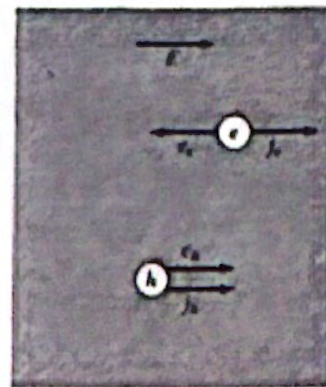


Figure 10 Motion of electrons in the conduction band and holes in the valence band in the electric field E . The hole and electron drift velocities are in opposite directions, but their electric currents are in the same direction, the direction of the electric field.

Effective Mass

When we look at the energy-wavevector relation $\epsilon = (\hbar^2/2m)k^2$ for free electrons, we see that the coefficient of k^2 determines the curvature of ϵ versus k . Turned about, we can say that $1/m$, the reciprocal mass, determines the curvature. For electrons in a band there can be regions of unusually high curvature near the band gap at the zone boundary, as we see from the solutions of the wave equation near the zone boundary. If the energy gap is small in comparison with the free electron energy λ at the boundary, the curvature is enhanced by the factor λ/E_g and the reciprocal mass is enhanced by the same factor.

In semiconductors the band width, which is like the free electron energy, is of the order of 20 eV, while the band gap is of the order of 0.2 to 2 eV. Thus the reciprocal mass is enhanced by a factor 10 to 100, and the effective mass is reduced to 0.1–0.01 of the free electron mass. These values apply near the band gap; as we go away from the gap the curvatures are likely to approach those of free electrons.

To summarize the solutions of Chapter 7 for U positive, an electron near the lower edge of the second band has an energy that may be written as

$$\epsilon(K) = \epsilon_c + (\hbar^2/2m_e)K^2 ; \quad m_e/m = 1/[(2\lambda/U) - 1] . \quad (24)$$

Here K is the wavevector measured from the zone boundary, and m_e denotes the effective mass of the electron near the edge of the second band. An electron near the top of the first band has the energy

$$\epsilon(K) = \epsilon_c - (\hbar^2/2m_h)K^2 ; \quad m_h/m = 1/[(2\lambda/U) + 1] . \quad (25)$$

The curvature and hence the mass will be negative near the top of the first band, but we have introduced a minus sign into (25) in order that the symbol m_h for the hole mass will have a positive value—see (20) above.

The crystal does not weigh any less if the effective mass of a carrier is less than the free electron mass, nor is Newton's second law violated for the crystal *taken as a whole*, ions plus carriers. The important point is that an electron in a periodic potential is accelerated relative to the lattice in an applied electric or magnetic field as if the mass of the electron were equal to an effective mass which we now define.

We differentiate the result (1) for the group velocity to obtain

$$\frac{dv_g}{dt} = \hbar^{-1} \frac{d^2\epsilon}{dk dt} = \hbar^{-1} \left(\frac{d^2\epsilon}{dk^2} \frac{dk}{dt} \right) \quad (26)$$

We know from (5) that $dk/dt = F/\hbar$, whence

$$\frac{dv_g}{dt} = \left(\frac{1}{\hbar^2} \frac{d^2\epsilon}{dk^2} \right) F ; \quad \text{or} \quad F = \frac{\hbar^2}{d^2\epsilon/dk^2} \frac{dv_g}{dt} . \quad (27)$$

If we identify $\hbar^2/(d^2\epsilon/dk^2)$ as a mass, then (27) assumes the form of Newton's second law. We define the effective mass m^* by

$$\boxed{\frac{1}{m^*} = \frac{1}{\hbar^2} \frac{d^2\epsilon}{dk^2}} . \quad (28)$$

It is easy to generalize this to take account of an anisotropic energy surface, as for electrons in Si or Ge. We introduce the components of the reciprocal effective mass tensor

$$\left(\frac{1}{m^*} \right)_{\mu\nu} = \frac{1}{\hbar^2} \frac{d^2\epsilon_k}{dk_\mu dk_\nu} ; \quad \frac{dv_\mu}{dt} = \left(\frac{1}{m^*} \right)_{\mu\nu} F_\nu , \quad (29)$$

where μ, ν are Cartesian coordinates.

Physical Interpretation of the Effective Mass

How can an electron of mass m when put into a crystal respond to applied fields as if the mass were m^* ? It is helpful to think of the process of Bragg reflection of electron waves in a lattice. Consider the weak interaction approxi-

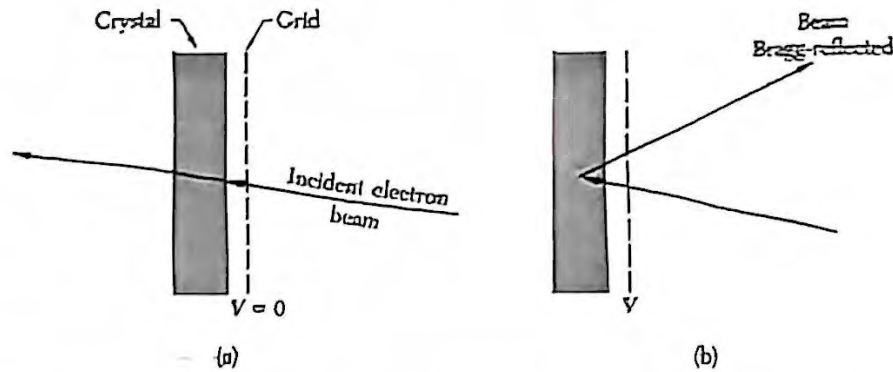


Figure 11 Explanation of negative effective masses which occur near a Brillouin zone boundary. In (a) the energy of the electron beam incident on a thin crystal is slightly too low to satisfy the condition for Bragg reflection and the beam is transmitted through the crystal. The application of a small voltage across the grid may, as in (b), cause the Bragg condition to be satisfied, and the electron beam will then be reflected from the appropriate set of crystal planes.

mation treated in Chapter 7. Near the bottom of the lower band the orbital is represented quite adequately by a plane wave $\exp(ikx)$ with momentum $\hbar k$; the wave component $\exp[i(k - G)x]$ with momentum $\hbar(k - G)$ is small and increases only slowly as k is increased, and in this region $m^* = m$. An increase in the reflected component $\exp[i(k - G)x]$ as k is increased represents momentum transfer to the electron from the lattice.

Near the boundary the reflected component is quite large; at the boundary it becomes equal in amplitude to the forward component, at which point the eigenfunctions are standing waves, rather than running waves. Here the momentum component $\hbar(-\frac{1}{2}G)$ cancels the momentum component $\hbar(\frac{1}{2}G)$.

It is not surprising to find negative values for m^* just below a zone boundary. A single electron in an energy band may have positive or negative effective mass: the states of positive effective mass occur near the bottom of a band because positive effective mass means that the band has upward curvature ($d^2\epsilon/dk^2$ is positive). States of negative effective mass occur near the top of the band. A negative effective mass means that on going from state k to state $k + \Delta k$, the momentum transfer to the lattice from the electron is larger than the momentum transfer from the applied force to the electron. Although k is increased by Δk by the applied electric field, the approach to Bragg reflection can give an overall decrease in the forward momentum of the electron; when this happens the effective mass is negative (Fig. 11).

As we proceed in the second band away from the boundary, the amplitude of $\exp[i(k - G)x]$ decreases rapidly and m^* assumes a small positive value. Here the increase in electron velocity resulting from a given external impulse is larger than that which a free electron would experience. The lattice makes up the difference through the reduced recoil it experiences when the amplitude of $\exp[i(k - G)x]$ is diminished.

If the energy in a band depends only slightly on k , then the effective mass will be very large. That is, $m^*/m \gg 1$ when $d^2\epsilon/dk^2$ is very small. The tight-binding approximation discussed in Chapter 9 gives quick insight into the formation of narrow bands. If the wavefunctions centered on neighboring atoms overlap very little, then the overlap integral will be small; the width of the band narrow; and the effective mass large.

The overlap of wavefunctions centered on neighboring atoms is small for the inner or core electrons. The 4f electrons of the rare earth metals, for example, overlap very little. The overlap integral determines the rate of quantum tunneling of an electron from one ion to another. When the effective mass is large, the electron tunnels slowly from one ion to an adjacent ion in the lattice.

Effective Masses in Semiconductors

In many semiconductors it has been possible to determine by cyclotron resonance the form of the energy surfaces of the conduction and valence bands near the band edges. The determination of the energy surface is equivalent to a determination of the effective mass tensor (29). Cyclotron resonance in a semiconductor is carried out with centimeter wave or millimeter wave radiation at low carrier concentration.

The current carriers are accelerated in helical orbits about the axis of a static magnetic field. The angular rotation frequency ω_c is

$$\text{(CGS)} \quad \omega_c = \frac{eB}{m^*c}, \quad \text{(SI)} \quad \omega_c = \frac{eB}{m^*}, \quad (30)$$

where m^* is the appropriate cyclotron effective mass. Resonant absorption of energy from an rf electric field perpendicular to the static magnetic field (Fig. 12) occurs when the rf frequency is equal to the cyclotron frequency. Holes and electrons rotate in opposite senses in a magnetic field.

We consider the experiment for $m^*/m \approx 0.1$. At $f_c = 24$ GHz, or $\omega_c = 1.5 \times 10^{11} \text{ s}^{-1}$, we have $B \approx 860$ G at resonance. The line width is determined by the collision relaxation time τ , and to obtain a distinctive resonance it is necessary that $\omega_c\tau \geq 1$. The mean free path must be long enough to permit the average carrier to get one radian around a circle between collisions. The requirements are met with the use of higher frequency radiation and higher magnetic fields, with high purity crystals in liquid helium.

In direct-gap semiconductors with band edges at the center of the Brillouin zone, the bands have the structure shown in Fig. 13. The conduction band edge is spherical with the effective mass m_e :

$$\epsilon_c = E_g + \hbar^2 k^2 / 2m_e, \quad (31)$$

referred to the valence band edge. The valence bands are characteristically

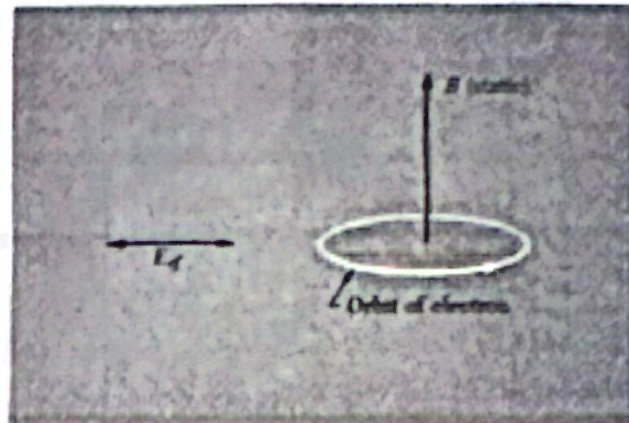


Figure 12 Arrangement of fields in a cyclotron resonance experiment in a semiconductor. The sense of the circulation is opposite for electrons and holes.

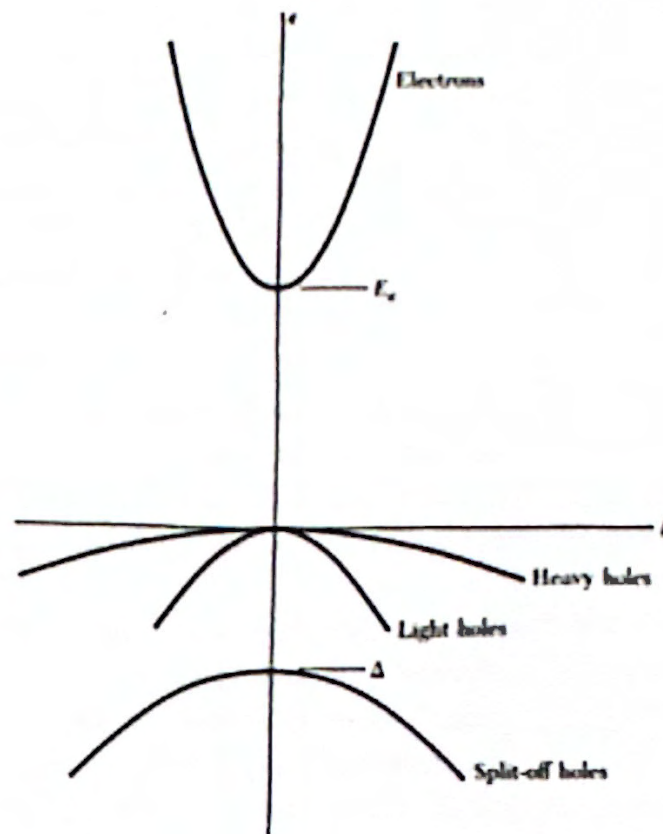


Figure 13 Simplified view of the band edge structure of a direct gap semiconductor.

Table 2 Effective masses of electrons and holes in direct gap semiconductors

Crystal	Electron m_e/m	Heavy hole m_{hh}/m	Light hole m_{lh}/m	Split-off hole m_{soh}/m	Spin-orbit Δ , eV
InSb	0.015	0.39	0.021	(0.11)	0.82
InAs	0.026	0.41	0.025	0.08	0.43
InP	0.073	0.4	(0.078)	(0.15)	0.11
GaSb	0.047	0.3	0.06	(0.14)	0.80
GaAs	0.066	0.5	0.082	0.17	0.34
Cu ₂ O	0.99	—	0.58	0.69	0.13

threefold near the edge, with the heavy hole hh and light hole lh bands degenerate at the center, and a band soh split off by the spin-orbit splitting Δ :

$$\begin{aligned}\epsilon_v(hh) &\cong -\hbar^2 k^2 / 2m_{hh} ; & \epsilon_v(lh) &\cong -\hbar^2 k^2 / 2m_{lh} ; \\ \epsilon_v(soh) &\cong -\Delta - \hbar^2 k^2 / 2m_{soh}\end{aligned}\quad (32)$$

Values of the mass parameters are given in Table 2. The forms (32) are only approximate, because even close to $k = 0$ the heavy and light hole bands are not spherical—see the discussion below for Ge and Si.

The perturbation theory of band edges (Problem 9.8) suggests that the electron effective mass should be proportional to the band gap, approximately, for a direct gap crystal. We use Tables 1 and 2 to find the fairly constant values $m_e/(mE_g) = 0.063$, 0.060, and 0.051 in $(\text{eV})^{-1}$ for the series InSb, InAs, and InP, in agreement with this suggestion.

Silicon and Germanium

The conduction and valence bands of germanium are shown in Fig. 14, based on a combination of theoretical and experimental results. The valence band edge in both crystals is at $k = 0$ and is derived from $p_{3/2}$ and $p_{1/2}$ states of the free atoms, as is clear from the tight-binding approximation (Chapter 9) to the wavefunctions.

The $p_{3/2}$ level is fourfold degenerate as in the atom; the four states correspond to m_j values $\pm \frac{3}{2}$ and $\pm \frac{1}{2}$. The $p_{1/2}$ level is doubly degenerate, with $m_j = \pm \frac{1}{2}$. The $p_{3/2}$ states are higher in energy than the $p_{1/2}$ states; the energy difference Δ is a measure of the spin-orbit interaction.

The valence band edges are not simple. Holes near the band edge are characterized by two effective masses, light and heavy. These arise from the two bands formed from the $p_{3/2}$ level of the atom. There is also a band formed from the $p_{1/2}$ level, split off from the $p_{3/2}$ level by the spin-orbit interaction. The energy surfaces are not spherical, but warped (QTS, p. 271):

$$\epsilon(k) = Ak^2 \pm [B^2k^4 + C^2(k_x^2k_y^2 + k_y^2k_z^2 + k_z^2k_x^2)]^{1/2} \quad (33)$$

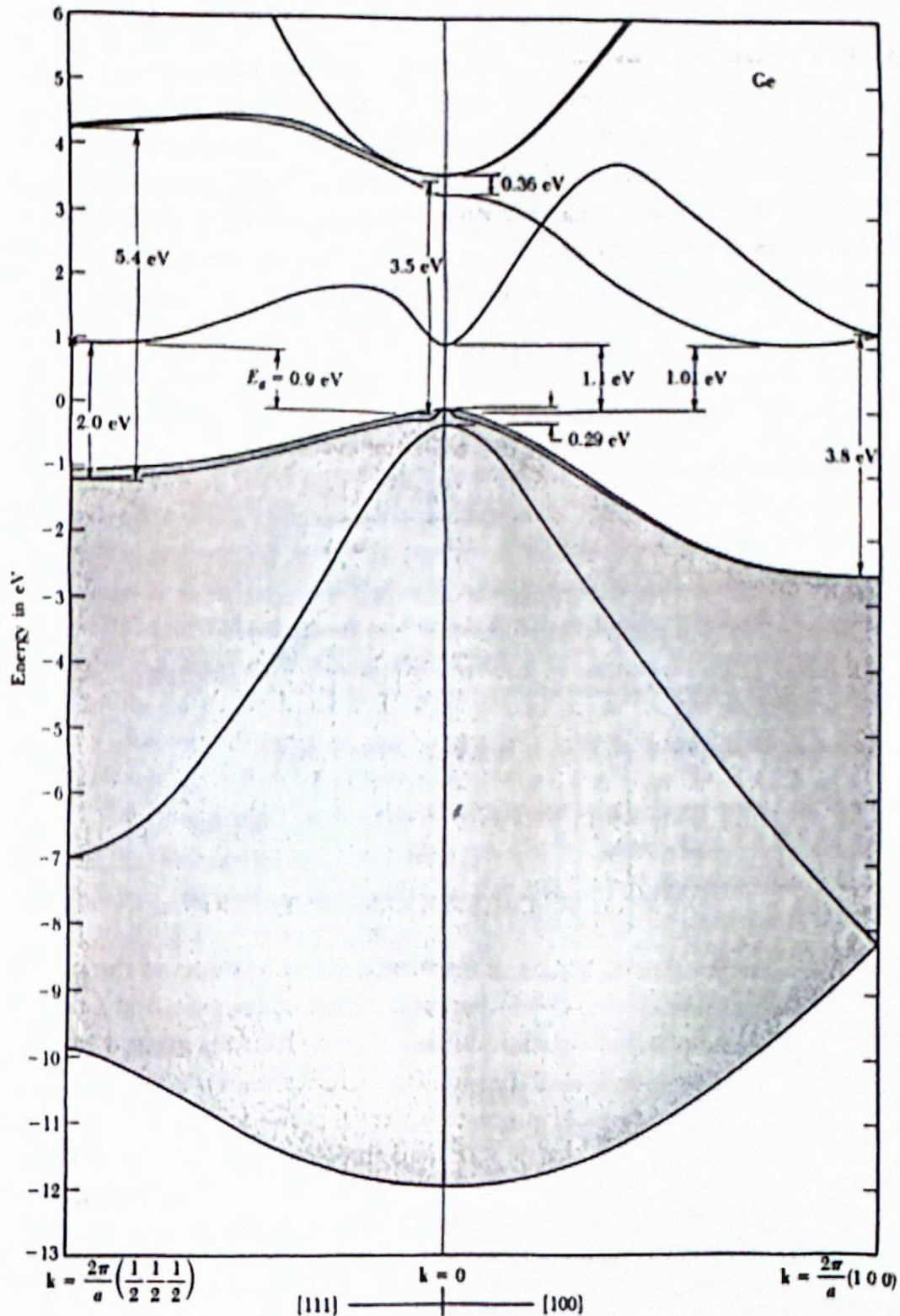


Figure 14 Calculated band structure of germanium, after C. Y. Fong. The general features are in good agreement with experiment. The four valence bands are shown in gray. The fine structure of the valence band edge is caused by spin-orbit splitting. The energy gap is indirect; the conduction band edge is at the point $(2\pi/a)(1\ 0\ 0)$. The constant energy surfaces around this point are ellipsoidal.

The choice of sign distinguishes the two masses. The split-off band has $\epsilon(k) = -\Delta + Ak^2$. The experiments give, in units $\hbar^2/2m$,

$$\begin{array}{llll} \text{Si:} & A = -4.29 ; & |B| = 0.68 ; & |C| = 4.87 ; & \Delta = 0.044 \text{ eV} \\ \text{Ge:} & A = -13.38 ; & |B| = 8.48 ; & |C| = 13.15 ; & \Delta = 0.29 \text{ eV} \end{array}$$

Roughly, the light and heavy holes in germanium have masses $0.043 m$ and $0.34 m$; in silicon $0.16 m$ and $0.52 m$; in diamond $0.7 m$ and $2.12 m$.

The conduction band edges in Ge are at the equivalent points L of the Brillouin zone, Fig. 15a. Each band edge has a spheroidal energy surface oriented along a $\langle 111 \rangle$ crystal axis, with a longitudinal mass $m_l = 1.59 m$ and a transverse mass $m_t = 0.082 m$. For a static magnetic field at an angle θ with the longitudinal axis of a spheroid, the effective cyclotron mass m_c is

$$\frac{1}{m_c^2} = \frac{\cos^2 \theta}{m_l^2} + \frac{\sin^2 \theta}{m_t m_l} \quad (34)$$

Results for Ge are shown in Fig. 16.

In silicon the conduction band edges are spheroids oriented along the equivalent $\langle 100 \rangle$ directions in the Brillouin zone, with mass parameters $m_l = 0.92 m$ and $m_t = 0.19 m$, as in Fig. 17a. The band edges lie along the lines labeled Δ in the zone of Fig. 15a, a little way in from the boundary points X .

In GaAs we have $A = -6.98$, $B = -4.5$, $|C| = 6.2$, $\Delta = 0.341 \text{ eV}$. The band structure is shown in Fig. 17b.

INTRINSIC CARRIER CONCENTRATION

We want the concentration of intrinsic carriers in terms of the band gap. We do the calculation for simple parabolic band edges. We first calculate in terms of the chemical potential μ the number of electrons excited to the conduction band at temperature T . In semiconductor physics μ is called the Fermi level. At the temperatures of interest we may suppose for the conduction band of a semiconductor that $\epsilon - \mu \gg k_B T$, and the Fermi-Dirac distribution function reduces to

$$f_e \approx \exp[(\mu - \epsilon)/k_B T] . \quad (35)$$

This is the probability that a conduction electron orbital is occupied, in an approximation valid when $f_e \ll 1$.

The energy of an electron in the conduction band is

$$\epsilon_k = E_c + \hbar^2 k^2 / 2m_e , \quad (36)$$

where E_c is the energy at the conduction band edge, as in Fig. 18. Here m_e is the effective mass of an electron. Thus from (6.20) the density of states at ϵ is

$$D_e(\epsilon) = \frac{1}{2\pi^2} \left(\frac{2m_e}{\hbar^2} \right)^{3/2} (\epsilon - E_c)^{1/2} . \quad (37)$$

SEVENTH EDITION

JOHN WILEY & SONS, INC.

New York Chichester Brisbane Toronto Singapore

ISBN 0-471-11181-3



90000>

9 780471 111818

Greenthread Ex. 2084, p. 27 of 27
Semiconductor v. Greenthread