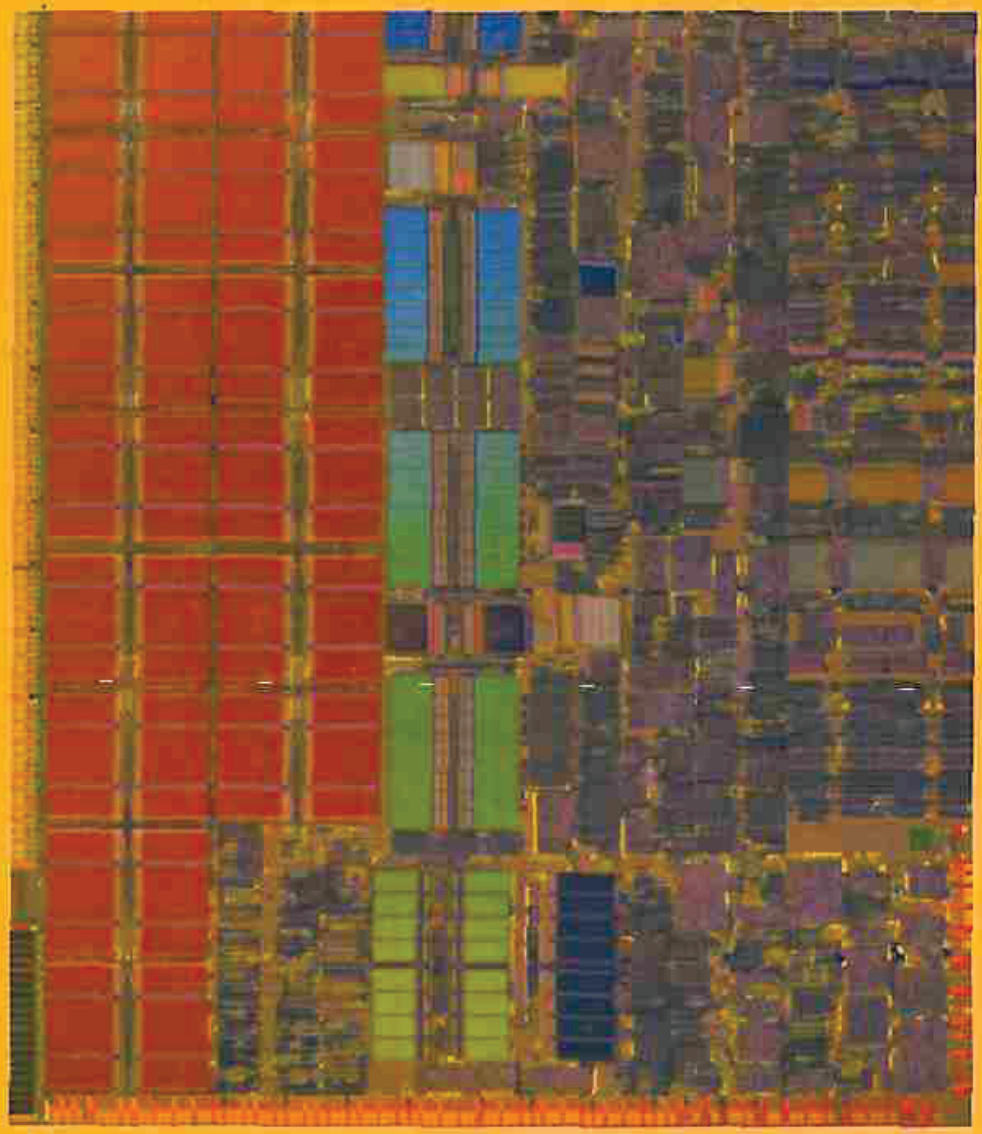
SOLID STATE ELECTRONIC DEVICES

Eastern Economy Edition

SIXTH EDITION



BEN G. STREETMAN • SANJAY KUMAR BANERJEE

CIRRUS EX. 1062 - 1/190

SIXTH EDITION

Solid State Electronic Devices

BEN G. STREETMAN AND SANJAY KUMAR BANERJEE

Microelectronics Research Center
Department of Electrical and Computer Engineering
The University of Texas at Austin

PHI Learning Private Limited
New Delhi-110001
2009

CIRRUS EX. 1062 - 2/190

This Indian Reprint—Rs. 295.00
(Original U.S. Edition—Rs. 5695.00)

SOLID STATE ELECTRONIC DEVICES, 6th Ed.
by Ben G. Streetman and Sanjay Kumar Banerjee

© 2006 by Prentice-Hall, Inc. (now known as Pearson Education Inc.), Upper Saddle River, New Jersey 07458, U.S.A. All rights reserved. No part of this book may be reproduced in any form, by mimeograph or any other means, without permission in writing from the publisher.

The author and publisher of this book have used their best efforts in preparing this book. These efforts include the development, research, and testing of the theories and programs to determine their effectiveness. The author and publisher shall not be liable in any event for incidental or consequential damages with, or arising out of, the furnishing, performance, or use of these programs.

ISBN-978-81-203-3020-7

Published by Asoke K. Ghosh, PHI Learning Private Limited, M-97, Connaught Circus, New Delhi-110001 and Printed by Mohan Makhijani at Rekha Printers Private Limited, New Delhi-110020.

Solid State Electronic Devices

CIRRUS EX. 1062 - 4/190

CONTENTS

PREFACE XIII

ABOUT THE AUTHORS XVII

1 CRYSTAL PROPERTIES AND GROWTH OF SEMICONDUCTORS 1

1.1 Semiconductor Materials 1
1.2 Crystal Lattices 3
1.2.1 Periodic Structures 3
1.2.2 Cubic Lattices 5
1.2.3 Planes and Directions 7
1.2.4 The Diamond Lattice 10
1.3 Bulk Crystal Growth 12
1.3.1 Starting Materials 13
1.3.2 Growth of Single-Crystal Ingots 13
1.3.3 Wafers 15
1.3.4 Doping 16
1.4 Epitaxial Growth 18
1.4.1 Lattice-Matching in Epitaxial Growth 18
1.4.2 Vapor-Phase Epitaxy 21
1.4.3 Molecular Beam Epitaxy 23

2 ATOMS AND ELECTRONS 31

2.1 Introduction to Physical Models 32
2.2 Experimental Observations 33
2.2.1 The Photoelectric Effect 33
2.2.2 Atomic Spectra 35
2.3 The Bohr Model 36
2.4 Quantum Mechanics 39
2.4.1 Probability and the Uncertainty Principle 40
2.4.2 The Schrödinger Wave Equation 41
2.4.3 Potential Well Problem 44
2.4.4 Tunneling 46
2.5 Atomic Structure and the Periodic Table 47
2.5.1 The Hydrogen Atom 47
2.5.2 The Periodic Table 50

CIRRUS EX. 1062 - 5/190

3 ENERGY BANDS AND CHARGE CARRIERS IN SEMICONDUCTORS 61

3.1 Bonding Forces and Energy Bands in Solids 61
3.1.1 Bonding Forces in Solids 62
3.1.2 Energy Bands 64
3.1.3 Metals, Semiconductors, and Insulators 67
3.1.4 Direct and Indirect Semiconductors 68
3.1.5 Variation of Energy Bands with Alloy Composition 71
3.2 Charge Carriers in Semiconductors 71
3.2.1 Electrons and Holes 73
3.2.2 Effective Mass 76
3.2.3 Intrinsic Material 80
3.2.4 Extrinsic Material 81
3.2.5 Electrons and Holes in Quantum Wells 85
3.3 Carrier Concentrations 86
3.3.1 The Fermi Level 86
3.3.2 Electron and Hole Concentrations at Equilibrium 89
3.3.3 Temperature Dependence of Carrier Concentrations 94
3.3.4 Compensation and Space Charge Neutrality 96
3.4 Drift of Carriers in Electric and Magnetic Fields 98
3.4.1 Conductivity and Mobility 98
3.4.2 Drift and Resistance 102
3.4.3 Effects of Temperature and Doping on Mobility 103
3.4.4 High-Field Effects 105
3.4.5 The Hall Effect 106
3.5 Invariance of the Fermi Level at Equilibrium 108

4 EXCESS CARRIERS IN SEMICONDUCTORS 118

4.1 Optical Absorption 118
4.2 Luminescence 121
4.2.1 Photoluminescence 122
4.2.2 Electroluminescence 124
4.3 Carrier Lifetime and Photoconductivity 124
4.3.1 Direct Recombination of Electrons and Holes 125
4.3.2 Indirect Recombination; Trapping 127
4.3.3 Steady State Carrier Generation; Quasi-Fermi Levels 130
4.3.4 Photoconductive Devices 132
4.4 Diffusion of Carriers 134
4.4.1 Diffusion Processes 134
4.4.2 Diffusion and Drift of Carriers; Built-in Fields 137

CIRRUS EX. 1062 - 6/190

4.4.3 Diffusion and Recombination; The Continuity Equation 140
4.4.4 Steady State Carrier Injection; Diffusion Length 141
4.4.5 The Haynes–Shockley Experiment 144
4.4.6 Gradients in the Quasi-Fermi Levels 147

5 JUNCTIONS 154

5.1 Fabrication of p-n Junctions 154
5.1.1 Thermal Oxidation 155
5.1.2 Diffusion 156
5.1.3 Rapid Thermal Processing 158
5.1.4 Ion Implantation 159
5.1.5 Chemical Vapor Deposition (CVD) 162
5.1.6 Photolithography 162
5.1.7 Etching 166
5.1.8 Metallization 167
5.2 Equilibrium Conditions 168
5.2.1 The Contact Potential 170
5.2.2 Equilibrium Fermi Levels 174
5.2.3 Space Charge at a Junction 175
5.3 Forward- and Reverse-Biased Junctions; Steady State Conditions 180
5.3.1 Qualitative Description of Current Flow at a Junction 180
5.3.2 Carrier Injection 184
5.3.3 Reverse Bias 193
5.4 Reverse-Bias Breakdown 196
5.4.1 Zener Breakdown 197
5.4.2 Avalanche Breakdown 198
5.4.3 Rectifiers 200
5.4.4 The Breakdown Diode 204
5.5 Transient and A-C Conditions 205
5.5.1 Time Variation of Stored Charge 205
5.5.2 Reverse Recovery Transient 208
5.5.3 Switching Diodes 211
5.5.4 Capacitance of p-n junctions 212
5.5.5 The Varactor Diode 217
5.6 Deviations from the Simple Theory 218
5.6.1 Effects of Contact Potential on Carrier Injection 219
5.6.2 Recombination and Generation in the Transition Region 221
5.6.3 Ohmic Losses 223
5.6.4 Graded Junctions 224

CIRRUS EX. 1062 - 7/190

5.7 Metal–Semiconductor Junctions 227
5.7.1 Schottky Barriers 227
5.7.2 Rectifying Contacts 229
5.7.3 Ohmic Contacts 231
5.7.4 Typical Schottky Barriers 233
5.8 Heterojunctions 234

6 FIELD-EFFECT TRANSISTORS 251
6.1 Transistor Operation 252
6.1.1 The Load Line 252
6.1.2 Amplification and Switching 254
6.2 The Junction FET 254
6.2.1 Pinch-off and Saturation 255
6.2.2 Gate Control 257
6.2.3 Current–Voltage Characteristics 259
6.3 The Metal–Semiconductor FET 261
6.3.1 The GaAs MESFET 261
6.3.2 The High Electron Mobility Transistor (HEMT) 262
6.3.3 Short Channel Effects 264
6.4 The Metal–Insulator–Semiconductor FET 265
6.4.1 Basic Operation and Fabrication 266
6.4.2 The Ideal MOS Capacitor 270
6.4.3 Effects of Real Surfaces 281
6.4.4 Threshold Voltage 284
6.4.5 MOS Capacitance–Voltage Analysis 286
6.4.6 Time-Dependent Capacitance Measurements 290
6.4.7 Current–Voltage Characteristics of MOS Gate Oxides 291
6.5 The MOS Field-Effect Transistor 294
6.5.1 Output Characteristics 295
6.5.2 Transfer Characteristics 297
6.5.3 Mobility Models 300
6.5.4 Short Channel MOSFET I–V Characteristics 302
6.5.5 Control of Threshold Voltage 303
6.5.6 Substrate Bias Effects 309
6.5.7 Subthreshold Characteristics 311
6.5.8 Equivalent Circuit for the MOSFET 313
6.5.9 MOSFET Scaling and Hot Electron Effects 315
6.5.10 Drain-Induced Barrier Lowering 320
6.5.11 Short Channel Effect and Narrow Width Effect 322
6.5.12 Gate-Induced Drain Leakage 323

CIRRUS EX. 1062 - 8/190

7 BIPOLAR JUNCTION TRANSISTORS 336
7.1 Fundamentals of BJT Operation 336
7.2 Amplification with BJTs 340
7.3 BJT Fabrication 343
7.4 Minority Carrier Distributions and Terminal Currents 346
7.4.1 Solution of the Diffusion Equation in the Base Region 347
7.4.2 Evaluation of the Terminal Currents 349
7.4.3 Approximations of the Terminal Currents 352
7.4.4 Current Transfer Ratio 354
7.5 Generalized Biasing 355
7.5.1 The Coupled-Diode Model 356
7.5.2 Charge Control Analysis 361
7.6 Switching 363
7.6.1 Cutoff 364
7.6.2 Saturation 365
7.6.3 The Switching Cycle 366
7.6.4 Specifications for Switching Transistors 368
7.7 Other Important Effects 368
7.7.1 Drift in the Base Region 369
7.7.2 Base Narrowing 370
7.7.3 Avalanche Breakdown 372
7.7.4 Injection Level; Thermal Effects 373
7.7.5 Base Resistance and Emitter Crowding 374
7.7.6 Gummel–Poon Model 376
7.7.7 Kirk Effect 380
7.8 Frequency Limitations of Transistors 382
7.8.1 Capacitance and Charging Times 382
7.8.2 Transit Time Effects 385
7.8.3 Webster Effect 386
7.8.4 High-Frequency Transistors 386
7.9 Heterojunction Bipolar Transistors 388

8 OPTOELECTRONIC DEVICES 398
8.1 Photodiodes 398
8.1.1 Current and Voltage in an Illuminated Junction 399
8.1.2 Solar Cells 402
8.1.3 Photodetectors 405
8.1.4 Gain, Bandwidth, and Signal-to-Noise Ratio of Photodetectors 407

CIRRUS EX. 1062 - 9/190

8.2 Light-Emitting Diodes 411
8.2.1 Light-Emitting Materials 411
8.2.2 Fiber-Optic Communications 414
8.3 Lasers 417
8.4 Semiconductor Lasers 422
8.4.1 Population Inversion at a Junction 422
8.4.2 Emission Spectra for p-n Junction Lasers 424
8.4.3 The Basic Semiconductor Laser 426
8.4.4 Heterojunction Lasers 426
8.4.5 Materials for Semiconductor Lasers 430

9 INTEGRATED CIRCUITS 437
9.1 Background 438
9.1.1 Advantages of Integration 438
9.1.2 Types of Integrated Circuits 440
9.2 Evolution of Integrated Circuits 441
9.3 Monolithic Device Elements 444
9.3.1 CMOS Process Integration 444
9.3.2 Silicon-on-Insulator (SOI) 459
9.3.3 Integration of Other Circuit Elements 461
9.4 Charge Transfer Devices 466
9.4.1 Dynamic Effects in MOS Capacitors 466
9.4.2 The Basic CCD 468
9.4.3 Improvements on the Basic Structure 469
9.4.4 Applications of CCDs 470
9.5 Ultra Large-Scale Integration (ULSI) 470
9.5.1 Logic Devices 474
9.5.2 Semiconductor Memories 483
9.6 Testing, Bonding, and Packaging 496
9.6.1 Testing 497
9.6.2 Wire Bonding 497
9.6.3 Flip-Chip Techniques 501
9.6.4 Packaging 502

10 HIGH-FREQUENCY AND HIGH-POWER DEVICES 508
10.1 Tunnel Diodes 508
10.1.1 Degenerate Semiconductors 509
10.2 The IMPATT Diode 512
10.3 The Gunn Diode 515
10.3.1 The Transferred-Electron Mechanism 515
10.3.2 Formation and Drift of Space Charge Domains 518

CIRRUS EX. 1062 - 10/190

10.4 The p-n-p-n Diode 520
10.4.1 Basic Structure 520
10.4.2 The Two-Transistor Analogy 521
10.4.3 Variation of α with Injection 522
10.4.4 Forward-Blocking State 523
10.4.5 Conducting State 524
10.4.6 Triggering Mechanisms 525
10.5 The Semiconductor-Controlled Rectifier 526
10.5.1 Turning off the SCR 527
10.6 Insulated-Gate Bipolar Transistor 528

APPENDICES

I. Definitions of Commonly Used Symbols 535
II. Physical Constants and Conversion Factors 539
III. Properties of Semiconductor Materials 540
IV. Derivation of the Density of States in the Conduction Band 541
V. Derivation of Fermi–Dirac Statistics 546
VI. Dry and Wet Thermal Oxide Thickness Grown on Si (100) as a Function of Time and Temperature 550
VII. Solid Solubilities of Impurities in Si 552
VIII. Diffusivities of Dopants in Si and SiO_2 554
IX. Projected Range and Straggle as Function of Implant Energy in Si 556

ANSWERS TO SELECTED SELF QUIZ QUESTIONS 559

INDEX 563

CIRRUS EX. 1062 - 11/190

Chapter 3

Energy Bands and Charge Carriers in Semiconductors

OBJECTIVES

1. Understand conduction and valence energy bands, and how bandgaps are formed
2. Appreciate the idea of doping in semiconductors
3. Use the density of states and Fermi Dirac statistics to calculate carrier concentrations
4. Calculate drift currents in an electric field in terms of carrier mobility, and how mobility is affected by scattering
5. Discuss the idea of "effective" masses

In this chapter we begin to discuss the specific mechanisms by which current flows in a solid. In examining these mechanisms we shall learn why some materials are good conductors of electric current, whereas others are poor conductors. We shall see how the conductivity of a semiconductor can be varied by changing the temperature or the number of impurities. These fundamental concepts of charge transport form the basis for later discussions of solid state device behavior.

3.1 BONDING FORCES AND ENERGY BANDS IN SOLIDS

In Chapter 2 we found that electrons are restricted to sets of discrete energy levels within atoms. Large gaps exist in the energy scale in which no energy states are available. In a similar fashion, electrons in solids are restricted to certain energies and are not allowed at other energies. The basic difference between the case of an electron in a solid and that of an electron in an isolated atom is that in the solid the electron has a *range*, or *band*, of available energies. The discrete energy levels of the isolated atom spread into bands of energies in the solid because in the solid the wave functions of electrons in neighboring atoms overlap, and an electron is not necessarily localized at a particular atom. Thus, for example, an electron in the outer orbit of one atom feels the influence of neighboring atoms, and its overall wave function is altered. Naturally, this influence affects the potential energy term and the boundary conditions in the Schrödinger equation, and we would

CIRRUS EX. 1062 - 12/190

expect to obtain different energies in the solution. Usually, the influence of neighboring atoms on the energy levels of a particular atom can be treated as a small perturbation, giving rise to shifting and splitting of energy states into energy bands.

3.1.1 Bonding Forces in Solids

The interaction of electrons in neighboring atoms of a solid serves the very important function of holding the crystal together. For example, alkali halides such as NaCl are typified by *ionic bonding*. In the NaCl lattice, each Na atom

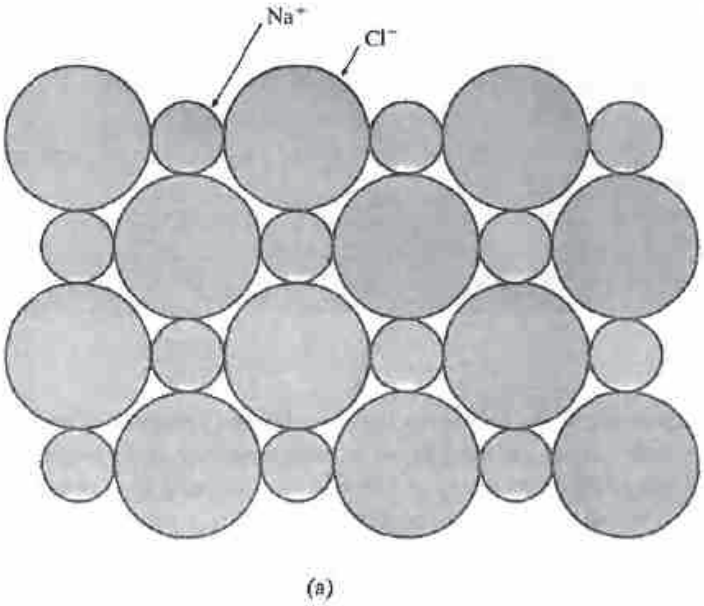


Figure 3–1 Different types of chemical bonding in solids: (a) an example of ionic bonding in NaCl; (b) covalent bonding in the Si crystal, viewed along a <100> direction (see also Figs. 1–8 and 1–9).

CIRRUS EX. 1062 - 13/190

is surrounded by six nearest neighbor Cl atoms, and vice versa. Four of the nearest neighbors are evident in the two-dimensional representation shown in Fig. 3-1a. The electronic structure of Na ($Z = 11$) is [Ne] $3s^1$, and Cl ($Z = 17$) has the structure [Ne]$3s^2 3p^5$. In the lattice each Na atom gives up its outer $3s$ electron to a Cl atom, so that the crystal is made up of ions with the electronic structures of the inert atoms Ne and Ar (Ar has the electronic structure [Ne]$3s^2 3p^6$). However, the ions have net electric charges after the electron exchange. The Na^+ ion has a net positive charge, having lost an electron, and the Cl^- ion has a net negative charge, having gained an electron.

Each Na^+ ion exerts an electrostatic attractive force upon its six Cl^- neighbors, and vice versa. These coulombic forces pull the lattice together until a balance is reached with repulsive forces. A reasonably accurate calculation of the atomic spacing can be made by considering the ions as hard spheres being attracted together (Example 1-1).

An important observation in the NaCl structure is that all electrons are tightly bound to atoms. Once the electron exchanges have been made between the Na and Cl atoms to form the Na^+ and Cl^- ions, the outer orbits of all atoms are completely filled. Since the ions have the closed-shell configurations of the inert atoms Ne and Ar, there are no loosely bound electrons to participate in current flow; as a result, NaCl is a good insulator.

In a metal atom the outer electronic shell is only partially filled, usually by no more than three electrons. We have already noted that the alkali metals (e.g., Na) have only one electron in the outer orbit. This electron is loosely bound and is given up easily in ion formation. This accounts for the great chemical activity in the alkali metals, as well as for their high electrical conductivity. In the metal the outer electron of each alkali atom is contributed to the crystal as a whole, so that the solid is made up of ions with closed shells immersed in a sea of free electrons. The forces holding the lattice together arise from an interaction between the positive ion cores and the surrounding free electrons. This is one type of *metallic bonding*. Obviously, there are complicated differences in the bonding forces for various metals, as evidenced by the wide range of melting temperatures (234 K for Hg, 3643 K for W). However, the metals have the sea of electrons in common, and these electrons are free to move about the crystal under the influence of an electric field.

A third type of bonding is exhibited by the diamond lattice semiconductors. We recall that each atom in the Ge, Si, or C diamond lattice is surrounded by four nearest neighbors, each with four electrons in the outer orbit. In these crystals each atom shares its valence electrons with its four neighbors (Fig. 3-1b). Bonding between nearest neighbor atoms is illustrated in the diamond lattice diagram of Fig. 1-9. The bonding forces arise from a quantum mechanical interaction between the shared electrons. This is known as *covalent bonding;* each electron pair constitutes a covalent bond. In the sharing process it is no longer relevant to ask which electron belongs to a particular atom—both belong to the bond. The two electrons are indistinguishable, except that they must have opposite spin to satisfy the Pauli exclusion principle. Covalent bonding is also found in certain molecules, such as H_2.

CIRRUS EX. 1062 - 14/190

As in the case of the ionic crystals, no free electrons are available to the lattice in the covalent diamond structure of Fig. 3-1b. By this reasoning Ge and Si should also be insulators. However, we have pictured an idealized lattice at 0 K in this figure. As we shall see in subsequent sections, an electron can be thermally or optically excited out of a covalent bond and thereby become free to participate in conduction. This is an important feature of semiconductors.

Compound semiconductors such as GaAs have mixed bonding, in which both ionic and covalent bonding forces participate. Some ionic bonding is to be expected in a crystal such as GaAs because of the difference in placement of the Ga and As atoms in the periodic table. The ionic character of the bonding becomes more important as the atoms of the compound become further separated in the periodic table, as in the II–VI compounds. Such electronic structure, and specifically the idea that the outermost valence shell is *complete* if it has a stable set of eight electrons (Ne, Ar, Kr), is the basis of most of chemistry and many of the semiconducting properties.

3.1.2 Energy Bands

As isolated atoms are brought together to form a solid, various interactions occur between neighboring atoms, including those described in the preceding section. The forces of attraction and repulsion between atoms will find a balance at the proper interatomic spacing for the crystal. In the process, important changes occur in the electron energy level configurations, and these changes result in the varied electrical properties of solids.

In Fig. 2–8, we showed the orbital model of a Si atom, along with the energy levels of the various electrons in the coulombic potential well of the nucleus. Let us focus on the outermost shell or valence shell, $\mathbf{n} = 3$, where two $3s$ and two $3p$ electrons interact to form the four "hybridized" sp^3 electrons when the atoms are brought close together. In Fig. 3–2, we schematically show the coulombic potential wells of two atoms close to each other, along with the wave functions of two electrons centered on the two nuclei. By solving the Schrödinger equation for such an interacting system, we find that the composite two-electron wave functions are *linear combinations* of the individual *atomic orbitals* (LCAO). The odd or antisymmetric combination is called the antibonding orbital, while the even or symmetric combination is the bonding orbital. It can be seen that the bonding orbital has a higher value of the wave function (and therefore the electron probability density) than the antibonding state in the region between the two nuclei. This corresponds to the covalent bond between the atoms.

To determine the energy levels of the bonding and the antibonding states, it is important to recognize that in the region between the two nuclei the coulombic potential energy $V(r)$ is lowered (solid line in Fig. 3–2) compared to isolated atoms (dashed lines). It is easy to see why the potential energy would be lowered in this region, because an electron here would be attracted by two nuclei, rather than just one. For the bonding state the electron

CIRRUS EX. 1062 - 15/190

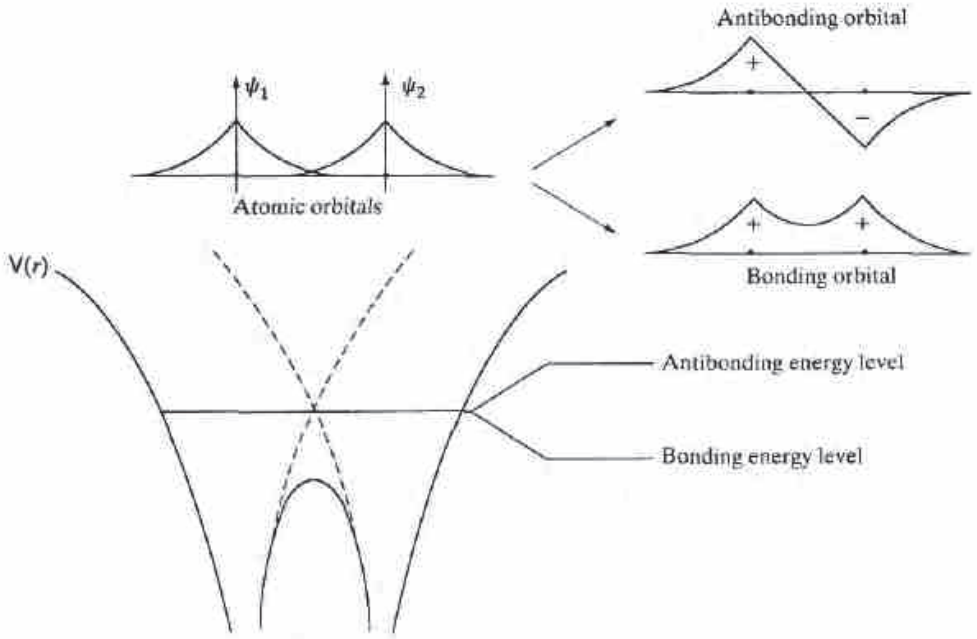


Figure 3–2
Linear combinations of atomic orbitals (LCAO): The LCAO when 2 atoms are brought together leads to 2 distinct "normal" modes—a higher energy antibonding orbital, and a lower energy bonding orbital. Note that the electron probability density is high in the region between the ion cores (covalent "bond"), leading to lowering of the bonding energy level and the cohesion of the crystal. If instead of 2 atoms, one brings together N atoms, there will be N distinct LCAO, and N closely spaced energy levels in a band.

probability density is higher in this region of lowered potential energy than for the antibonding state. As a result, the original isolated atomic energy level would be split into two, a lower bonding energy level and a higher antibonding level. It is the lowering of the energy of the bonding state that gives rise to cohesion of the crystal. For even smaller interatomic spacings, the energy of the crystal goes up because of repulsion between the nuclei, and other electronic interactions. Since the probability density is given by the square of the wave function, if the entire wave function is multiplied by -1, it does not lead to a different LCAO. The important point to note in this discussion is that the number of distinct LCAO, and the number of distinct energy levels, depends on the number of atoms that are brought together. The lowest energy level corresponds to the totally symmetric LCAO, the highest corresponds to the totally antisymmetric case, and the other combinations lead to energy levels in between.

Qualitatively, we can see that as atoms are brought together, the application of the Pauli exclusion principle becomes important. When two atoms are completely isolated from each other so that there is no interaction of electron wave functions between them, they can have identical electronic

CIRRUS EX. 1062 - 16/190

structures. As the spacing between the two atoms becomes smaller, however, electron wave functions begin to overlap. The exclusion principle dictates that no two electrons in a given interacting system may have the same quantum state; thus there must be at most one electron per level after there is a splitting of the discrete energy levels of the isolated atoms into new levels belonging to the pair rather than to individual atoms.

In a solid, many atoms are brought together, so that the split energy levels form essentially continuous *bands* of energies. As an example, Fig. 3–3 illustrates the imaginary formation of a silicon crystal from isolated silicon atoms. Each isolated silicon atom has an electronic structure $1s^2 2s^2 2p^6 3s^2 3p^2$ in the ground state. Each atom has available two $1s$ states, two $2s$ states, six $2p$ states, two $3s$ states, six $3p$ states, and higher states (see Tables 2–1 and 2–2). If we consider N atoms, there will be $2N$, $2N$, $6N$, $2N$, and $6N$ states of type $1s$, $2s$, $2p$, $3s$, and $3p$, respectively. As the interatomic spacing decreases, these energy levels split into bands, beginning with the outer ($\mathbf{n} = 3$) shell. As the "$3s$" and "$3p$" bands grow, they merge into a single band composed of a mixture of energy levels. This band of "$3s$–$3p$" levels contains $8N$ available states. As the

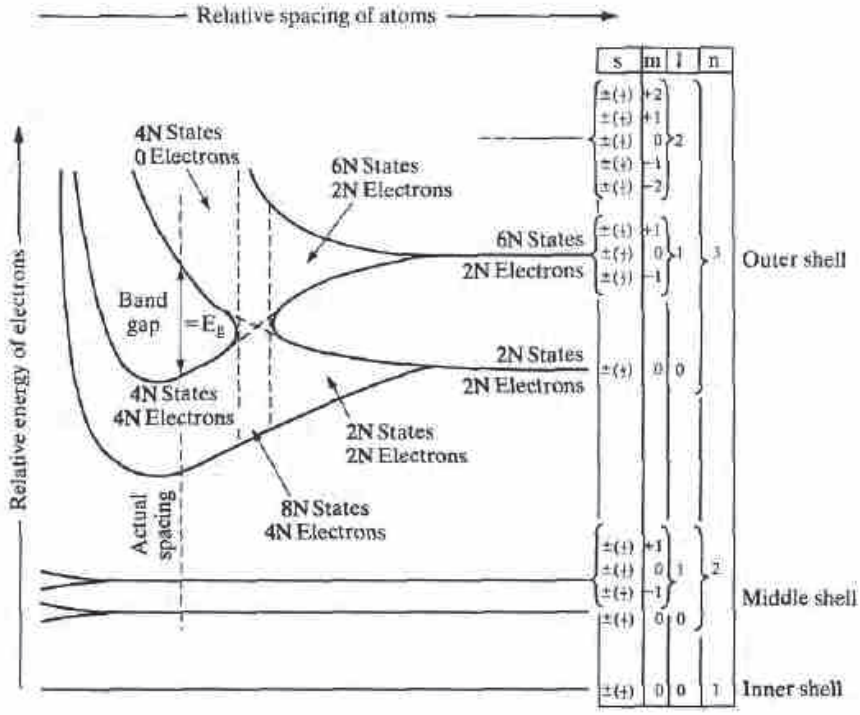


Figure 3–3
Energy levels in Si as a function of interatomic spacing. The core levels ($\mathbf{n} = 1,2$) in Si are completely filled with electrons. At the actual atomic spacing of the crystal, the $2N$ electrons in the $3s$ subshell and the $2N$ electrons in the $3p$ subshell undergo sp^3 hybridization, and all end up in the lower $4N$ states (valence band), while the higher-lying $4N$ states (conduction band) are empty, separated by a band gap.

CIRRUS EX. 1062 - 17/190

distance between atoms approaches the equilibrium interatomic spacing of silicon, this band splits into two bands separated by an *energy gap* E_g. The upper band (called the *conduction band*) contains $4N$ states, as does the lower (*valence*) band. Thus, apart from the low-lying and tightly bound "core" levels, the silicon crystal has two bands of available energy levels separated by an energy gap E_g wide, which contains no allowed energy levels for electrons to occupy. This gap is sometimes called a "forbidden band," since in a perfect crystal it contains no electron energy states.

We should pause at this point and count electrons. The lower "1*s*" band is filled with the $2N$ electrons which originally resided in the collective 1*s* states of the isolated atoms. Similarly, the 2*s* band and the 2*p* bands will have $2N$ and $6N$ electrons in them, respectively. However, there were $4N$ electrons in the original isolated $\mathbf{n} = 3$ shells ($2N$ in 3*s* states and $2N$ in 3*p* states). These $4N$ electrons must occupy states in the valence band or the conduction band in the crystal. At 0 K the electrons will occupy the lowest energy states available to them. In the case of the Si crystal, there are exactly $4N$ states in the valence band available to the $4N$ electrons. Thus at 0 K, *every* state in the valence band will be filled, while the conduction band will be completely empty of electrons. As we shall see, this arrangement of completely filled and empty energy bands has an important effect on the electrical conductivity of the solid.

3.1.3 Metals, Semiconductors, and Insulators

Every solid has its own characteristic energy band structure. This variation in band structure is responsible for the wide range of electrical characteristics observed in various materials. The silicon band structure of Fig. 3-3, for example, can give a good picture of why silicon in the diamond lattice is a good insulator. To reach such a conclusion, we must consider the properties of completely filled and completely empty energy bands in the current conduction process.

Before discussing the mechanisms of current flow in solids further, we can observe here that for electrons to experience acceleration in an applied electric field, they must be able to move into new energy states. This implies there must be empty states (allowed energy states which are not already occupied by electrons) available to the electrons. For example, if relatively few electrons reside in an otherwise empty band, ample unoccupied states are available into which the electrons can move. On the other hand, the silicon band structure is such that the valence band is completely filled with electrons at 0 K and the conduction band is empty. There can be no charge transport within the valence band, since no empty states are available into which electrons can move. There are no electrons in the conduction band, so no charge transport can take place there either. Thus silicon has a high resistivity typical of insulators.

Semiconductor materials at 0 K have basically the same structure as insulators—a filled valence band separated from an empty conduction band by a band gap containing no allowed energy states (Fig. 3-4). The difference lies in the size of the band gap E_g, which is much smaller in semiconductors

CIRRUS EX. 1062 - 18/190

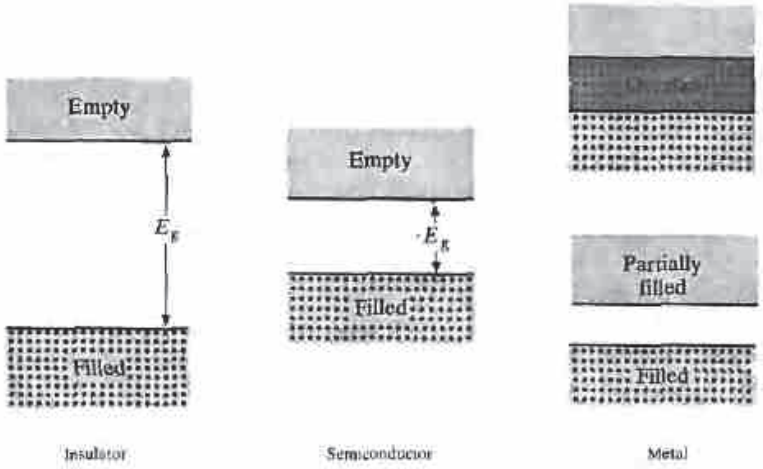


Figure 3–4
Typical band structures at 0 K.

than in insulators. For example, the semiconductor Si has a band gap of about 1.1 eV compared with 5 eV for diamond. The relatively small band gaps of semiconductors (Appendix III) allow for excitation of electrons from the lower (valence) band to the upper (conduction) band by reasonable amounts of thermal or optical energy. For example, at room temperature a semiconductor with a 1-eV band gap will have a significant number of electrons excited thermally across the energy gap into the conduction band, whereas an insulator with $E_g = 10$ eV will have a negligible number of such excitations. Thus an important difference between semiconductors and insulators is that the number of electrons available for conduction can be increased greatly in semiconductors by thermal or optical energy.

In metals the bands either overlap or are only partially filled. Thus electrons and empty energy states are intermixed within the bands so that electrons can move freely under the influence of an electric field. As expected from the metallic band structures of Fig. 3–4, metals have a high electrical conductivity.

3.1.4 Direct and Indirect Semiconductors

The "thought experiment" of Section 3.1.2, in which isolated atoms were brought together to form a solid, is useful in pointing out the existence of energy bands and some of their properties. Other techniques are generally used, however, when quantitative calculations are made of band structures. In a typical calculation, a single electron is assumed to travel through a perfectly periodic lattice. The wave function of the electron is assumed to be in the form of a plane wave[1] moving, for example, in the x- direction

[1]Discussions of plane waves are available in most sophomore physics texts or in introductory electromagnetics texts.

CIRRUS EX. 1062 - 19/190

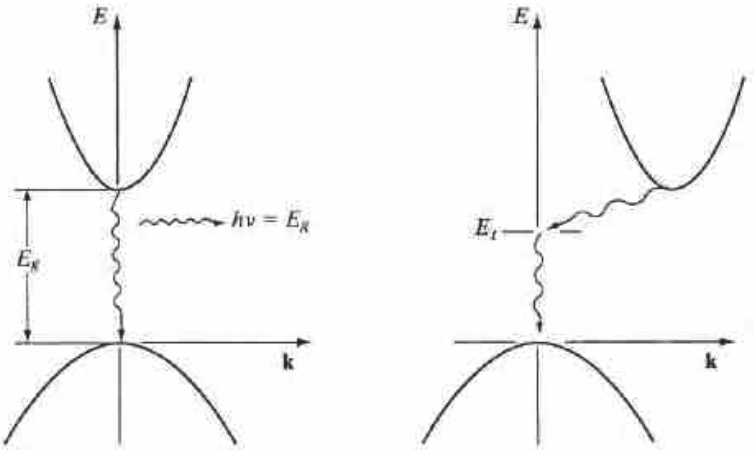


Figure 3–5
Direct and indirect electron transitions in semiconductors: (a) direct transition with accompanying photon emission; (b) indirect transition via a defect level.

with propagation constant **k**, also called a *wave vector.* The space-dependent wave function for the electron is

$$\psi_{\mathbf{k}}(x) = U(\mathbf{k}_x, x)e^{j\mathbf{k}_x x} \tag{3-1}$$

where the function $U(\mathbf{k}_x, x)$ modulates the wave function according to the periodicity of the lattice.

In such a calculation, allowed values of energy can be plotted vs. the propagation constant **k**. Since the periodicity of most lattices is different in various directions, the $(E, \mathbf{k})$ diagram must be plotted for the various crystal directions, and the full relationship between E and **k** is a complex surface which should be visualized in three dimensions.

The band structure of GaAs has a minimum in the conduction band and a maximum in the valence band for the same **k** value ($\mathbf{k} = 0$). On the other hand, Si has its valence band maximum at a different value of **k** than its conduction band minimum. Thus an electron making a smallest-energy transition from the conduction band to the valence band in GaAs can do so without a change in **k** value; on the other hand, a transition from the minimum point in the Si conduction band to the maximum point of the valence band requires some change in **k**. Thus there are two classes of semiconductor energy bands; *direct* and *indirect* (Fig. 3-5). We can show that an indirect transition, involving a change in **k**, requires a change of momentum for the electron.

EXAMPLE 3–1

Assuming that U is constant in Eq. (3-1) for an essentially free electron, show that the x-component of the electron momentum in the crystal is given by $\langle p_x \rangle = \hbar \mathbf{k}_x$.

CIRRUS EX. 1062 - 20/190

SOLUTION From Eq. (3–1)

$$\psi_{\mathbf{k}}(x) = Ue^{j\mathbf{k}_x x}$$

Using Eq. (2–21b) and the momentum operator,

$$\langle p_x \rangle = \frac{\int_{-\infty}^{\infty} U^2 e^{-j\mathbf{k}_x x} \frac{\hbar}{j} \frac{\partial}{\partial x} (e^{j\mathbf{k}_x x})\, dx}{\int_{-\infty}^{\infty} U^2\, dx}$$

$$= \frac{\hbar\mathbf{k}_x \int_{-\infty}^{\infty} U^2\, dx}{\int_{-\infty}^{\infty} U^2\, dx} = \hbar\mathbf{k}_x$$

With these limits of integration, both the numerator and denominator are infinite. For problems of this type, one integrates between the finite limits $-L/2$ and $+L/2$ and, in the final result, then assumes that L approaches infinity.

This result implies that $(E, \mathbf{k})$ diagrams such as shown in Fig. 3–5 can be considered plots of electron energy vs. momentum, with a scaling factor $\hbar$.

The direct and indirect semiconductors are identified in Appendix III In a direct semiconductor such as GaAs, an electron in the conduction band can fall to an empty state in the valence band, giving off the energy difference E_g as a photon of light. On the other hand, an electron in the conduction band minimum of an indirect semiconductor such as Si cannot fall directly to the valence band maximum but must undergo a momentum change as well as changing its energy. For example, it may go through some defect state (E_t) within the band gap. We shall discuss such defect states in Sections 4.2.1 and 4.3.2. In an indirect transition which involves a change in $\mathbf{k}$, part of the energy is generally given up as heat to the lattice rather than as an emitted photon. This difference between direct and indirect band structures is very important for deciding which semiconductors can be used in devices requiring light output. For example, semiconductor light emitters and lasers (Chapter 8) generally must be made of materials capable of direct band-to-band transitions or of indirect materials with vertical transitions between defect states.

Band diagrams such as those shown in Fig. 3–5 are cumbersome to draw in analyzing devices, and do not provide a view of the variation of electron energy with distance in the sample. Therefore, in most discussions we shall use simple band pictures such as those shown in Fig. 3–4, remembering that electron transitions across the band gap may be direct or indirect.

CIRRUS EX. 1062 - 21/190

3.1.5 Variation of Energy Bands with Alloy Composition

As III–V ternary and quaternary alloys are varied over their composition ranges (see Sections 1.2.4 and 1.4.1), their band structures change. For example, Fig. 3–6 illustrates the band structure of GaAs and AlAs, and the way in which the bands change with composition x in the ternary compound $Al_xGa_{1-x}As$. The binary compound GaAs is a direct material, with a band gap of 1.43 eV at room temperature. For reference, we call the direct ($\mathbf{k} = 0$) conduction band minimum Γ. There are also two higher-lying indirect minima in the GaAs conduction band, but these are sufficiently far above Γ that few electrons reside there (we discuss an important exception in Chapter 10 in which high-field excitation of electrons into the indirect minima leads to the Gunn effect). We call the lowest-lying GaAs indirect minimum L and the other X. In AlAs the direct Γ minimum is much higher than the indirect X minimum, and this material is therefore indirect with a band gap of 2.16 eV at room temperature.

In the ternary alloy $Al_xGa_{1-x}As$ all of these conduction band minima move up relative to the valence band as the composition x varies from 0 (GaAs) to 1 (AlAs). However, the indirect minimum X moves up less than the others, and for compositions above about 38 percent Al this indirect minimum becomes the lowest-lying conduction band. Therefore, the ternary alloy AlGaAs is a direct semiconductor for Al compositions on the column III sublattice up to about 38 percent, and is an indirect semiconductor for higher Al mole fractions. The band gap energy E_g is shown in color on Fig. 3–6c.

The variation of energy bands for the ternary alloy $GaAs_{1-x}P_x$ is generally similar to that of AlGaAs shown in Fig. 3–6. GaAsP is a direct semiconductor from GaAs to about $GaAs_{.55}P_{.45}$ and is indirect from this composition to GaP (see Fig. 8–11). This material is often used in visible LEDs.

Since light emission is most efficient for direct materials, in which electrons can drop from the conduction band to the valence band without changing $\mathbf{k}$ (and therefore momentum), LEDs in GaAsP are generally made in material grown with a composition less than $x = 0.45$. For example, most red LEDs in this material are made at about $x = 0.4$, where the Γ minimum is still the lowest-lying conduction band edge, and where the photon resulting from a direct transition from this band to the valence band is in the red portion of the spectrum (about 1.9 eV). The use of impurities to enhance radiative recombination in indirect material will be discussed in Section 8.2.

3.2 CHARGE CARRIERS IN SEMICONDUCTORS

The mechanism of current conduction is relatively easy to visualize in the case of a metal; the metal atoms are imbedded in a "sea" of relatively free electrons, and these electrons can move as a group under the influence of an electric field. This free electron view is oversimplified, but many important conduction properties of metals can be derived from just such a model. However, we cannot account for all of the electrical properties of semiconductors

CIRRUS EX. 1062 - 22/190

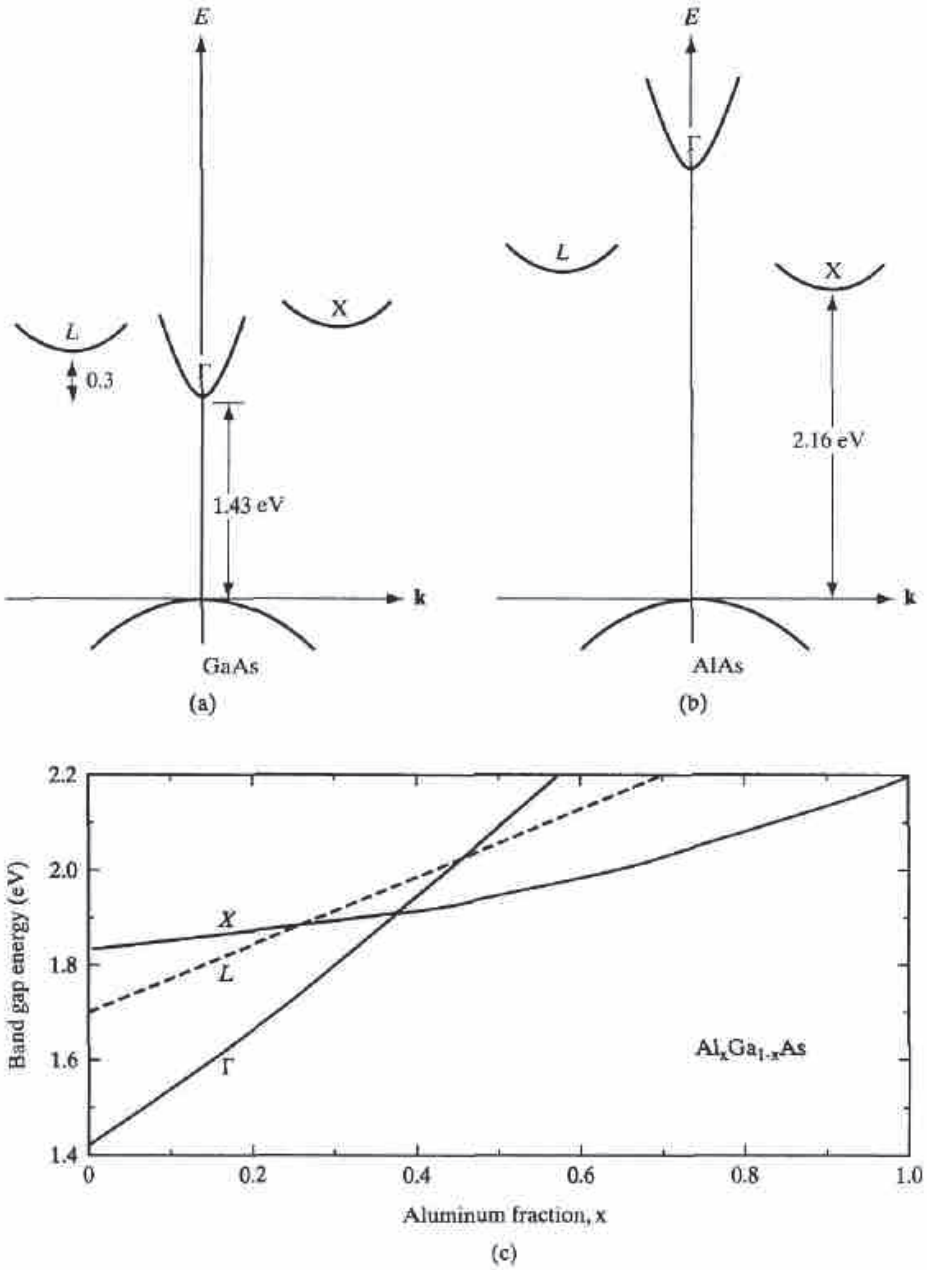


Figure 3–6
Variation of direct and indirect conduction bands in AlGaAs as a function of composition: (a) the (E,**k**) diagram for GaAs, showing three minima in the conduction band; (b) AlAs band diagram; (c) positions of the three conduction band minima in $Al_xGa_{1-x}As$ as x varies over the range of compositions from GaAs ($x = 0$) to AlAs ($x = 1$). The smallest band gap, E_g (shown in color), follows the direct Γ band to $x = 0.38$, and then follows the indirect X band.

CIRRUS EX. 1062 - 23/190

in this way. Since the semiconductor has a filled valence band and an empty conduction band at 0 K, we must consider the increase in conduction band electrons by thermal excitations across the band gap as the temperature is raised. In addition, after electrons are excited to the conduction band, the empty states left in the valence band can contribute to the conduction process. The introduction of impurities has an important effect on the energy band structure and on the availability of charge carriers. Thus there is considerable flexibility in controlling the electrical properties of semiconductors.

3.2.1 Electrons and Holes

As the temperature of a semiconductor is raised from 0 K, some electrons in the valence band receive enough thermal energy to be excited across the band gap to the conduction band. The result is a material with some electrons in an otherwise empty conduction band and some unoccupied states in an otherwise filled valence band (Fig. 3-7).[2] For convenience, an empty state in the valence band is referred to as a *hole*. If the conduction band electron and the hole are created by the excitation of a valence band electron to the conduction band, they are called an *electron–hole pair* (abbreviated EHP).

After excitation to the conduction band, an electron is surrounded by a large number of unoccupied energy states. For example, the equilibrium number of electron-hole pairs in pure Si at room temperature is only about 10^{10} EHP/cm^3, compared to the Si atom density of 5×10^{22} atoms/cm^3. Thus the few electrons in the conduction band are free to move about via the many available empty states.

The corresponding problem of charge transport in the valence band is somewhat more complicated. However, it is possible to show that the effects of current in a valence band containing holes can be accounted for by simply keeping track of the holes themselves.

In a filled band, all available energy states are occupied. For every electron moving with a given velocity, there is an equal and opposite electron motion elsewhere in the band. If we apply an electric field, the net current is zero

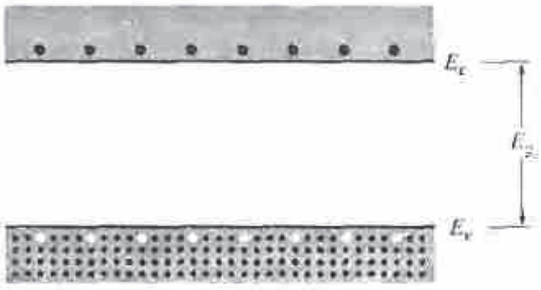


Figure 3–7
Electron–hole pairs in a semiconductor.

[2] In Fig. 3–7 and in subsequent discussions, we refer to the bottom of the conduction band as E_c and the top of the valence band as E_v.

CIRRUS EX. 1062 - 24/190

because for every electron j moving with velocity v_j there is a corresponding electron j' with velocity $-\mathsf{v}_j$. Figure 3-8 illustrates this effect in terms of the electron energy vs. wave vector plot for the valence band. Since $\mathbf{k}$ is proportional to electron momentum, it is clear the two electrons have oppositely directed velocities. With N electrons/cm^3 in the band we express the current density using a sum over all of the electron velocities, and including the charge $-q$ on each electron. In a unit volume,

$$J = (-q)\sum_{i}^{N} \mathsf{v}_i = 0 \quad (\textit{filled band}) \tag{3-2a}$$

Now if we create a hole by removing the jth electron, the net current density in the valence band involves the sum over all velocities, minus the contribution of the electron we have removed:

$$J = (-q)\sum_{i}^{N} \mathsf{v}_i - (-q)\mathsf{v}_j \quad (j\text{th } \textit{electron missing}) \tag{3-2b}$$

But the first term is zero, from Eq. (3-2a). Thus the net current is $+q\mathsf{v}_j$. In other words, the current contribution of the hole is equivalent to that of a positively charged particle with velocity v_j, that of the missing electron. Of course, the charge transport is actually due to the motion of the new uncompensated electron (j'). Its current contribution $(-q)\,(-\mathsf{v}_j)$ is equivalent to that of a positively charged particle with velocity $+\mathsf{v}_j$. For simplicity, it is customary to treat empty states in the valence band as charge carriers with positive charge and positive mass.

A simple analogy may help in understanding the behavior of holes. If we have two bottles, one completely filled with water and one completely empty, we can ask ourselves "Will there be any net transport of water when we tilt the bottles?" The answer is "no". In the case of the empty bottle, the answer is obvious. In the case of the completely full bottle also, there cannot be any net motion of water because there is no empty space for water to move into. Similarly, an empty conduction band completely devoid of electrons or a valence band completely full of electrons cannot give rise to a net motion of electrons, and thus to current conduction.

Next, we imagine transferring some water droplets from the full bottle into the empty bottle, leaving behind some air bubbles, and ask ourselves the same question. Now when we tilt the bottles there will be net transport of water: the water droplets will roll downhill in one bottle and the air bubbles will move uphill in the other. Similarly, a few electrons in an otherwise empty conduction band move opposite to an electric field, while holes in an otherwise filled valence band move in the direction of the field. The bubble analogy is imperfect, but it may provide a physical feel for why the charge and mass of a hole have opposite signs from those of an electron.

In all the following discussions we shall concentrate on the electrons in the conduction band and on the holes in the valence band. We can account for the current flow in a semiconductor by the motion of these two types of

CIRRUS EX. 1062 - 25/190

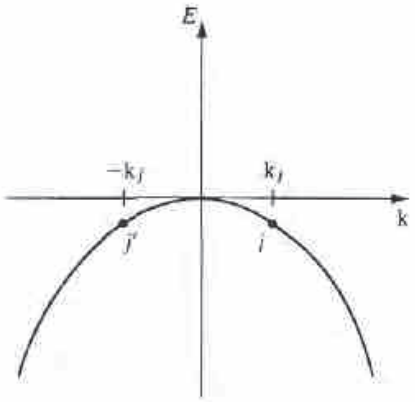


Figure 3–8
A valence band with all states filled, including states j and j', marked for discussion. The jth electron with wave vector k_j is matched by an electron at j' with the opposite wave vector $-k_j$. There is no net current in the band unless an electron is removed. For example, if the jth electron is removed, the motion of the electron at j' is no longer compensated.

charge carriers. We draw valence and conduction bands on an electron energy scale E, as in Fig. 3-8. However, we should remember that in the valence band, hole energy increases oppositely to electron energy, because the two carriers have opposite charge. Thus hole energy increases downward in Fig. 3-8 and holes, seeking the lowest energy state available, are generally found at the *top* of the valence band. In contrast, conduction band electrons are found at the bottom of the conduction band.

It would be instructive to compare the $(E, \mathbf{k})$ band diagrams with the "simplified" band diagrams that are used for routine device analysis (Fig. 3-9). As discussed in Examples 3-1 and 3-2, an $(E, \mathbf{k})$ diagram is a plot of the total electron energy (potential plus kinetic) as a function of the crystal-direction-dependent electron wave vector (which is proportional to the momentum and therefore the velocity) at some point in space. Hence, the bottom of the conduction band corresponds to zero electron velocity or kinetic energy, and simply gives us the potential energy at that point in space. For holes, the top of the valence band corresponds to zero kinetic energy. For simplified band diagrams, we plot the edges of the conduction and valence bands (i.e., the potential energy) as a function of position in the device. Energies higher in the band correspond to additional kinetic energy of the electron. Also, the fact that the band edge corresponds to the electron potential energy tells us that the variation of the band edge in space is related to the electric field at different points in the semiconductor. We will show this relationship explicitly in Section 4.4.2.

In Fig. 3-9, an electron at location A sees an electric field given by the slope of the band edge (potential energy), and gains kinetic energy (at the expense of potential energy) by moving to point B. Correspondingly, in the $(E, \mathbf{k})$ diagram, the electron starts at $k = 0$, but moves to a nonzero wave vector $\mathbf{k}_B$.

CIRRUS EX. 1062 - 26/190

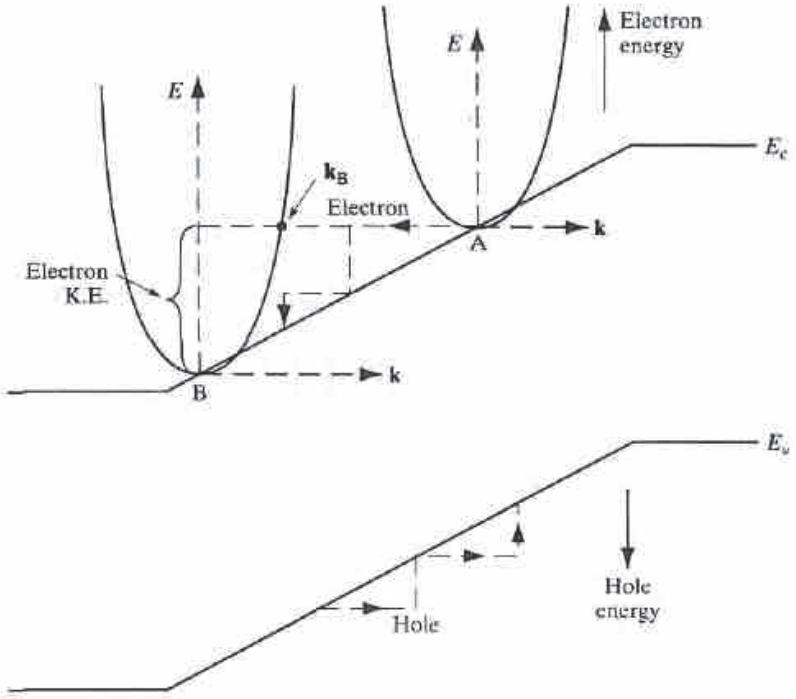


Figure 3–9 Superimposition of the (*E*,**k**) band structure on the *E*-versus-position simplified band diagram for a semiconductor in an electric field. Electron energies increase going up, while hole energies increase going down. Similarly, electron and hole wave vectors point in opposite directions and these charge carriers move opposite to each other, as shown.

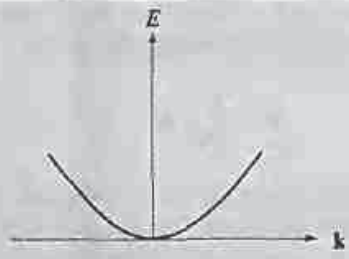
The electron then loses kinetic energy to heat by scattering mechanisms (discussed in Section 3.4.3) and returns to the bottom of the band at B. The slopes of the (E, x) band edges at different points in space reflect the local electric fields at those points. In practice, the electron may lose its kinetic energy in stages by a series of scattering events, as shown by the colored dashed lines.

3.2.2 Effective Mass

The electrons in a crystal are not completely free, but instead interact with the periodic potential of the lattice. As a result, their "wave-particle" motion cannot be expected to be the same as for electrons in free space. Thus, in applying the usual equations of electrodynamics to charge carriers in a solid, we must use altered values of particle mass. In doing so, we account for most of the influences of the lattice, so that the electrons and holes can be treated as "almost free" carriers in most computations. The calculation of effective mass must take into account the shape of the energy bands in three-dimensional **k**-space, taking appropriate averages over the various energy bands.

EXAMPLE 3–2 Find the (E, **k**) relationship for a free electron and relate it to the electron mass.

CIRRUS EX. 1062 - 27/190



From Example 3–1, the electron momentum is $\mathbf{p} = m\mathbf{v} = \hbar\mathbf{k}$. Then **SOLUTION**

$$E = \frac{1}{2}m\mathbf{v}^2 = \frac{1}{2}\frac{\mathbf{p}^2}{m} = \frac{\hbar^2}{2m}\mathbf{k}^2$$

Thus the electron energy is parabolic with wave vector $\mathbf{k}$. The electron mass is inversely related to the curvature (second derivative) of the (E, $\mathbf{k}$) relationship, since

$$\frac{d^2E}{d\mathbf{k}^2} = \frac{\hbar^2}{m}$$

Although electrons in solids are not free, most energy bands are close to parabolic at their minima (for conduction bands) or maxima (for valence bands). We can also approximate effective mass near those band extrema from the curvature of the band.

The effective mass of an electron in a band with a given (E, $\mathbf{k}$) relationship is found in Example 3–2 to be

$$m^* = \frac{\hbar^2}{d^2E/d\mathbf{k}^2} \tag{3–3}$$

Thus the curvature of the band determines the electron effective mass. For example, in Fig. 3–6a it is clear that the electron effective mass in GaAs is much smaller in the direct Γ conduction band (strong curvature) than in the L or X minima (weaker curvature, smaller value in the denominator of the m^* expression).

A particularly interesting feature of Figs. 3–5 and 3–6 is that the curvature of $d^2E/d\mathbf{k}^2$ is positive at the conduction band minima, but is negative at the valence band maxima. Thus, the electrons near the top of the valence band have *negative effective mass*, according to Eq. (3–3). Valence band electrons with negative charge and negative mass move in an electric field in the same direction as holes with positive charge and positive mass. As discussed in Section 3.2.1, we can fully account for charge transport in the valence band by considering hole motion.

CIRRUS EX. 1062 - 28/190

For a band centered at $\mathbf{k} = 0$ (such as the Γ band in GaAs), the $(E, \mathbf{k})$ relationship near the minimum is usually parabolic:

$$E = \frac{\hbar^2}{2m^*}\mathbf{k}^2 + E_c \tag{3-4}$$

Comparing this relation to Eq. (3-3) indicates that the effective mass m^* is constant in a parabolic band. On the other hand, many conduction bands have complex $(E, \mathbf{k})$ relationships that depend on the direction of electron transport with respect to the principal crystal directions. In this case, the effective mass is a tensor quantity. However, we can use appropriate averages over such bands in most calculations.

Figure 3-10a shows the band structures for Si and GaAs viewed along two major directions. While the shape is parabolic near the band edges (as indicated in Figure 3-5 and Example 3-2), there are significant non-parabolicities at higher energies. The energies are plotted along the high symmetry [111] and [100] directions in the crystal. The $\mathbf{k} = 0$ point is denoted as Γ. When we go along

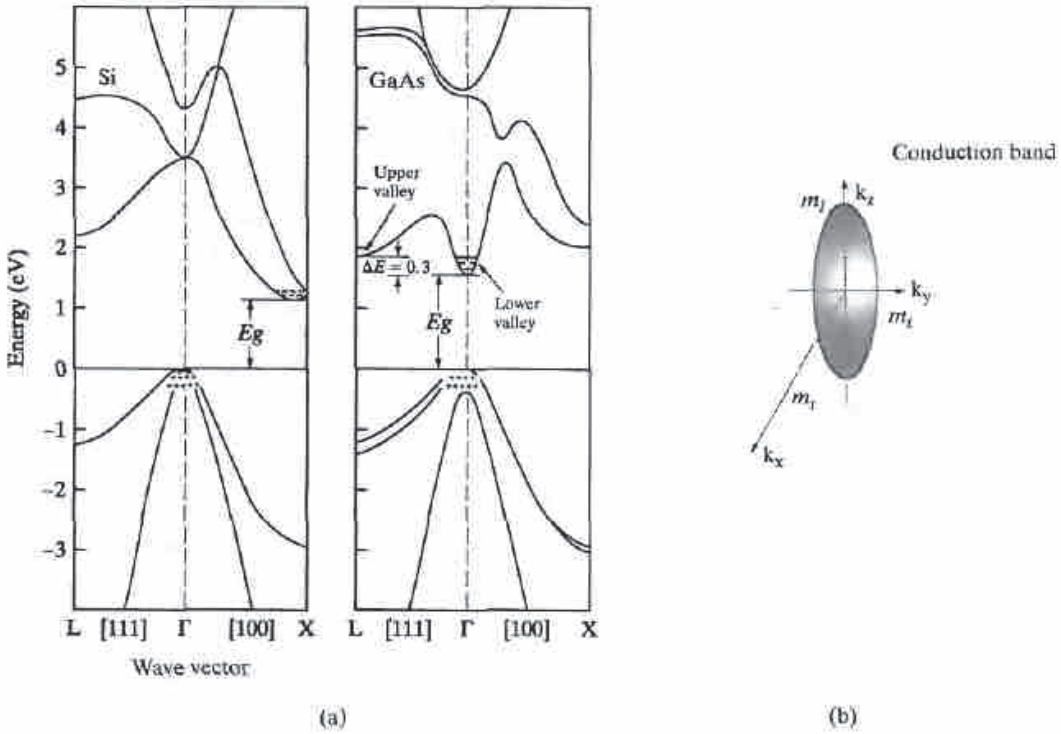


Figure 3–10
Realistic band structures in semiconductors: (a) Conduction and valence bands in Si and GaAs along [111] and [100]; (b) ellipsoidal constant energy surface for Si, near the 6 conduction band minima along the *X* directions. (From Chelikowsky and Cohen, Phys. Rev. B14, 556, 1976).

CIRRUS EX. 1062 - 29/190

the [100] direction, we reach a valley near X, while we reach the L valley along the [111] direction. (Since the energies are plotted along different directions, the curves do not look symmetric.) The valence band maximum in most semiconductors is at the Γ point. It has three branches: the *heavy hole* band with the smallest curvature, a *light hole band* with a larger curvature, and a *split-off band* at a different energy. We notice that for GaAs the conduction band minimum and the valence band maximum are both at $\mathbf{k} = 0$; therefore it is direct band gap. Silicon, on the other hand, has 6 equivalent conduction minima at X along the 6 equivalent $\langle 100 \rangle$ directions; therefore, it is indirect.

Figure 3-10b shows the constant energy surface for electrons in one of the six conduction bands for Si. The way to relate these surfaces to the band structures shown in Fig. 3-10a is to consider a certain value of energy, and determine all the $\mathbf{k}$ vectors in 3 dimensions for which we get this energy. We find that for Si we have 6 cigar-shaped ellipsoidal equi-energy surfaces near the conduction band minima along the six equivalent X-directions, with a longitudinal effective mass, m_l, along the major axis, and two transverse effective masses, m_t, along the minor axes. For GaAs, the conduction band is more or less spherical for low energies. On the other hand, we have warped spherical surfaces in the valence band. The importance of these surfaces will be clear in Sections 3.3.2 and 3.4.1 when we consider different types of effective masses in semiconductors.

In any calculation involving the mass of the charge carriers, we must use effective mass values for the particular material involved. In all subsequent discussions, the electron effective mass is denoted by m_n^* and the hole effective mass by m_p^*. The n subscript indicates the electron as a negative charge carrier, and the p subscript indicates the hole as a positive charge carrier.

There is nothing mysterious about the concept of an "effective" mass, m_n^*, and about the fact that it is different in different semiconductors. Indeed, the "true" mass of an electron, m, is the same in Si, Ge, or GaAs—it is the same as for a free electron in vacuum. To understand why the effective mass is different from the true mass, consider Newton's second law, which states that the time rate of change of momentum is the force.

$$dp/dt = d(m\mathbf{v})/dt = \text{Force} \tag{3-5a}$$

An electron in a crystal experiences a total force $\mathrm{F_{int}} + \mathrm{F_{ext}}$, where $\mathrm{F_{int}}$ is the collection of internal periodic crystal forces, and $\mathrm{F_{ext}}$ is the externally applied force. It is inefficient to solve this complicated problem involving the periodic crystal potential (which is obviously different in different semiconductors) every time we try to solve a semiconductor device problem. It is better to solve the complicated problem of carrier motion in the periodic crystal potential just once, and encapsulate that information in what is called the bandstructure, $(E, \mathbf{k})$, whose curvature gives us the effective mass, m_n^*. The electron then responds to external forces with this new m_n^*. Newton's law is then written as:

$$d(m_n^*\mathbf{v})/dt = \mathrm{F_{ext}} \tag{3-5b}$$

CIRRUS EX. 1062 - 30/190

This is clearly an enormous simplification compared to the more detailed problem. Obviously, the periodic crystal forces depend on the details of a specific semiconductor; therefore, the effective mass is different in different materials.

Once we determine the band curvature effective mass components from the orientation-dependent $(E, \mathbf{k})$, we have to combine them appropriately for different types of calculations. We shall see in Section 3.3.2 that when we are interested in determining the numbers of carriers in the bands, we have to use a "density-of-states" effective mass by taking the geometric mean of the band curvature effective masses, and the number of equivalent band extrema. On the other hand we will find in Section 3.4.1 that in problems involving the motion of carriers, one must take the harmonic mean of the band curvature effective masses to get the "conductivity" effective mass.

3.2.3 Intrinsic Material

A perfect semiconductor crystal with no impurities or lattice defects is called an *intrinsic* semiconductor. In such material there are no charge carriers at 0 K, since the valence band is filled with electrons and the conduction band is empty. At higher temperatures electron-hole pairs are generated as valence band electrons are excited thermally across the band gap to the conduction band. These EHPs are the only charge carriers in intrinsic material.

The generation of EHPs can be visualized in a qualitative way by considering the breaking of covalent bonds in the crystal lattice (Fig. 3-11). If one of the Si valence electrons is broken away from its position in the bonding structure such that it becomes free to move about in the lattice, a conduction electron is created and a broken bond (hole) is left behind. The energy required to break the bond is the band gap energy E_g. This model helps in visualizing the physical mechanism of EHP creation, but the energy band model is more productive for purposes of quantitative calculation. One important difficulty in the "broken bond" model is that the free electron and the hole seem deceptively localized in the lattice. Actually, the positions of the free electron and the hole are spread out over several lattice spacings and should be considered quantum mechanically by probability distributions (see Section 2.4).

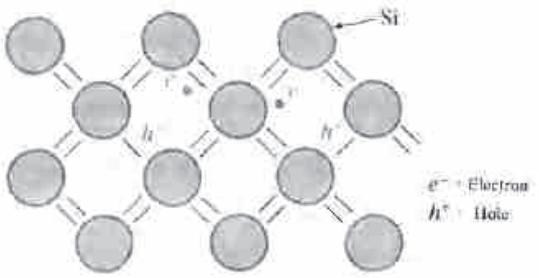


Figure 3–11 Electron–hole pairs in the covalent bonding model of the Si crystal.

CIRRUS EX. 1062 - 31/190

Since the electrons and holes are created in pairs, the conduction band electron concentration n (electrons per cm^3) is equal to the concentration of holes in the valence band p (holes per cm^3). Each of these intrinsic carrier concentrations is commonly referred to as n_i. Thus *for intrinsic material*

$$n = p = n_i \tag{3-6}$$

At a given temperature there is a certain concentration of electron - hole pairs n_i. Obviously, if a steady state carrier concentration is maintained, there must be *recombination* of EHPs at the same rate at which they are generated. Recombination occurs when an electron in the conduction band makes a transition (direct or indirect) to an empty state (hole) in the valence band, thus annihilating the pair. If we denote the generation rate of EHPs as g_i (EHP/cm^3-s) and the recombination rate as r_i, equilibrium requires that

$$r_i = g_i \tag{3-7a}$$

Each of these rates is temperature dependent. For example, $g_i(T)$ increases when the temperature is raised, and a new carrier concentration n_i is established such that the higher recombination rate $r_i(T)$ just balances generation. At any temperature, we can predict that the rate of recombination of electrons and holes r_i is proportional to the equilibrium concentration of electrons n_0 and the concentration of holes p_0:

$$r_i = \alpha_r n_0 p_0 = \alpha_r n_i^2 = g_i \tag{3-7b}$$

The factor α_r is a constant of proportionality which depends on the particular mechanism by which recombination takes place. We shall discuss the calculation of n_i as a function of temperature in Section 3.3.3; recombination processes will be discussed in Chapter 4.

3.2.4 Extrinsic Material

In addition to the intrinsic carriers generated thermally, it is possible to create carriers in semiconductors by purposely introducing impurities into the crystal. This process, called *doping*, is the most common technique for varying the conductivity of semiconductors. By doping, a crystal can be altered so that it has a predominance of either electrons or holes. Thus there are two types of doped semiconductors, n-type (mostly electrons) and p-type (mostly holes). When a crystal is doped such that the equilibrium carrier concentrations n_0 and p_0 are different from the intrinsic carrier concentration n_i, the material is said to be *extrinsic*.

When impurities or lattice defects are introduced into an otherwise perfect crystal, additional levels are created in the energy band structure, usually within the band gap. For example, an impurity from column V of the periodic table (P, As, and Sb) introduces an energy level very near the conduction band in Ge or Si. This level is filled with electrons at 0 K, and very little thermal energy is required to excite these electrons to the conduction band (Fig. 3-12a). Thus at about 50–100 K virtually all of the electrons in the

CIRRUS EX. 1062 - 32/190

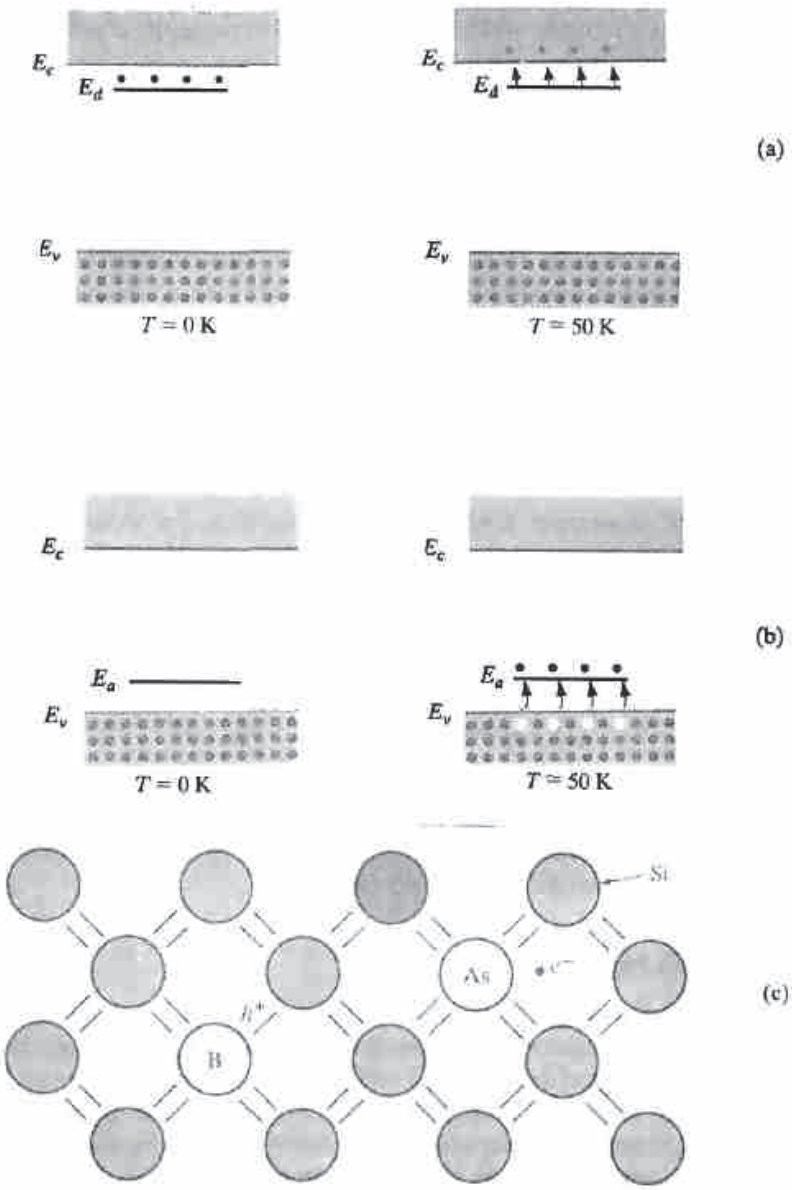


Figure 3–12
Energy band model and chemical bond model of dopants in semiconductors: (a) donation of electrons from donor level to conduction band; (b) acceptance of valence band electrons by an acceptor level, and the resulting creation of holes; (c) donor and acceptor atoms in the covalent bonding model of a Si crystal.

impurity level are "donated" to the conduction band. Such an impurity level is called a *donor* level, and the column V impurities in Ge or Si are called donor impurities. From Fig. 3–12a we note that the material doped with donor impurities can have a considerable concentration of electrons in the conduction band, even when the temperature is too low for the intrinsic EHP concentration to be appreciable. Thus semiconductors doped with a significant number of donor atoms will have $n_0 \gg (n_i, p_0)$ at room temperature. This is n-type material.

CIRRUS EX. 1062 - 33/190

Atoms from column III (B, Al, Ga, and In) introduce impurity levels in Ge or Si near the valence band. These levels are empty of electrons at 0 K (Fig. 3-12b). At low temperatures, enough thermal energy is available to excite electrons from the valence band into the impurity level, leaving behind holes in the valence band. Since this type of impurity level "accepts" electrons from the valence band, it is called an *acceptor* level, and the column III impurities are acceptor impurities in Ge and Si. As Fig. 3-12b indicates, doping with acceptor impurities can create a semiconductor with a hole concentration p_0 much greater than the conduction band electron concentration n_0 (this type is p-type material).

In the covalent bonding model, donor and acceptor atoms can be visualized as shown in Fig. 3-12c. An As atom (column V) in the Si lattice has the four necessary valence electrons to complete the covalent bonds with the neighboring Si atoms, plus one extra electron. This fifth electron does not fit into the bonding structure of the lattice and is therefore loosely bound to the As atom. A small amount of thermal energy enables this extra electron to overcome its coulombic binding to the impurity atom and be donated to the lattice as a whole. Thus it is free to participate in current conduction. This process is a qualitative model of the excitation of electrons out of a donor level and into the conduction band (Fig. 3-12a). Similarly, the column III impurity B has only three valence electrons to contribute to the covalent bonding (Fig. 3-12c), thereby leaving one bond incomplete. With a small amount of thermal energy, this incomplete bond can be transferred to other atoms as the bonding electrons exchange positions. Again, the idea of an electron "hopping" from an adjacent bond into the incomplete bond at the B site provides some physical insight into the behavior of an acceptor, but the model of Fig. 3-12b is preferable for most discussions.

We can calculate rather simply the approximate energy required to excite the fifth electron of a donor atom into the conduction band (the donor *binding energy*). Let us assume for rough calculations that the As atom of Fig. 3-12c has its four covalent bonding electrons rather tightly bound and the fifth "extra" electron loosely bound to the atom. We can approximate this situation by using the Bohr model results, considering the loosely bound electron as ranging about the tightly bound "core" electrons in a hydrogen-like orbit. From Eq. (2-15) the magnitude of the ground-state energy ($\mathbf{n} = 1$) of such an electron is

$$E = \frac{mq^4}{2K^2\hbar^2} \tag{3-8}$$

The value of K must be modified from the free-space value $4\pi\epsilon_0$ used in the hydrogen atom problem to

$$K = 4\pi\epsilon_0\,\epsilon_r \tag{3-9}$$

where ϵ_r is the relative dielectric constant of the semiconductor material. In addition, we must use the conductivity effective mass m_n^* typical of the semiconductor, discussed in more detail in Section 3.4.1.

CIRRUS EX. 1062 - 34/190

EXAMPLE 3–3 It was mentioned in Section 3.2 that the covalent bonding model gives a false impression of the localization of carriers. As an illustration, calculate the radius of the electron orbit around the donor in Fig. 3–12c, assuming a ground state hydrogen-like orbit in Si. Compare with the Si lattice constant. Use $m_n^* = 0.26m_0$ for Si.

SOLUTION From Eq. (2–10) with $n = 1$ and using $\epsilon_r = 11.8$ for Si,

$$r = \frac{4\pi\epsilon_r\epsilon_0\hbar^2}{m_n^* q^2} = \frac{11.8(8.85 \times 10^{-12})(6.63 \times 10^{-34})^2}{\pi(0.26)(9.11 \times 10^{-31})(1.6 \times 10^{-19})^2}$$

$$r = 2.41 \times 10^{-9} m = 24.1 \text{ Å}.$$

Note that this is more than four lattice spacings $a = 5.43$ Å.

Generally, the column V donor levels lie approximately 0.01 eV below the conduction band in Ge, and the column III acceptor levels lie about 0.01 eV above the valence band. In Si the usual donor and acceptor levels lie about 0.03–0.06 eV from a band edge.

In III–V compounds, column VI impurities occupying column V sites serve as donors. For example, S, Se, and Te are donors in GaAs, since they substitute for As and provide an extra electron compared with the As atom. Similarly, impurities from column II (Be, Zn, Cd) substitute for column III atoms to form acceptors in the III–V compounds. A more ambiguous case arises when a III–V material is doped with Si or Ge, from column IV. These impurities are called *amphoteric*, meaning that Si or Ge can serve as donors or acceptors depending on whether they reside on the column III or column V sublattice of the crystal. In GaAs it is common for Si impurities to occupy Ga sites. Since the Si has an extra electron compared with the Ga it replaces, it serves as a donor. However, an excess of As vacancies arising during growth or processing of the GaAs can cause Si impurities to occupy As sites, where they serve as acceptors.

The importance of doping will become obvious when we discuss electronic devices made from junctions between p-type and n-type semiconductor material. The extent to which doping controls the electronic properties of semiconductors can be illustrated here by considering changes in the sample resistance which occur with doping. In Si, for example, the intrinsic carrier concentration n_i is about 10^{10} cm^{-3} at room temperature. If we dope Si with 10^{15} As atoms/cm^3, the conduction electron concentration changes by five orders of magnitude. The resistivity of Si changes from about 2×10^5 Ω-cm to 5 Ω-cm with this doping.

When a semiconductor is doped n-type or p-type, one type of carrier dominates. In the example given above, the conduction band electrons outnumber the holes in the valence band by many orders of magnitude. We refer

CIRRUS EX. 1062 - 35/190

to the small number of holes in n-type material as *minority carriers* and the relatively large number of conduction band electrons as *majority carriers*. Similarly, electrons are the minority carriers in p-type material, and holes are the majority carriers.

3.2.5 Electrons and Holes in Quantum Wells

We have discussed single-valued (*discrete*) energy levels in the band gap arising from doping, and a *continuum* of allowed states in the valence and conduction bands. A third possibility is the formation of discrete levels for electrons and holes as a result of quantum-mechanical confinement.

One of the most useful applications of MBE or OMVPE growth of multi-layer compound semiconductors, as described in Section 1.4, is the fact that a continuous single crystal can be grown in which adjacent layers have different band gaps. For example, Fig. 3–13 shows the spatial variation in conduction and valence bands for a multilayer structure in which a very thin layer of GaAs is sandwiched between two layers of AlGaAs, which has a wider band gap that the GaAs. We will discuss the details of such *heterojunctions* (junctions between dissimilar materials) in Section 5.8. It is interesting to point out here, however, that a consequence of confining electrons and holes in a very thin layer is that these particles behave according to the *particle in a potential well* problem, with quantum states calculated in Section 2.4.3. Therefore,

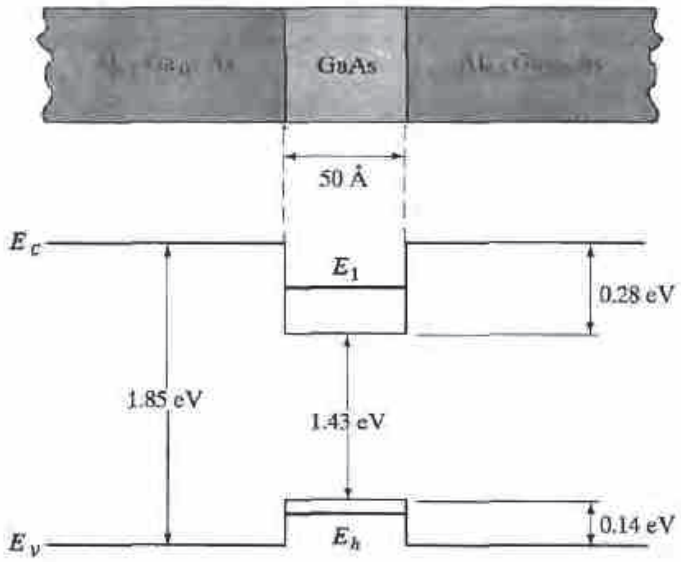


Figure 3–13
Energy band discontinuities for a thin layer of GaAs sandwiched between layers of wider band gap AlGaAs. In this case, the GaAs region is so thin that quantum states are formed in the valence and conduction bands. Electrons in the GaAs conduction band reside on "particle in a potential well" states such as E_1 shown here, rather than in the usual conduction band states. Holes in the quantum well occupy similar discrete states, such as E_h.

CIRRUS EX. 1062 - 36/190

instead of having the continuum of states normally available in the conduction band, the conduction band electrons in the narrow-gap material are confined to discrete quantum states as described by Eq. (2-33), modified for effective mass and finite barrier height. Similarly, the states in the valence band available for holes are restricted to discrete levels in the quantum well. This is one of the clearest demonstrations of the quantum mechanical results discussed in Chapter 2. From a practical device point of view, the formation of discrete quantum states in the GaAs layer of Fig. 3-13 changes the energy at which photons can be emitted. An electron on one of the discrete conduction band states (E_1 in Fig. 3-13) can make a transition to an empty discrete valence band state in the GaAs quantum well (such as E_h), giving off a photon of energy $E_g + E_1 + E_h$, greater than the GaAs band gap. Semiconductor lasers have been made in which such a quantum well is used to raise the energy of the transition from the infrared, typical of GaAs, to the red portion of the spectrum. We will see other examples of quantum wells in semiconductor devices in later chapters.

3.3 CARRIER CONCENTRATIONS

In calculating semiconductor electrical properties and analyzing device behavior, it is often necessary to know the number of charge carriers per cm^3 in the material. The majority carrier concentration is usually obvious in heavily doped material, since one majority carrier is obtained for each impurity atom (for the standard doping impurities). The concentration of minority carriers is not obvious, however, nor is the temperature dependence of the carrier concentrations.

To obtain equations for the carrier concentrations we must investigate the distribution of carriers over the available energy states. This type of distribution is not difficult to calculate, but the derivation requires some background in statistical methods. Since we are primarily concerned here with the application of these results to semiconductor materials and devices, we shall accept the distribution function as given.

3.3.1 The Fermi Level

Electrons in solids obey *Fermi-Dirac* statistics.[3] In the development of this type of statistics, one must consider the indistinguishability of the electrons,

[3]Examples of other types of statistics are *Maxwell-Boltzmann* for classical particles (e.g., gas) and *Bose-Einstein* for photons. For two discrete energy levels, E_2 and E_1 (with $E_2 > E_1$), classical gas atoms follow a Boltzmann distribution; the number n_2 of atoms in state E_2 is related to the number n_1 in E_1 at thermal equilibrium by

$$\frac{n_2}{n_1} = \frac{N_2 e^{-E_2/kT}}{N_1 e^{-E_1/kT}} = \frac{N_2}{N_1} e^{-(E_2 - E_1)/kT}$$

assuming the two levels have N_2 and N_1 number of states, respectively. The exponential term $\exp[-\Delta E/kT]$ is commonly called the *Boltzmann factor*. It appears also in the denominator of the Fermi-Dirac distribution function. We shall return to the Boltzmann distribution in Chapter 8 in discussions of the properties of lasers.

CIRRUS EX. 1062 - 37/190

their wave nature, and the Pauli exclusion principle. The rather simple result of these statistical arguments is that the distribution of electrons over a range of allowed energy levels at thermal equilibrium is

$$\boxed{f(E) = \frac{1}{1 + e^{(E-E_F)/kT}}} \tag{3-10}$$

where k is Boltzmann's constant ($k = 8.62 \times 10^{-5}$ eV/K $= 1.38 \times 10^{-23}$ J/K). The function $f(E)$, the *Fermi–Dirac distribution function*, gives the probability that an available energy state at E will be occupied by an electron at absolute temperature T. The quantity E_F is called the *Fermi level*, and it represents an important quantity in the analysis of semiconductor behavior. We notice that, for an energy E equal to the Fermi level energy E_F, the occupation probability is

$$f(E_F) = [1 + e^{(E_F - E_F)/kT}]^{-1} = \frac{1}{1+1} = \frac{1}{2} \tag{3-11}$$

Thus an energy state at the Fermi level has a probability of $^1/_2$ of being occupied by an electron.

A closer examination of $f(E)$ indicates that at 0 K the distribution takes the simple rectangular form shown in Fig. 3-14. With $T = 0$ in the denominator of the exponent, $f(E)$ is $1/(1 + 0) = 1$ when the exponent is negative ($E < E_F$), and is $1/(1 + \infty) = 0$ when the exponent is positive ($E > E_F$). This rectangular distribution implies that at 0 K every available energy state up to E_F is filled with electrons, and all states above E_F are empty.

At temperatures higher than 0 K, some probability exists for states above the Fermi level to be filled. For example, at $T = T_1$ in Fig. 3-14 there is some probability f(E) that states above E_F are filled, and there is a corresponding probability $[1 - f(E)]$ that states below E_F are empty. The Fermi function is symmetrical about E_F for all temperatures; that is, the probability $f(E_F + \Delta E)$ that a state ΔE above E_F is filled is the same as the probability $[1 - f(E_F - \Delta E)]$ that a state ΔE below E_F is empty. The symmetry of the

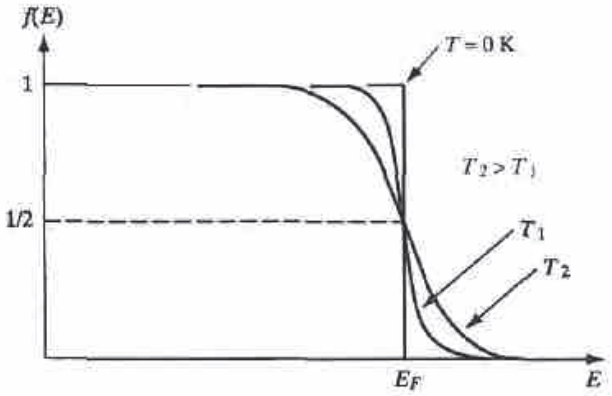


Figure 3–14
The Fermi-Dirac distribution function.

CIRRUS EX. 1062 - 38/190

distribution of empty and filled states about E_F makes the Fermi level a natural reference point in calculations of electron and hole concentrations in semiconductors.

In applying the Fermi–Dirac distribution to semiconductors, we must recall that $f(E)$ is the probability of occupancy of an *available* state at E. Thus if there is no available state at E (e.g., in the band gap of a semiconductor), there is no possibility of finding an electron there. We can best visualize the relation between $f(E)$ and the band structure by turning the $f(E)$ vs. E diagram on its side so that the E scale corresponds to the energies of the band diagram (Fig. 3-15). For intrinsic material we know that the concentration of holes in the valence band is equal to the concentration of electrons in the conduction band. Therefore, the Fermi level E_F must lie at the middle of the band gap in intrinsic material.[4] Since $f(E)$ is symmetrical about E_F, the electron probability "tail" of $f(E)$ extending into the conduction band of Fig. 3-15a is symmetrical with the hole probability tail $[1 - f(E)]$ in the valence band. The distribution function has values within the band gap between

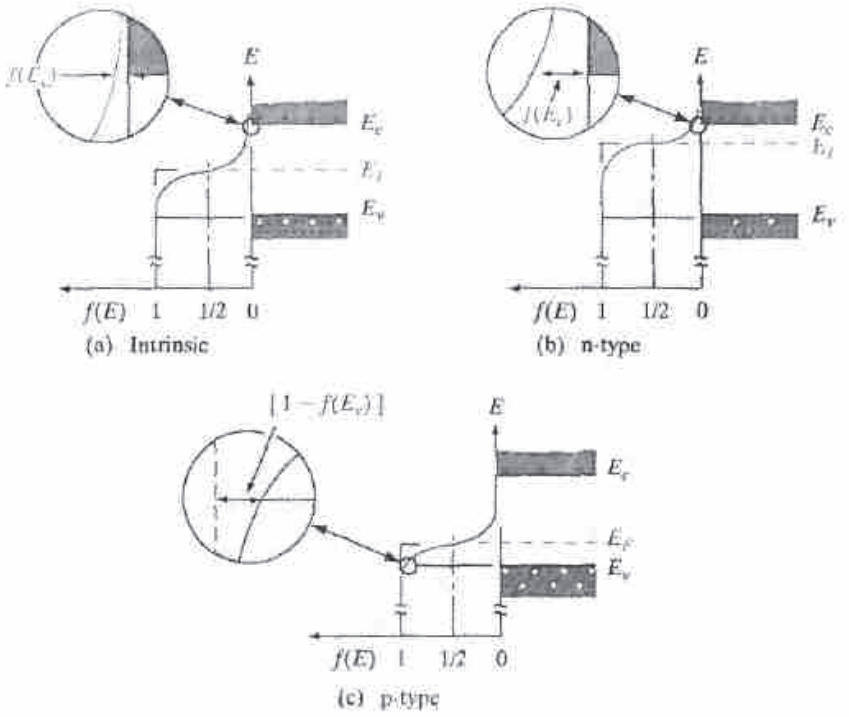


Figure 3–15
The Fermi distribution function applied to semiconductors: (a) intrinsic material; (b) n-type material; (c) p-type material.

[4]Actually the intrinsic E_F is displaced slightly from the middle of the gap, since the densities of available states in the valence and conduction bands are not equal (Section 3.3.2).

CIRRUS EX. 1062 - 39/190

E_v and E_c, but there are no energy states available, and no electron occupancy results from $f(E)$ in this range.

The tails in $f(E)$ are exaggerated in Fig. 3-15 for illustrative purposes. Actually, the probability values at E_v and E_c are quite small for intrinsic material at reasonable temperatures. For example, in Si at 300 K, $n_i = p_i = 10^{10}$ cm^{-3}, whereas the densities of available states at E_v and E_c are on the order of 10^{19} cm^{-3}. Thus the probability of occupancy $f(E)$ for an individual state in the conduction band and the hole probability $[1 - f(E)]$ for a state in the valence band are quite small. Because of the relatively large density of states in each band, small changes in $f(E)$ can result in significant changes in carrier concentration.

In n-type material there is a high concentration of electrons in the conduction band compared with the hole concentration in the valence band (recall Fig. 3-12a). Thus in n-type material the distribution function $f(E)$ must lie above its intrinsic position on the energy scale (Fig. 3-15b). Since $f(E)$ retains its shape for a particular temperature, the larger concentration of electrons at E_c in n-type material implies a correspondingly smaller hole concentration at E_v. We notice that the value of $f(E)$ for each energy level in the conduction band (and therefore the total electron concentration n_0) increases as E_F moves closer to E_c. Thus the energy difference $(E_c - E_F)$ gives a measure of n; we shall express this relation mathematically in the following section.

For p-type material the Fermi level lies near the valence band (Fig. 3-15c) such that the $[1 - f(E)]$ tail below E_v is larger than the $f(E)$ tail above E_c. The value of $(E_F - E_v)$ indicates how strongly p-type the material is.

It is usually inconvenient to draw $f(E)$ vs. E on every energy band diagram to indicate the electron and hole distributions. Therefore, it is common practice merely to indicate the position of E_F in band diagrams. This is sufficient information, since for a particular temperature the position of E_F implies the distributions in Fig. 3-15.

3.3.2 Electron and Hole Concentrations at Equilibrium

The Fermi distribution function can be used to calculate the concentrations of electrons and holes in a semiconductor, if the densities of available states in the valence and conduction bands are known. For example, the concentration of electrons in the conduction band is

$$n_0 = \int_{E_c}^{\infty} f(E)N(E)dE \tag{3-12}$$

where $N(E)dE$ is the density of states (cm^{-3}) in the energy range dE. The subscript 0 used with the electron and hole concentration symbols (n_0, p_0) indicates equilibrium conditions. The number of electrons per unit volume in the energy range dE is the product of the density of states and the probability of occupancy $f(E)$. Thus the total electron concentration is the integral

CIRRUS EX. 1062 - 40/190

over the entire conduction band, as in Eq. (3–12).[5] The function $N(E)$ can be calculated by using quantum mechanics and the Pauli exclusion principle (Appendix IV).

It is shown in Appendix IV that $N(E)$ is proportional to $E^{1/2}$, so the density of states in the conduction band increases with electron energy. On the other hand, the Fermi function becomes extremely small for large energies. The result is that the product $f(E)N(E)$ decreases rapidly above E_c, and very few electrons occupy energy states far above the conduction band edge. Similarly, the probability of finding an empty state (hole) in the valence band $[1 - f(E)]$ decreases rapidly below E_v, and most holes occupy states near the top of the valence band. This effect is demonstrated in Fig. 3–16, which shows the density of available states, the Fermi function, and the resulting number of electrons and holes occupying available energy states in the conduction and valence bands at thermal equilibrium (i.e., with no excitations except thermal energy). For holes, increasing energy points down in Fig. 3–16, since the E scale refers to electron energy.

The result of the integration of Eq. (3–12) is the same as that obtained if we represent all of the distributed electron states in the conduction band by an *effective density of states* N_c located at the conduction band edge E_c. Therefore, the conduction band electron concentration is simply the effective density of states at E_c times the probability of occupancy at E_c.[6]

$$n_0 = N_c f(E_c) \tag{3-13}$$

In this expression we assume the Fermi level E_F lies at least several kT below the conduction band. Then the exponential term is large compared with unity, and the Fermi function $f(E_c)$ can be simplified as

$$f(E_c) = \frac{1}{1 + e^{(E_c - E_F)/kT}} \simeq e^{-(E_c - E_F)/kT} \tag{3-14}$$

Since kT at room temperature is only 0.026 eV, this is generally a good approximation. For this condition the concentration of electrons in the conduction band is

$$\boxed{n_0 = N_c e^{-(E_c - E_F)/kT}} \tag{3-15}$$

The effective density of states N_c is shown in Appendix IV to be

$$N_c = 2\left(\frac{2\pi m_n^* kT}{h^2}\right)^{3/2} \tag{3-16a}$$

[5]The upper limit is actually improper in Eq. (3–12), since the conduction band does not extend to infinite energy. This is unimportant in the calculation of n_0, however, since $f(E)$ becomes negligibly small for large values of E. Most electrons occupy states near the bottom of the conduction band at equilibrium.

[6]The simple expression for n_0 obtained in Eq. (3–13) is the direct result of integrating Eq. (3–12), as in Appendix IV. Equations (3–15) and (3–19) properly include the effects of the conduction and valence bands through the density-of-states terms.

CIRRUS EX. 1062 - 41/190

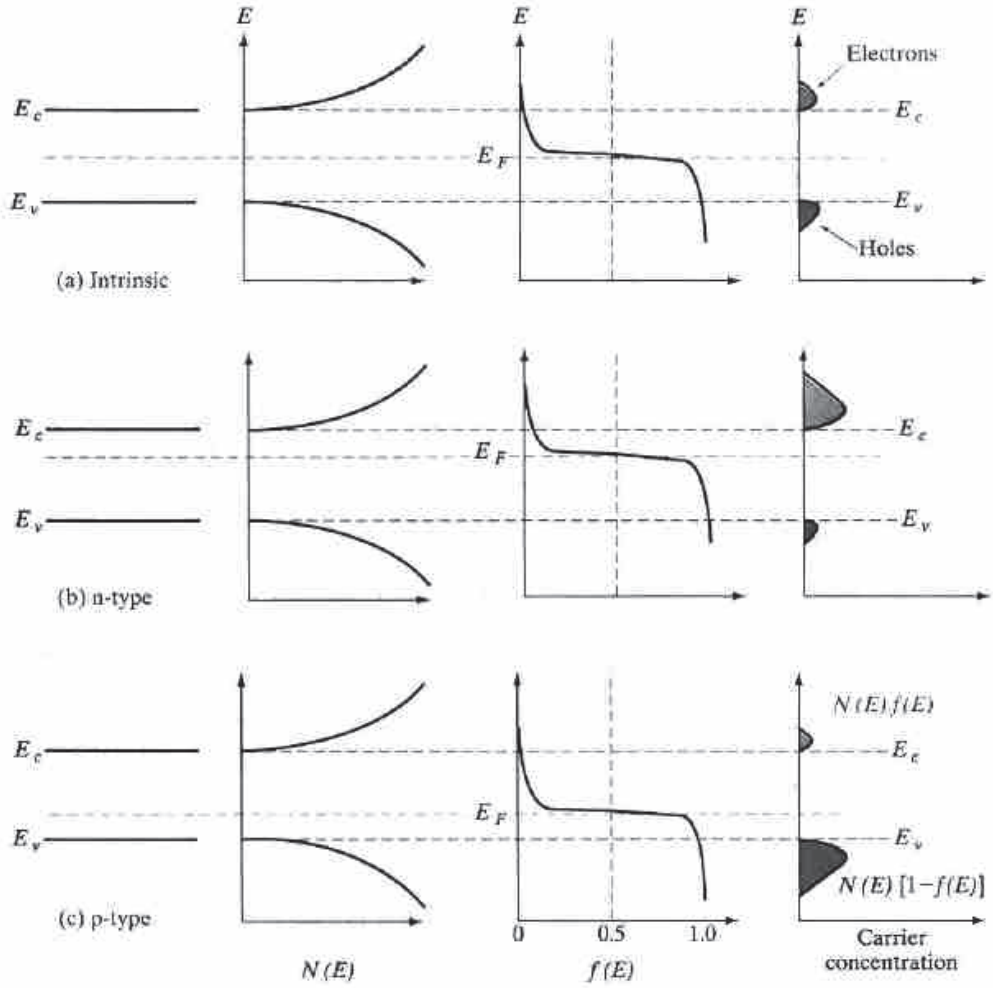


Figure 3–16
Schematic band diagram, density of states, Fermi–Dirac distribution, and the carrier concentrations for (a) intrinsic, (b) n-type, and (c) p-type semiconductors at thermal equilibrium.

Since the quantities in Eq. (3-16a) are known, values of N_c can be tabulated as a function of temperature. As Eq. (3-15) indicates, the electron concentration increases as E_F moves closer to the conduction band. This is the result we would predict from Fig. 3-15b.

In Eq. (3-16a), m_n^* is the density-of-states effective mass for electrons. To illustrate how it is obtained from the band curvature effective masses mentioned in Section 3.2.2, let us consider the 6 equivalent conduction

CIRRUS EX. 1062 - 42/190

band minima along the X-directions for Si. Looking at the cigar-shaped equi-energy surfaces in Fig. 3-10b, we find that we have more than one band curvature to deal with in calculating effective masses. There is a longitudinal effective mass m_l along the major axis of the ellipsoid, and the transverse effective mass m_t along the two minor axes. Since we have $(m_n^*)^{3/2}$ appearing in the density-of-states expression Eq. (3-16a), by using dimensional equivalence and adding contributions from all 6 valleys, we get

$$(m_n^*)^{3/2} = 6(m_l m_t^2)^{1/2} \tag{3-16b}$$

It can be seen that this is the geometric mean of the effective masses.

EXAMPLE 3–4 Calculate the density-of-states effective mass of electrons in Si.

SOLUTION For Si, $m_l = 0.98\ m_0$; $m_t = 0.19\ m_0$ from Appendix III.
There are six equivalent X valleys in the conduction band.

$$m_n^* = 6^{2/3}[0.98(0.19)^2]^{1/3} m_0 = 1.1\ m_0$$

Note: For GaAs, the conduction band equi-energy surfaces are spherical. So there is only one band curvature effective mass, and it is equal to the density-of-states effective mass ($= 0.067\ m_0$).

By similar arguments, the concentration of holes in the valence band is

$$p_0 = N_v[1 - f(E_v)] \tag{3-17}$$

where N_v is the effective density of states in the valence band. The probability of finding an empty state at E_v is

$$1 - f(E_v) = 1 - \frac{1}{1 + e^{(E_v - E_F)/kT}} \simeq e^{-(E_F - E_v)/kT} \tag{3-18}$$

for E_F larger than E_v by several kT. From these equations, the concentration of holes in the valence band is

$$\boxed{p_0 = N_v e^{-(E_F - E_v)/kT}} \tag{3-19}$$

The effective density of states in the valence band reduced to the band edge is

$$N_v = 2\left(\frac{2\pi m_p^* kT}{h^2}\right)^{3/2} \tag{3-20}$$

As expected from Fig. 3-15c, Eq. (3-19) predicts that the hole concentration increases as E_F moves closer to the valence band.

CIRRUS EX. 1062 - 43/190

The electron and hole concentrations predicted by Eqs. (3-15) and (3-19) are valid whether the material is intrinsic or doped, provided thermal equilibrium is maintained. Thus *for intrinsic material, E_F lies at some intrinsic level E_i* near the middle of the band gap (Fig. 3-15a), and the intrinsic electron and hole concentrations are

$$n_i = N_c e^{-(E_c - E_i)/kT}, \quad p_i = N_v e^{-(E_i - E_v)/kT} \tag{3-21}$$

The product of n_0 and p_0 at equilibrium is a constant for a particular material and temperature, even if the doping is varied:

$$n_0 p_0 = (N_c e^{-(E_c - E_F)/kT})(N_v e^{-(E_F - E_v)/kT}) = N_c N_v e^{-(E_c - E_v)/kT} \tag{3-22a}$$
$$= N_c N_v e^{-E_g/kT}$$
$$n_i p_i = (N_c e^{-(E_c - E_i)/kT})(N_v e^{-(E_i - E_v)/kT}) = N_c N_v e^{-E_g/kT} \tag{3-22b}$$

The intrinsic electron and hole concentrations are equal (since the carriers are created in pairs), $n_i = p_i$; thus the intrinsic concentration is

$$n_i = \sqrt{N_c N_v}\, e^{-E_g/2kT} \tag{3-23}$$

The constant product of electron and hole concentrations in Eq. (3-22) can be written conveniently as

$$\boxed{n_0 p_0 = n_i^2} \tag{3-24}$$

This is an important relation, and we shall use it extensively in later calculations. The intrinsic concentration for Si at room temperature is approximately $n_i = 1.5 \times 10^{10}\ cm^{-3}$.

Comparing Eqs. (3-21) and (3-23), we note that the intrinsic level E_i is the middle of the band gap ($E_c - E_i = E_g/2$), if the effective densities of states N_c and N_v are equal. There is usually some difference in effective mass for electrons and holes, however, and N_c and N_v are slightly different as Eqs. (3-16) and (3-20) indicate. The intrinsic level E_i is displaced from the middle of the band gap, more for GaAs than for Ge or Si.

Another convenient way of writing Eqs. (3-15) and (3-19) is

$$\boxed{n_0 = n_i e^{(E_F - E_i)/kT}} \tag{3-25a}$$
$$\boxed{p_0 = n_i e^{(E_i - E_F)/kT}} \tag{3-25b}$$

obtained by the application of Eq. (3-21). This form of the equations indicates directly that the electron concentrations is n_i when E_F is at the intrinsic level E_i, and that n_0 increases exponentially as the Fermi level moves away from E_i toward the conduction band. Similarly, the hole concentration p_0 varies from n_i to larger values as E_F moves from E_i toward the valence band. Since these equations reveal the qualitative features of carrier concentration so directly, they are particularly convenient to remember.

CIRRUS EX. 1062 - 44/190

EXAMPLE 3–5 A Si sample is doped with 10^{17} As atoms/cm^3. What is the equilibrium hole concentration p_0 at 300 K? Where is E_F relative to E_i?

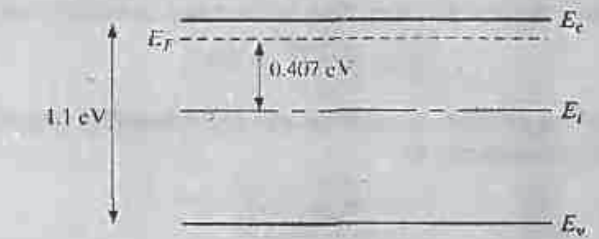
SOLUTION Since $N_d \gg n_i$, we can approximate $n_0 = N_d$ and

$$p_0 = \frac{n_i^2}{n_0} = \frac{2.25 \times 10^{20}}{10^{17}} = 2.25 \times 10^3 \text{ cm}^{-3}$$

From Eq. (3-25a), we have

$$E_F - E_i = kT \ln \frac{n_0}{n_i} = 0.0259 \ln \frac{10^{17}}{1.5 \times 10^{10}} = 0.407 \text{ eV}$$

The resulting band diagram is:



3.3.3 Temperature Dependence of Carrier Concentrations

The variation of carrier concentration with temperature is indicated by Eq. (3-25). Initially, the variation of n_0 and p_0 with T seems relatively straightforward in these relations. The problem is complicated, however, by the fact that n_i has a strong temperature dependence [Eq. (3-23)] and that E_F can also vary with temperature. Let us begin by examining the intrinsic carrier concentration. By combining Eqs. (3-23), (3-16a), and (3-20) we obtain

$$n_i(T) = 2\left(\frac{2\pi kT}{h^2}\right)^{3/2} (m_n^* m_p^*)^{3/4} e^{-E_g/2kT} \tag{3-26}$$

The exponential temperature dependence dominates $n_i(T)$, and a plot of ln n_i vs. $10^3/T$ appears linear (Fig. 3-17).[7] In this figure we neglect variations due to the $T^{3/2}$ dependence of the density-of-states function and the fact

[7]When plotting quantities such as carrier concentration, which involve a Boltzmann factor, it is common to use an inverse temperature scale. This allows terms which are exponential in $1/T$ to appear linear in the semilogarithmic plot. When reading such graphs, remember that temperature increases from right to left.

CIRRUS EX. 1062 - 45/190

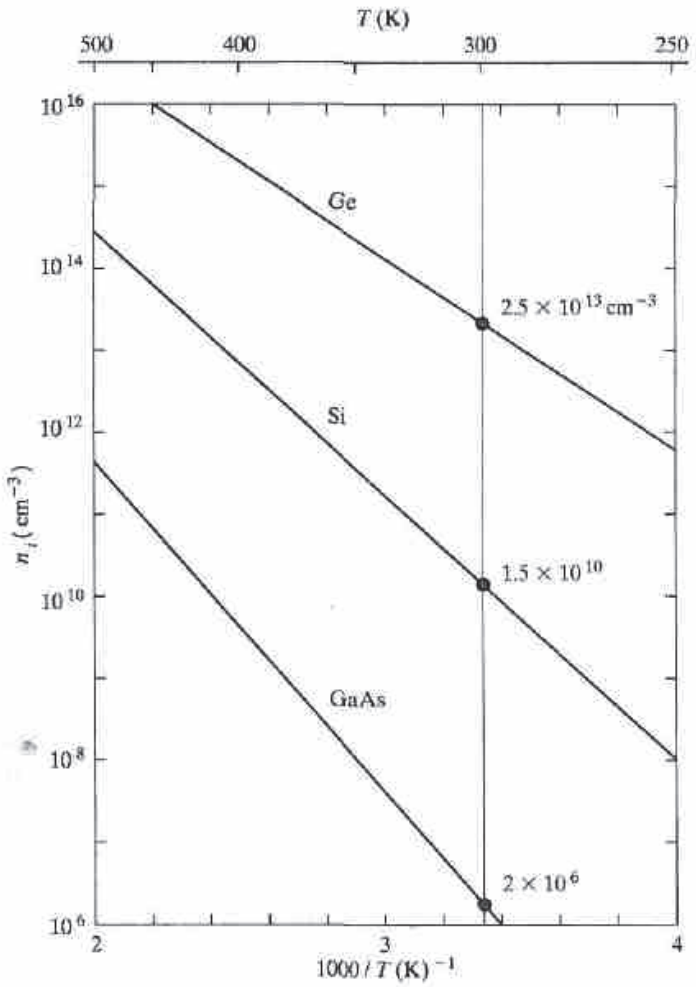


Figure 3–17
Intrinsic carrier concentration for Ge, Si, and GaAs as a function of inverse temperature. The room temperature values are marked for reference.

that E_g varies somewhat with temperature.[8] The value of n_i at any temperature is a definite number for a given semiconductor, and is known for most materials. Thus we can take n_i as given in calculating n_0 or p_0 from Eq. (3-25).[9]

With n_i and T given, the unknowns in Eq. (3-25) are the carrier concentrations and the Fermi level position relative to E_i. One of these two

[8]For Si the band gap E_g varies from about 1.11 eV at 300 K to about 1.16 eV at 0 K.

[9]Care must be taken to use consistent units in these calculations. For example, if an energy such as E_g is expressed in electron volts (eV), it should be multiplied by q (1.6×10^{-19} C) to convert to joules if k is in J/K; alternatively, E_g can be kept in eV and the value of k in eV/K can be used. At 300 K we can use $kT = 0.0259$ eV and E_g in eV.

CIRRUS EX. 1062 - 46/190

quantities must be given if the other is to be found. If the carrier concentration is held at a certain value, as in heavily doped extrinsic material, E_F can be obtained from Eq. (3-25). The temperature dependence of electron concentration in a doped semiconductor can be visualized as shown in Fig. 3-18. In this example, Si is doped n-type with a donor concentration N_d of 10^{15} cm^{-3}. At very low temperatures (large $1/T$), negligible intrinsic EHPs exist, and the donor electrons are bound to the donor atoms. As the temperature is raised, these electrons are donated to the conduction band, and at about 100 K ($1000/T = 10$) all the donor atoms are ionized. This temperature range is called the *ionization* region. Once the donors are ionized, the conduction band electron concentration is $n_0 \simeq N_d = 10^{15}$ cm^{-3}, since one electron is obtained for each donor atom. When every available extrinsic electron has been transferred to the conduction band, n_0 is virtually constant with temperature until the concentration of intrinsic carriers n_i becomes comparable to the extrinsic concentration N_d. Finally, at higher temperatures n_i is much greater than N_d, and the intrinsic carriers dominate. In most devices it is desirable to control the carrier concentration by doping rather than by thermal EHP generation. Thus one usually dopes the material such that the extrinsic range extends beyond the highest temperature at which the device is to be used.

3.3.4 Compensation and Space Charge Neutrality

When the concept of doping was introduced, we assumed the material contained either N_d donors or N_a acceptors, so that the extrinsic majority carrier concentrations were $n_0 \simeq N_d$ or $p_0 \simeq N_a$, respectively, for the n-type or

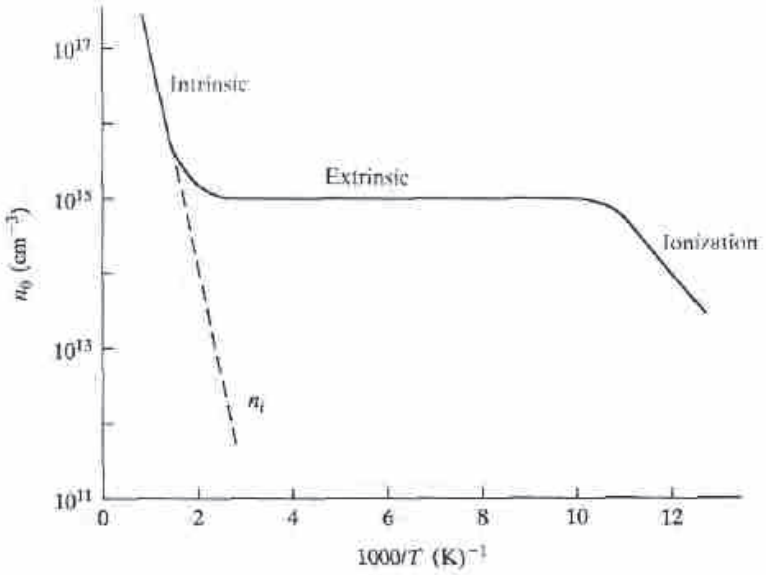


Figure 3-18 Carrier concentration vs. inverse temperature for Si doped with 10^{15} donors/cm^3.

CIRRUS EX. 1062 - 47/190

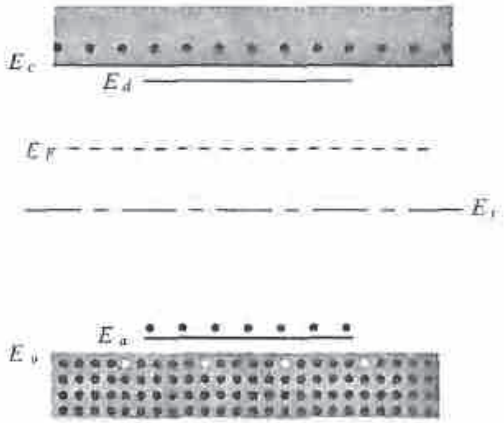


Figure 3–19
Compensation in an n-type semiconductor ($N_d > N_a$).

p-type material. It often happens, however, that a semiconductor contains both donors and acceptors. For example, Fig. 3–19 illustrates a semiconductor for which both donors and acceptors are present, but $N_d > N_a$. The predominance of donors makes the material n-type, and the Fermi level is therefore in the upper part of the band gap. Since E_F is well above the acceptor level E_a, this level is essentially filled with electrons. However, with E_F above E_i, we cannot expect a hole concentration in the valence band commensurate with the acceptor concentration. In fact, the filling of the E_a states occurs at the expense of the donated conduction band electrons. The mechanism can be visualized as follows: Assume an acceptor state is filled with a valence band electron as described in Fig. 3–12b, with a hole resulting in the valence band. This hole is then filled by recombination with one of the conduction band electrons. Extending this logic to all the acceptor atoms, we expect the resultant concentration of electrons in the conduction band to be $N_d - N_a$ instead of the total N_d. This process is called *compensation*. By this process it is possible to begin with an *n*-type semiconductor and add acceptors until $N_a = N_d$ and no donated electrons remain in the conduction band. In such compensated material, $n_0 = n_i = p_0$ and intrinsic conduction is obtained. With further acceptor doping the semiconductor becomes p-type with a hole concentration of essentially $N_a - N_d$.

The exact relationship among the electron, hole, donor, and acceptor concentrations can be obtained by considering the requirements for *space charge neutrality*. If the material is to remain electrostatically neutral, the sum of the positive charges (holes and ionized donor atoms) must balance the sum of the negative charges (electrons and ionized acceptor atoms):

$$p_0 + N_d^+ = n_0 + N_a^- \tag{3-27}$$

Thus in Fig. 3–19 the net electron concentration in the conduction band is

$$n_0 = p_0 + (N_d^+ - N_a^-) \tag{3-28}$$

CIRRUS EX. 1062 - 48/190

If the material is doped n-type ($n_0 \gg p_0$) and all the impurities are ionized, we can approximate Eq. (3-28) by $n_0 \simeq N_d - N_a$.

Since the intrinsic semiconductor itself is electrostatically neutral and the doping atoms we add are also neutral, the requirement of Eq. (3-27) must be maintained at equilibrium. The electron and hole concentrations and the Fermi level adjust such that Eqs. (3-27) and (3-25) are satisfied.

3.4 DRIFT OF CARRIERS IN ELECTRIC AND MAGNETIC FIELDS

Knowledge of carrier concentrations in a solid is necessary for calculating current flow in the presence of electric or magnetic fields. In addition to the values of n and p, we must be able to take into account the collisions of the charge carriers with the lattice and with the impurities. These processes will affect the ease with which electrons and holes can flow through the crystal, that is, their *mobility* within the solid. As should be expected, these collision and scattering processes depend on temperature, which affects the thermal motion of the lattice atoms and the velocity of the carriers.

3.4.1 Conductivity and Mobility

The charge carriers in a solid are in constant motion, even at thermal equilibrium. At room temperature, for example, the thermal motion of an individual electron may be visualized as random scattering from lattice vibrations, impurities, other electrons, and defects (Fig. 3-20a). Since the scattering is random, there is no net motion of the group of n electrons/cm^3 over any period of time. This is not true of an individual electron, of course. The probability of

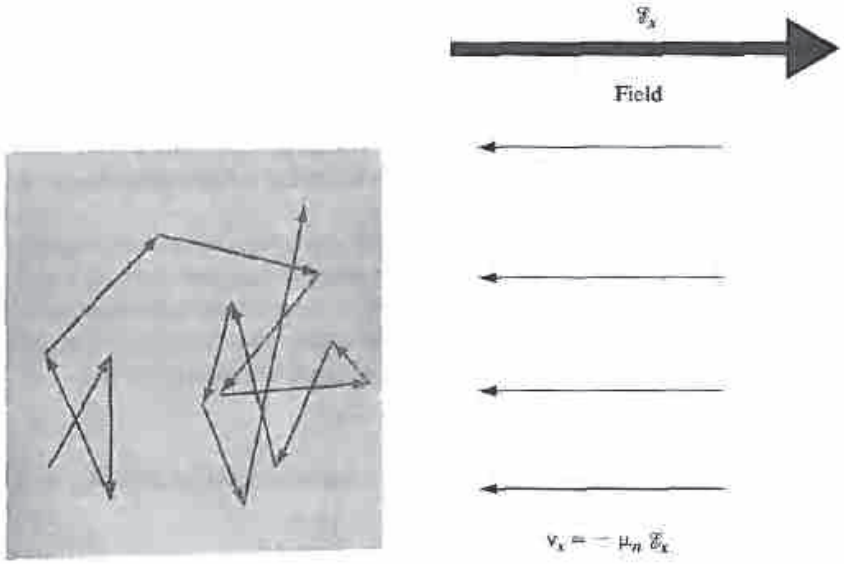


Figure 3-20 (a) Random thermal motion of an electron in a solid; (b) well-directed drift velocity with an applied electric field.

CIRRUS EX. 1062 - 49/190

the electron in Fig. 3–20a returning to its starting point after some time t is negligibly small. However, if a large number of electrons is considered (e.g., 10^{16} cm^{-3} in an n-type semiconductor), there will be no preferred direction of motion for the group of electrons and no net current flow.

If an electric field $\mathscr{E}_x$ is applied in the x-direction, each electron experiences a net force $-q\mathscr{E}_x$ from the field. This force may be insufficient to alter appreciably the random path of an individual electron; the effect when averaged over all the electrons, however, is a net motion of the group in the $-x$-direction (Fig. 3–20b). If p_x is the x-component of the total momentum of the group, the force of the field on the n electrons/cm^3 is

$$-nq\mathscr{E}_x = \left.\frac{dp_x}{dt}\right|_{\text{field}} \tag{3-29}$$

Initially, Eq. (3–29) seems to indicate a continuous acceleration of the electrons in the $-x$-direction. This is not the case, however, because the net acceleration of Eq. (3–29) is just balanced in steady state by the decelerations of the collision processes. Thus while the steady field $\mathscr{E}_x$ does produce a net momentum p_{-x}, the net rate of change of momentum when collisions are included must be zero in the case of steady state current flow.

To find the total rate of momentum change from collisions, we must investigate the collision probabilities more closely. If the collisions are truly random, there will be a constant probability of collision at any time for each electron. Let us consider a group of N_0 electrons at time $t = 0$ and define $N(t)$ as the number of electrons that *have not* undergone a collision by time t. The rate of decrease in $N(t)$ at any time t is proportional to the number left unscattered at t,

$$-\frac{dN(t)}{dt} = \frac{1}{\bar{t}}N(t) \tag{3-30}$$

where $\bar{t}^{-1}$ is a constant of proportionality.

The solution to Eq. (3–30) is an exponential function

$$N(t) = N_0 e^{-t/\bar{t}} \tag{3-31}$$

and $\bar{t}$ represents the mean time between scattering events,[10] called the *mean free time*. The probability that any electron has a collision in the time interval dt is $dt/\bar{t}$. Thus the differential change in p_x due to collisions in time dt is

$$dp_x = -p_x \frac{dt}{\bar{t}} \tag{3-32}$$

[10]Equations (3–30) and (3–31) are typical of events dominated by random processes, and the forms of these equations occur often in many branches of physics and engineering. For example, in the radioactive decay of unstable nuclear isotopes, N_0 nuclides decay exponentially with a mean lifetime $\bar{t}$. Other examples will be found in this text, including the absorption of light in a semiconductor and the recombination of excess EHPs.

CIRRUS EX. 1062 - 50/190

The rate of change of p_x due to the decelerating effect of collisions is

$$\left.\frac{dp_x}{dt}\right|_{\text{collisions}} = -\frac{p_x}{\bar{t}} \tag{3-33}$$

The sum of acceleration and deceleration effects must be zero for steady state. Taking the sum of Eqs. (3-29) and (3-33), we have

$$-\frac{p_x}{\bar{t}} - nq\mathscr{E}_x = 0 \tag{3-34}$$

The average momentum per electron is

$$\langle p_x \rangle = \frac{p_x}{n} = -q\bar{t}\mathscr{E}_x \tag{3-35}$$

where the angular brackets indicate an average over the entire group of electrons. As expected for steady state, Eq. (3-35) indicates that the electrons have *on the average* a constant net velocity in the negative x-direction:

$$\langle v_x \rangle = \frac{\langle p_x \rangle}{m_n^*} = -\frac{q\bar{t}}{m_n^*}\mathscr{E}_x \tag{3-36}$$

Actually, the individual electrons move in many directions by thermal motion during a given time period, but Eq. (3-36) tells us the *net drift* of an average electron in response to the electric field. The drift speed described by Eq. (3-36) is usually much smaller than the random speed due to the thermal motion v_{th}.

The current density resulting from this net drift is just the number of electrons crossing a unit area per unit time ($n\langle v_x \rangle$) multiplied by the charge on the electron ($-q$):

$$\boxed{\begin{gathered} J_x = -qn\langle v_x \rangle \\ \frac{\text{ampere}}{\text{cm}^2} = \frac{\text{coulomb}}{\text{electron}} \cdot \frac{\text{electrons}}{\text{cm}^3} \cdot \frac{\text{cm}}{\text{s}} \end{gathered}} \tag{3-37}$$

Using Eq. (3-36) for the average velocity, we obtain

$$J_x = \frac{nq^2\bar{t}}{m_n^*}\mathscr{E}_x \tag{3-38}$$

Thus the current density is proportional to the electric field, as we expect from Ohm's law:

$$J_x = \sigma\mathscr{E}_x, \quad \text{where } \sigma \equiv \frac{nq^2\bar{t}}{m_n^*} \tag{3-39}$$

The conductivity $\sigma(\Omega\text{-cm})^{-1}$ can be written

$$\sigma = qn\mu_n, \quad \text{where } \mu_n \equiv \frac{q\bar{t}}{m_n^*} \tag{3-40a}$$

CIRRUS EX. 1062 - 51/190

The quantity μ_n, called the *electron mobility*, describes the ease with which electrons drift in the material. Mobility is a very important quantity in characterizing semiconductor materials and in device development.

Here m_n^* is the conductivity effective mass for electrons, different from the density-of-states effective mass mentioned in Eq. (3-16b). While we use the density-of-states effective mass to count the number of carriers in bands, we must use the conductivity effective mass for charge transport problems. To illustrate how it is obtained from the band curvature effective masses mentioned in Section 3.2.2, once again let us consider the 6 equivalent conduction band minima along the X-directions for Si, with the band curvature longitudinal effective mass, m_l, along the major axis of the ellipsoid, and the transverse effective mass, m_t, along the two minor axes (Fig. 3-10b). Since we have $1/m_n^*$ in the mobility expression Eq. (3-40a), by using dimensional equivalence, we can write the conductivity effective mass as the harmonic mean of the band curvature effective masses.

$$\frac{1}{m_n^*} = \frac{1}{3}\left(\frac{1}{m_l} + \frac{2}{m_t}\right) \tag{3-40b}$$

EXAMPLE 3-6

Calculate the conductivity effective mass of electrons in Si.

SOLUTION

For Si, $m_l = 0.98\, m_0$; $m_t = 0.19\, m_0$ (Appendix III).
There are 6 equivalent X valleys in the conduction band.

$$1/m_n^* = 1/3(1/m_x + 1/m_y + 1/m_z) = 1/3(1/m_l + 2/m_t)$$

$$1/m_n^* = \frac{1}{3}\left(\frac{1}{0.98\, m_0} + \frac{2}{0.19\, m_0}\right)$$

$$m_n^* = 0.26\, m_0$$

Note: For GaAs, the conduction band equi-energy surfaces are spherical. So there is only one band curvature effective mass. (The density of states effective mass and the conductivity effective mass are both $0.067\, m_0$.)

The mobility defined in Eq. (3-40a) can be expressed as the average particle drift velocity per unit electric field. Comparing Eqs. (3-36) and (3-40a), we have

$$\mu_n = -\frac{\langle v_x \rangle}{\mathscr{E}_x} \tag{3-41}$$

The units of mobility are (cm/s)/(V/cm) = cm^2/V-s, as Eq. (3-41) suggests. The minus sign in the definition results in a positive value of mobility, since electrons drift opposite to the field.

CIRRUS EX. 1062 - 52/190

The current density can be written in terms of mobility as

$$J_x = qn\mu_n \mathscr{E}_x \tag{3-42}$$

This derivation has been based on the assumption that the current is carried primarily by electrons. For hole conduction we change n to p, $-q$ to $+q$, and μ_n to μ_p, where $\mu_p = +\langle v_x \rangle / \mathscr{E}_x$ is the mobility for holes. If both electrons and holes participate, we must modify Eq. (3-42) to

$$\boxed{J_x = q(n\mu_n + p\mu_p)\mathscr{E}_x = \sigma \mathscr{E}_x} \tag{3-43}$$

Values of μ_n and μ_p are given for many of the common semiconductor materials in Appendix III. According to Eq. (3-40), the parameters determining mobility are m^* and mean free time t. Effective mass is a property of the material's band structure, as described by Eq. (3-3). Thus we expect m_n^* to be small in the strongly curved Γ minimum of the GaAs conduction band (Fig. 3-6), with the result that μ_n is very high. In a more gradually curved band, a larger m^* in the denominator of Eq. (3-40) leads to a smaller value of mobility. It is reasonable to expect that lighter particles are more mobile than heavier particles (which is satisfying, since the common-sense value of effective mass is not always apparent). The other parameter determining mobility is the mean time between scattering events, t. In Section 3.4.3 we shall see that this is determined primarily by temperature and impurity concentration in the semiconductor.

3.4.2 Drift and Resistance

Let us look more closely at the drift of electrons and holes. If the semiconductor bar of Fig. 3-21 contains both types of carrier, Eq. (3-43) gives the conductivity of the material. The resistance of the bar is then

$$R = \frac{\rho L}{wt} = \frac{L}{wt}\frac{1}{\sigma} \tag{3-44}$$

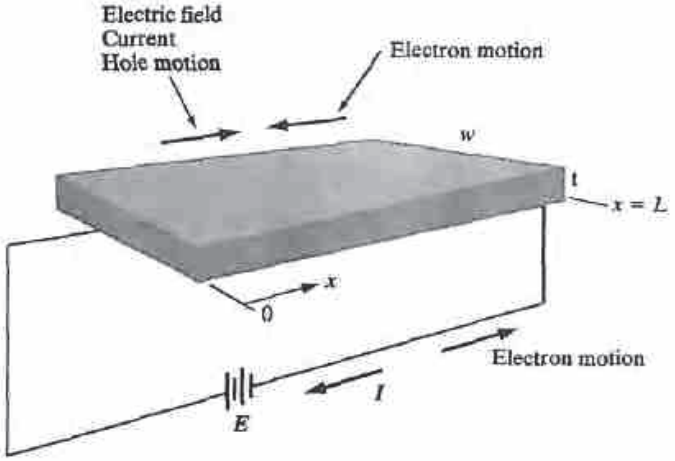


Figure 3-21 Drift of electrons and holes in a semiconductor bar.

CIRRUS EX. 1062 - 53/190

where ρ is the resistivity (Ω-cm). The physical mechanism of carrier drift requires that the holes in the bar move as a group in the direction of the electric field and that the electrons move as a group in the opposite direction. Both the electron and the hole components of current are in the direction of the $\mathscr{E}$ field, since conventional current is positive in the direction of hole flow and opposite to the direction of electron flow. The drift current described by Eq. (3-43) is constant throughout the bar. A valid question arises, therefore, concerning the nature of the electron and hole flow at the contacts and in the external circuit. We should specify that the contacts to the bar of Fig. 3-21 are *ohmic*, meaning that they are perfect sources and sinks of both carrier types and have no special tendency to inject or collect either electrons or holes.

If we consider that current is carried around the external circuit by electrons, there is no problem in visualizing electrons flowing into the bar at one end and out at the other (always opposite to I). Thus for every electron leaving the left end ($x = 0$) of the bar in Fig. 3-21, there is a corresponding electron entering at $x = L$, so that the electron concentration in the bar remains constant at n. But what happens to the holes at the contacts? As a hole reaches the ohmic contact at $x = L$, it recombines with an electron, which must be supplied through the external circuit. As this hole disappears, a corresponding hole must appear at $x = 0$ to maintain space charge neutrality. It is reasonable to consider the source of this hole as the generation of an EHP at $x = 0$, with the hole flowing into the bar and the electron flowing into the external circuit.

3.4.3 Effects of Temperature and Doping on Mobility

The two basic types of scattering mechanisms that influence electron and hole mobility are *lattice scattering* and *impurity scattering*. In lattice scattering a carrier moving through the crystal is scattered by a vibration of the lattice, resulting from the temperature.[11] The frequency of such scattering events increases as the temperature increases, since the thermal agitation of the lattice becomes greater. Therefore, we should expect the mobility to decrease as the sample is heated (Fig. 3-22). On the other hand, scattering from crystal defects such as ionized impurities becomes the dominant mechanism at low temperatures. Since the atoms of the cooler lattice are less agitated, lattice scattering is less important; however, the thermal motion of the carriers is also slower. Since a slowly moving carrier is likely to be scattered more strongly by an interaction with a charged ion than is a carrier with greater momentum, impurity scattering events cause a decrease in mobility with decreasing temperature. As Fig. 3-22 indicates, the approximate temperature dependencies are $T^{-3/2}$ for lattice scattering and $T^{3/2}$ for impurity scattering. Since the scattering probability of Eq. (3-32) is inversely proportional to the

[11] Collective vibrations of atoms in the crystal are called *phonons*. Thus lattice scattering is also known as *phonon scattering*.

CIRRUS EX. 1062 - 54/190

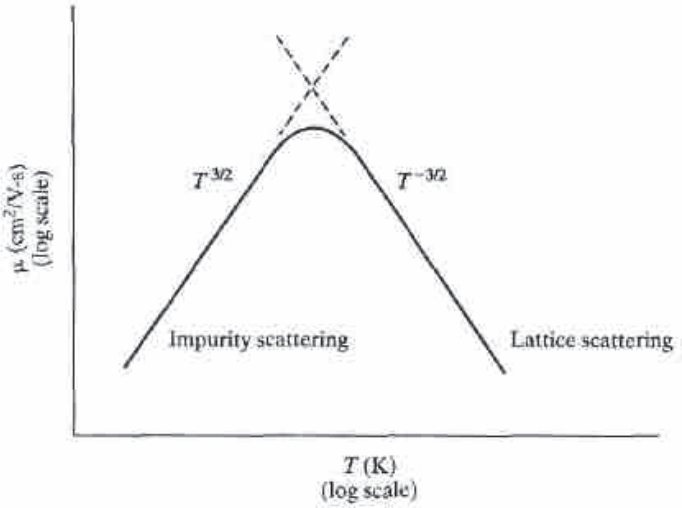


Figure 3–22
Approximate temperature dependence of mobility with both lattice and impurity scattering.

mean free time and therefore to mobility, the mobilities due to two or more scattering mechanisms add inversely:

$$\frac{1}{\mu} = \frac{1}{\mu_1} + \frac{1}{\mu_2} + \cdots \tag{3-45}$$

As a result, the mechanism causing the lowest mobility value dominates, as shown in Fig. 3–22.

As the concentration of impurities increases, the effects of impurity scattering are felt at higher temperatures. For example, the electron mobility μ_n of intrinsic silicon at 300 K is 1350 cm^2/(V-s). With a donor doping concentration of 10^{17} cm^{-3}, however, μ_n is 700 cm^2/(V-s). Thus, the presence of the 10^{17} ionized donors/cm^3 introduces a significant amount of impurity scattering. This effect is illustrated in Fig. 3-23, which shows the variation of mobility with doping concentration at room temperature.

EXAMPLE 3–7 A Si bar 0.1 cm long and 100 μm^2 in cross-sectional area is doped with 10^{17} cm^{-3} phosphorus. Find the current at 300 K with 10 V applied.

SOLUTION For this applied voltage, we see that the electric field is low enough that we are in the ohmic regime. From Fig. 3-23, for this doping, the $\mu_n = 700$ cm^2/V-s.

$$\sigma = q\mu_n n_0 = 1.6 \times 10^{-19} \times 700 \times 10^{17} = 11.2(\Omega \cdot \text{cm})^{-1} = \rho^{-1}$$

$$\rho = 0.0893\ \Omega \cdot \text{cm}$$

$$R = \rho L/A = 0.0893 \times 0.1/10^{-6} = 8.93 \times 10^3 \Omega$$

$$I = V/R = 10/(8.93 \times 10^3) = \mathbf{1.12\ mA}$$

CIRRUS EX. 1062 - 55/190

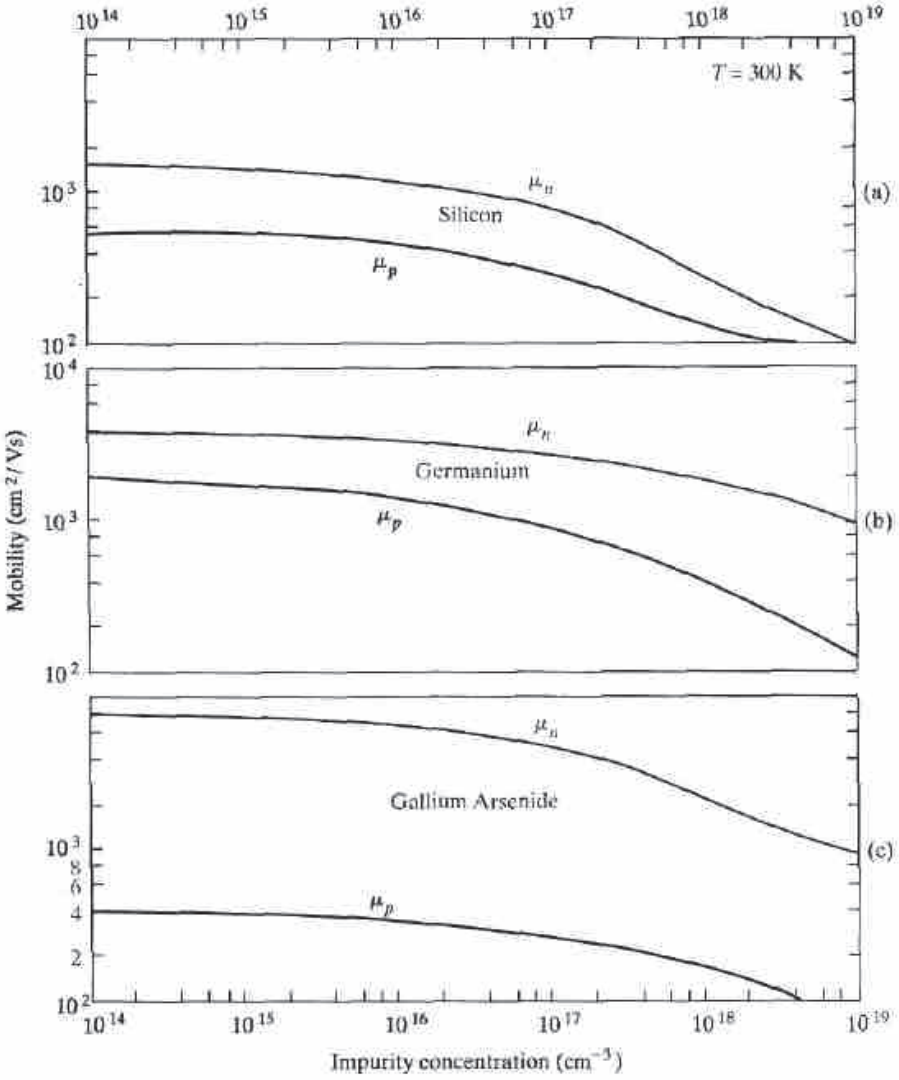


Figure 3–23
Variation of mobility with total doping impurity concentration ($N_a + N_d$) for Ge, Si, and GaAs at 300 K.

3.4.4 High-Field Effects

One assumption implied in the derivation of Eq. (3-39) was that Ohm's law is valid in the carrier drift processes. That is, it was assumed that the drift current is proportional to the electric field and that the proportionality constant (σ) is not a function of field $\mathscr{E}$. This assumption is valid over a wide range of $\mathscr{E}$. However, large electric fields ($> 10^3$V/cm) can cause the drift velocity and therefore the current $J = -qnv_d$ to exhibit a sublinear dependence on the

CIRRUS EX. 1062 - 56/190

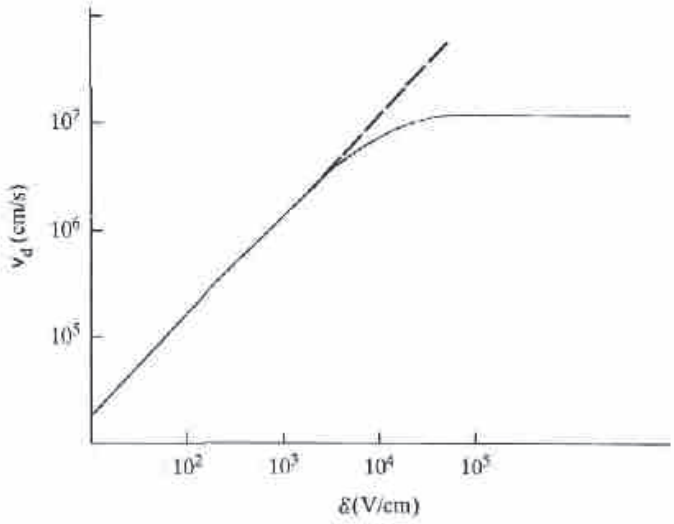


Figure 3–24
Saturation of electron drift velocity at high electric fields for Si.

electric field. This dependence of σ upon $\mathscr{E}$ is an example of a *hot carrier* effect, which implies that the carrier drift velocity v_d is comparable to the thermal velocity v_{th}.

In many cases an upper limit is reached for the carrier drift velocity in a high field (Fig. 3–24). This limit occurs near the mean thermal velocity ($= 10^7$ cm/s) and represents the point at which added energy imparted by the field is transferred to the lattice rather than increasing the carrier velocity. The result of this *scattering limited velocity* is a fairly constant current at high field. This behavior is typical of Si, Ge, and some other semiconductors. However, there are other important effects in some materials; for example, in Chapter 10 we shall discuss a *decrease* in electron velocity at high fields for GaAs and certain other materials, which results in negative conductivity and current instabilities in the sample. Another important high-field effect is avalanche multiplication, which we shall discuss in Section 5.4.2.

3.4.5 The Hall Effect

If a magnetic field is applied perpendicular to the direction in which holes drift in a p-type bar, the path of the holes tends to be deflected (Fig. 3–25). Using vector notation, the total force on a single hole due to the electric and magnetic fields is

$$\mathbf{F} = q(\mathscr{E} + \mathbf{v} \times \mathscr{B}) \tag{3-46}$$

In the y-direction the force is

$$F_y = q(\mathscr{E}_y - v_x \mathscr{B}_z) \tag{3-47}$$

The important result of Eq. (3–47) is that unless an electric field $\mathscr{E}_y$ is established along the width of the bar, each hole will experience a net force

CIRRUS EX. 1062 - 57/190

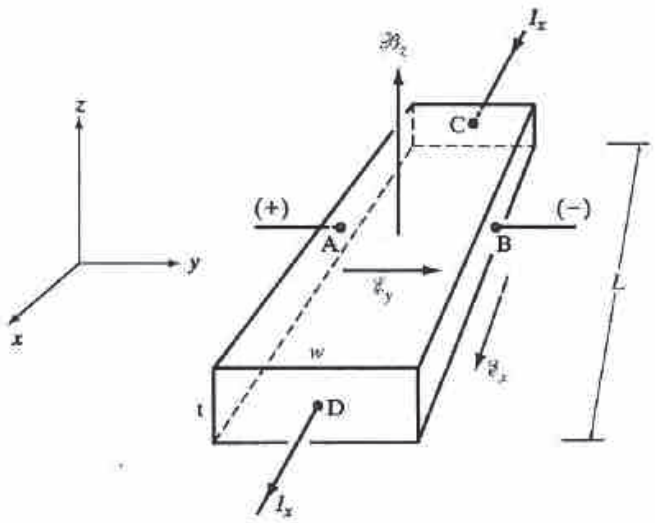


Figure 3–25
The Hall effect.

(and therefore an acceleration) in the $-y$-direction due to the $qv_x\mathcal{B}_z$ product. Therefore, to maintain a steady state flow of holes down the length of the bar, the electric field $\mathscr{E}_y$ must just balance the product $v_x\mathcal{B}_z$:

$$\mathscr{E}_y = v_x\mathcal{B}_z \tag{3-48}$$

so that the net force F_y is zero. Physically, this electric field is set up when the magnetic field shifts the hole distribution slightly in the $-y$-direction. Once the electric field $\mathscr{E}_y$ becomes as large as $v_x\mathcal{B}_z$, no net lateral force is experienced by the holes as they drift along the bar. The establishment of the electric field $\mathscr{E}_y$ is known as the *Hall effect*, and the resulting voltage $V_{AB} = \mathscr{E}_y w$ is called the *Hall voltage*. If we use the expression derived in Eq. (3–37) for the drift velocity (using $+q$ and p_0 for holes), the field $\mathscr{E}_y$ becomes

$$\mathscr{E}_y = \frac{J_x}{qp_0}\mathcal{B}_z = R_H J_x \mathcal{B}_z, \quad R_H \equiv \frac{1}{qp_0} \tag{3-49}$$

Thus the Hall field is proportional to the product of the current density and the magnetic flux density. The proportionality constant $R_H = (qp_0)^{-1}$ is called the *Hall coefficient*. A measurement of the Hall voltage for a known current and magnetic field yields a value for the hole concentration p_0

$$p_0 = \frac{1}{qR_H} = \frac{J_x\mathcal{B}_z}{q\mathscr{E}_y} = \frac{(I_x/wt)\mathcal{B}_z}{q(V_{AB}/w)} = \frac{I_x\mathcal{B}_z}{qtV_{AB}} \tag{3-50}$$

Since all of the quantities in the right-hand side of Eq. (3–50) can be measured, the Hall effect can be used to give quite accurate values for carrier concentration.

CIRRUS EX. 1062 - 58/190

If a measurement of resistance R is made, the sample resistivity ρ can be calculated:

$$\rho(\Omega\text{-cm}) = \frac{Rwt}{L} = \frac{V_{CD}/I_x}{L/wt} \tag{3-51}$$

Since the conductivity $\sigma = 1/\rho$ is given by $q\mu_p p_0$, the mobility is simply the ratio of the Hall coefficient and the resistivity:

$$\mu_p = \frac{\sigma}{qp_0} = \frac{1/\rho}{q(1/qR_H)} = \frac{R_H}{\rho} \tag{3-52}$$

Measurements of the Hall coefficient and the resistivity over a range of temperatures yield plots of majority carrier concentration and mobility vs. temperature. Such measurements are extremely useful in the analysis of semiconductor materials. Although the discussion here has been related to p-type material, similar results are obtained for n-type material. A negative value of q is used for electrons, and the Hall voltage V_{AB} and Hall coefficient R_H are negative. In fact, measurement of the sign of the Hall voltage is a common technique for determining if an unknown sample is p-type or n-type.

EXAMPLE 3-8 Referring to Fig. 3-25, consider a semiconductor bar with $w = 0.1$ mm, $t = 10$ μm, and $L = 5$ mm. For $\mathcal{B} = 10$ kg in the direction shown (1 kG = 10^{-5} Wb/cm^2) and a current of 1 mA, we have $V_{AB} = -2$ mV, $V_{CD} = 100$ mV. Find the type, concentration, and mobility of the majority carrier.

SOLUTION $\mathcal{B}_z = 10^{-4}$ Wb/cm^2

From the sign of V_{AB}, we can see that the majority carriers are electrons:

$$n_0 = \frac{I_x \mathcal{B}_z}{qt(-V_{AB})} = \frac{(10^{-3})(10^{-4})}{1.6 \times 10^{-19}(10^{-3})(2 \times 10^{-3})} = 3.125 \times 10^{17}\ \text{cm}^{-3}$$

$$\rho = \frac{R}{L/wt} = \frac{V_{CD}/I_x}{L/wt} = \frac{0.1/10^{-3}}{0.5/0.01 \times 10^{-3}} = 0.002\ \Omega \cdot \text{cm}$$

$$\mu_n = \frac{1}{\rho q n_0} = \frac{1}{(0.002)(1.6 \times 10^{-19})(3.125 \times 10^{17})} = \mathbf{10{,}000\ cm^2(V \cdot s)^{-1}}$$

3.5 INVARIANCE OF THE FERMI LEVEL AT EQUILIBRIUM

In this chapter we have discussed homogeneous semiconductors, without variations in doping and without junctions between dissimilar materials. In the following chapters we will be considering cases in which nonuniform doping occurs in a given semiconductor, or junctions occur between different semiconductors or a semiconductor and a metal. These cases are crucial

CIRRUS EX. 1062 - 59/190

to the various types of electronic and optoelectronic devices made in semiconductors. In anticipation of those discussions, an important concept should be established here regarding the demands of equilibrium. That concept can be summarized by noting that *no discontinuity or gradient can arise in the equilibrium Fermi level* E_F.

To demonstrate this assertion, let us consider two materials in intimate contact such that electrons can move between the two (Fig. 3–26). These may be, for example, dissimilar semiconductors, n- and p-type regions, a metal and a semiconductor, or simply two adjacent regions of a nonuniformly doped semiconductor. Each material is described by a Fermi-Dirac distribution function and some distribution of available energy states that electrons can occupy.

There is no current, and therefore no net charge transport, at thermal equilibrium. There is also no net transfer of energy. Therefore, for each energy E in Fig. 3–26 any transfer of electrons from material 1 to material 2 must be exactly balanced by the opposite transfer of electrons from 2 to 1. We will let the density of states at energy E in material 1 be called $N_1(E)$ and in material 2 we will call it $N_2(E)$. At energy E the rate of transfer of electrons from 1 to 2 is proportional to the number of filled states at E in material 1 times the number of empty states at E in material 2:

$$\text{rate from 1 to 2} \propto N_1(E)f_1(E) \cdot N_2(E)[1 - f_2(E)] \tag{3-53}$$

where $f(E)$ is the probability of a state being filled at E in each material, i.e., the Fermi-Dirac distribution function given by Eq. (3-10). Similarly,

$$\text{rate from 2 to 1} \propto N_2(E)f_2(E) \cdot N_1(E)[1 - f_1(E)] \tag{3-54}$$

At equilibrium these must be equal:

$$N_1(E)f_1(E) \cdot N_2(E)[1 - f_2(E)] = N_2(E)f_2(E) \cdot N_1(E)[1 - f_1(E)] \tag{3-55}$$

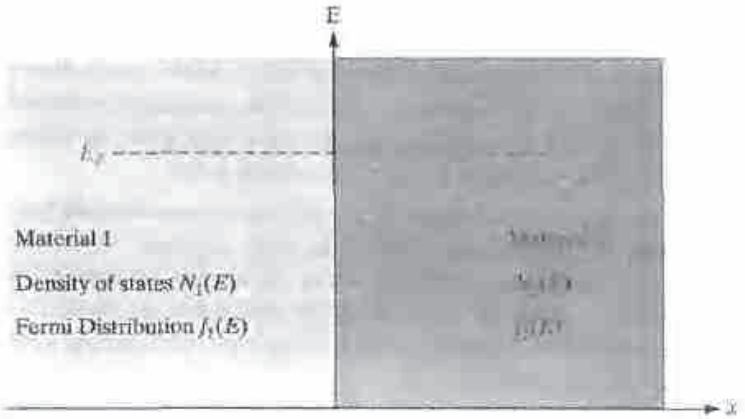


Figure 3–26
Two materials in intimate contact at equilibrium. Since the net motion of electrons is zero, the equilibrium Fermi level must be constant throughout.

CIRRUS EX. 1062 - 60/190

Rearranging terms, we have, at energy E,

$$N_1 f_1 N_2 - N_1 f_1 N_2 f_2 = N_2 f_2 N_1 - N_2 f_2 N_1 f_1 \tag{3-56}$$

which results in

$$f_1(E) = f_2(E), \quad \text{that is,} \quad [1 + e^{(E - E_{F1})/kT}]^{-1} = [1 + e^{(E - E_{F2})/kT}]^{-1} \tag{3-57}$$

Therefore, we conclude that $E_{F1} = E_{F2}$. That is, there is no discontinuity in the equilibrium Fermi level. More generally, we can state that the Fermi level at equilibrium must be constant throughout materials in intimate contact. One way of stating this is that no gradient exists in the Fermi level at equilibrium:

$$\boxed{\frac{dE_F}{dx} = 0} \tag{3-58}$$

We will make considerable use of this result in the chapters to follow.

SUMMARY

3.1 In a diamond lattice, each Si atom (with four valence electrons) is surrounded by four Si atoms that form four *covalent* bonds consisting of shared electron pairs, thereby forming an *octet* of electrons in the valence shell. In zinc blende structures such as GaAs, electrons are partly shared (covalent bonding) and partly transferred from Ga to As (*ionic* bonding).

3.2 In crystals, electronic *wave functions overlap* to give various linear combinations of atomic orbitals (LCAO). Bonding or symmetric combinations of the wave functions of valence-shell electrons form (*almost*) *continuous allowed bands* of energies in the (*almost*) *filled valence* band, separated by an energy gap from higher energy states in an (*almost*) *empty conduction* band that correspond to the antibonding or antisymmetric LCAOs. Empty electronic states in the valence band can be considered to be positively charged carriers (*holes*), while filled states in the conduction band are negatively charged (*conduction*) *electrons*.

3.3 If the band gaps are large, we get insulators; if they are small (~1 eV), we get semiconductors; and if they are zero, we get conductors (metals).

3.4 *Simplified* band diagrams plot *electron energy* in the conduction band (*increasing upwards*) as a function of position. The *band edge* corresponds to *potential* energy, and the distance from the band edge gives the *kinetic* energy. *Hole energies increase downward* in the valence band.

3.5 Carrier energies can also be plotted as a function of wave vector $\mathbf{k}$ (proportional to velocity or momentum) to give $(E, \mathbf{k})$ *band structures*, which can be direct (conduction band minimum directly above valence band maximum) or indirect. The *curvature* of the $(E, \mathbf{k})$ is inversely proportional to the *effective mass* m^* of the carriers. The m^* accounts for the interactions of the carriers with the periodic crystal potential.

CIRRUS EX. 1062 - 61/190

3.6 In a pure semiconductor, we have an *intrinsic* concentration of electrons (or holes), n_i, that result from *thermal* generation–recombination between the valence and conduction band (or bond breaking). If we replace some Si atoms (with four valence electrons) with *donor* impurities with five valence electrons, they donate conduction electrons, n ($= N_d^+$); similarly, *acceptors* create holes p.

3.7 The number of electrons, n, is the integral with respect to energy from the bottom to the top of the conduction band of the product of the available *density of states* (DOS) and the *Fermi–Dirac* (FD) distribution function. For parabolic band structures, we get a parabolic DOS. The FD function is the average occupancy of an electronic state. The electron concentration n can also be expressed as the product of an *effective* DOS at the band edge and the FD occupancy at E_c, and similarly for holes p. The np product in equilibrium is constant (n_i^2).

3.8 Electrons in a solid execute *random Brownian* motion with an average kinetic energy related to the thermal energy kT. In an electric field, electrons *drift* (on top of the random motion) with a velocity equal to *mobility* times field in the *ohmic* regime and *saturation velocity* at *high fields*. The drift current is proportional to carrier concentration times drift velocity. Negatively charged electrons drift opposite to the electric field, and the current is opposite to the motion. Positively charged holes drift in the direction of the electric field, and the current goes in the same direction as the hole flow.

3.9 Carrier mobility is determined by *scattering*, caused by *deviations* from a periodic lattice potential, such as lattice vibrations (phonons) or ionized impurities. Carrier mobility and concentration can be obtained by the Hall effect and resistivity measurements.

PROBLEMS

3.1 Calculate the approximate donor binding energy for GaAs ($\epsilon_r = 13.2$, $m_n^* = 0.067\, m_0$).

3.2 Calculate values for the Fermi function $f(C)$ at 300 K and plot vs. energy in eV as in Fig. 3–14. Choose $E_F = 1$ eV and make the calculated points closer together near the Fermi level to obtain a smooth curve. Notice that $f(E)$ varies quite rapidly within a few kT of E_F. Show that the probability that a state ΔE above E_F is occupied is the same as the probability that the state ΔE below E_F is empty.

3.3 An unknown semiconductor has $E_g = 1.1$ eV and $N_c = N_v$. It is doped with 10^{15} cm^{-3} donors, where the donor level is 0.2 eV below E_c. Given that E_F is 0.25 eV below E_c, calculate n_i and the concentration of electrons and holes in the semiconductor at 300 K.

3.4 At room temperature, an unknown indirect band gap, intrinsic, cubic semiconductor has the following band structure: There are 6 X minima along the ⟨100⟩ directions. If $m_n^*(\Gamma) = 0.065 m_0$, $m_n^*(X) = 0.30\, m_0$ (for each of the X minima) and $m_p^* = 0.47 m_0$, at what temperature is the number of electrons in the Γ minima and the X minima equal if the Γ to X energy separation is 0.35 eV, and the bandgap is 1.7 eV (m_0 = free electron mass)?

CIRRUS EX. 1062 - 62/190

3.5 Since the effective mass of electrons in a conduction band decreases with increasing curvature of the band according to Eq. (3-3), comment on the electron effective mass in the Γ valley of GaAs compared with the indirect X or L valleys. (See Fig. 3-10.) How is this effective mass difference reflected in the electron mobility for GaAs and GaP shown in Appendix III? From Fig. 3-10, what would you expect to happen to the conductivity of GaAs if Γ-valley electrons drifting in an electric field were suddenly promoted to the L valley?

3.6 Calculate the band gap of Si from Eq. (3-23) and the plot of n_i vs. $1000/T$ (Fig. 3-17). *Hint:* The slope cannot be measured directly from a semilogarithmic plot; read the values from two points on the plot and take the natural logarithm as needed for the solution.

3.7 (a) A Si sample is doped with 10^{16} cm^{-3} boron atoms and a certain number of shallow donors. The Fermi level is 0.36 eV above E_i at 300 K. What is the donor concentration N_d?

(b) A Si sample contains 10^{16} cm^{-3} In acceptor atoms and a certain number of shallow donors. The In acceptor level is 0.16 eV above E_v, and E_F is 0.26 eV above E_v at 300 K. How many (cm^{-3}) In atoms are un-ionized (i.e., neutral)?

3.8 Show that Eq. (3-25) results from Eqs. (3-15) and (3-19). If $n_0 = 10^{16}$ cm^{-3}, where is the Fermi level relative to E_i in Si at 300 K?

3.9 Derive an expression relating the intrinsic level E_i to the center of the band gap $E_g/2$. Calculate the displacement of E_i from $E_g/2$ for Si at 300 K, assuming the effective mass values for electrons and holes are $1.1m_0$ and $0.56m_0$, respectively.

3.10 A semiconductor device requires n-type material; it is to be operated at 400 K. Would Si doped with 10^{15} atoms/cm^3 of arsenic be useful in this application? Could Ge doped with 10^{15} cm^{-3} antimony be used?

3.11 A new semiconductor has $N_c = 10^{19}$ cm^{-3}, $N_v = 5 \times 10^{18}$ cm^{-3}, and $E_g = 2$ eV. If it is doped with 10^{17} donors (fully ionized), calculate the electron, hole, and intrinsic carrier concentrations at 627°C. Sketch the simplified band diagram, showing the position of E_F.

3.12 (a) Show that the minimum conductivity of a semiconductor sample occurs when $n_0 = n_i\sqrt{\mu_p/\mu_n}$. *Hint:* begin with Eq. (3-43) and apply Eq. (3-24).

(b) What is the expression for the minimum conductivity σ_{min}?

(c) Calculate σ_{min} for Si at 300 K and compare with the intrinsic conductivity.

3.13 (a) A Si bar 0.1 μm long and 100 μm^2 in cross-sectional area is doped with 10^{17} cm^{-3} phosphorus. Find the current at 300 K with 10 V applied. Repeat for a Si bar 1 μm long.

(b) How long does it take an average electron to drift 1 μm in pure Si at an electric field of 100 V/cm? Repeat for 10^5 V/cm.

3.14. (a) A Si sample is doped with 10^{17} boron atoms/cm^3. What is the electron concentration n_0 at 300 K? What is the resistivity?

(b) A Ge sample is doped with 3×10^{13} Sb atoms/cm^3. Using the requirements of space charge neutrality, calculate the electron concentration n_0 at 300 K.

CIRRUS EX. 1062 - 63/190

3.15 For a Si conductor of length 5 μm, doped n-type at 10^{15} cm^{-3}, calculate the current density for an applied voltage of 2.5 V across its length. How about for a voltage of 2500 V? The electron and hole mobilities are 1500 cm^2/V-s and 500 cm^2/V-s, respectively, in the ohmic region for electric fields below 10^4 V/cm. For higher fields, electrons and holes have a saturation velocity of 10^7 cm/s.

3.16 In a long semiconductor bar ($E_G = 2$ eV), conduction band electrons come in from the left in the positive x direction with a kinetic energy of 3 eV. They move from location A to B to C to D. Between A and B, the electric field is zero; between locations B and C, there is a linearly varying voltage increase of 4 V; between C and D, the field is again zero. Assuming no scattering, sketch a *simplified* band diagram describing the motion of these electrons. Assuming that these electrons can be described as plane waves, with a free-electron mass, write down the wave function of the electrons at D. Leave your result in terms of an arbitrary normalization constant.

3.17 Assume that a conduction electron in Si ($\mu_n = 1350$ cm^2/V-s) has a thermal energy of kT, related to its mean thermal velocity by $E_{th} = (m_0 v_{th}^2)/2$. This electron is placed in an electric field of 100 V/cm. Show that the drift velocity of the electron in this case is small compared with its thermal velocity. Repeat for a field of 10^4 V/cm, using the same value of μ_n. Comment on the actual mobility effects at this higher value of the field.

3.18 Use Eq. (3-45) to calculate and plot the mobility vs. temperature $\mu(T)$ from 10 K to 500 K for Si doped with $N_d = 10^{14}$, 10^{16}, and 10^{18} donors cm^{-3}. Consider the mobility to be determined by impurity and phonon (lattice) scattering. Impurity scattering limited mobility can be described by

$$\mu_I = 3.29 \times 10^{15} \frac{\epsilon_r^2 T^{3/2}}{N_d^+ (m_n^*/m_0)^{1/2} \left[\ln(1+z) - \dfrac{z}{1+z} \right]}$$

where

$$z = 1.3 \times 10^{13} \epsilon_r T^2 (m_n^*/m_0)(N_d^+)^{-1}$$

Assume that the ionized impurity concentration N_d^+ is equal to N_d at all temperatures.

The conductivity effective mass m_n^* for Si is $0.26\, m_0$. Acoustic phonon (lattice) scattering limited mobility can be described by

$$\mu_{AC} = 1.18 \times 10^{-5} c_1 (m_n^*/m_0)^{-5/2} T^{-3/2} (E_{AC})^{-2}$$

where the stiffness (c_1) is given by

$$c_1 = 1.9 \times 10^{12} \text{ dyne cm}^{-2} \text{ for Si}$$

and the conduction band acoustic deformation potential (E_{AC}) is

$$E_{AC} = 9.5 \text{ eV for Si}$$

3.19 Rework Prob. 3.18 considering carrier freeze-out onto donors at low T. That is, consider

$$N_d^+ = \frac{N_d}{1 + \exp(E_d/kT)}$$

CIRRUS EX. 1062 - 64/190

as the ionized impurity concentration. Consider the donor ionization energy (E_d) to be 45 meV for Si.

3.20 Hall measurements are made on a p-type semiconductor bar 500 μm wide and 20 μm thick. The Hall contacts A and B are displaced 2 μm with respect to each other in the direction of current flow of 3 mA. The voltage between A and B with a magnetic field of 10 kG (1kG = 10^{-5} Wb/cm^2) pointing out of the plane of the sample is 3.2 mV. When the magnetic field direction is reversed the voltage changes to −2.8 mV. What is the hole concentration and mobility?

3.21 In soldering wires to a sample such as that shown in Fig. 3–25, it is difficult to align the Hall probes A and B precisely. If B is displaced slightly down the length of the bar from A, an erroneous Hall voltage results. Show that the true Hall voltage V_H can be obtained from two measurements of V_{AB}, with the magnetic field first in the $+z$-direction and then in the $-z$-direction.

3.22 A sample of Si is doped with 10^{17} phosphorus atoms/cm^3. What would you expect to measure for its resistivity? What Hall voltage would you expect in a sample 100 μm thick if I_x = 1 mA and B_z = 1 kG = 10^{-5} Wb/cm^2?

READING LIST

Blakemore, J. S. *Semiconductor Statistics.* New York: Dover Publications, 1987.

Hess, K. *Advanced Theory of Semiconductor Devices,* 2d ed. New York: IEEE Press, 2000.

Kittel, C. *Introduction to Solid State Physics,* 7th ed. New York: Wiley, 1996.

Neamen, D. A. *Semiconductor Physics and Devices: Basic Principles.* Homewood, IL: Irwin, 2003.

Pierret, R. F. *Semiconductor Device Fundamentals.* Reading, MA: Addison-Wesley, 1996.

Schubert, E. F. *Doping in III–V Semiconductors.* Cambridge: Cambridge University Press, 1993.

Singh, J. *Semiconductor Devices.* New York: McGraw-Hill, 1994.

Wang, S. *Fundamentals of Semiconductor Theory and Device Physics.* Englewood Cliffs, NJ: Prentice Hall, 1989.

Wolfe, C. M., G. E. Stillman, and N. Holonyak, Jr. *Physical Properties of Semiconductors.* Englewood Cliffs, NJ: Prentice Hall, 1989.

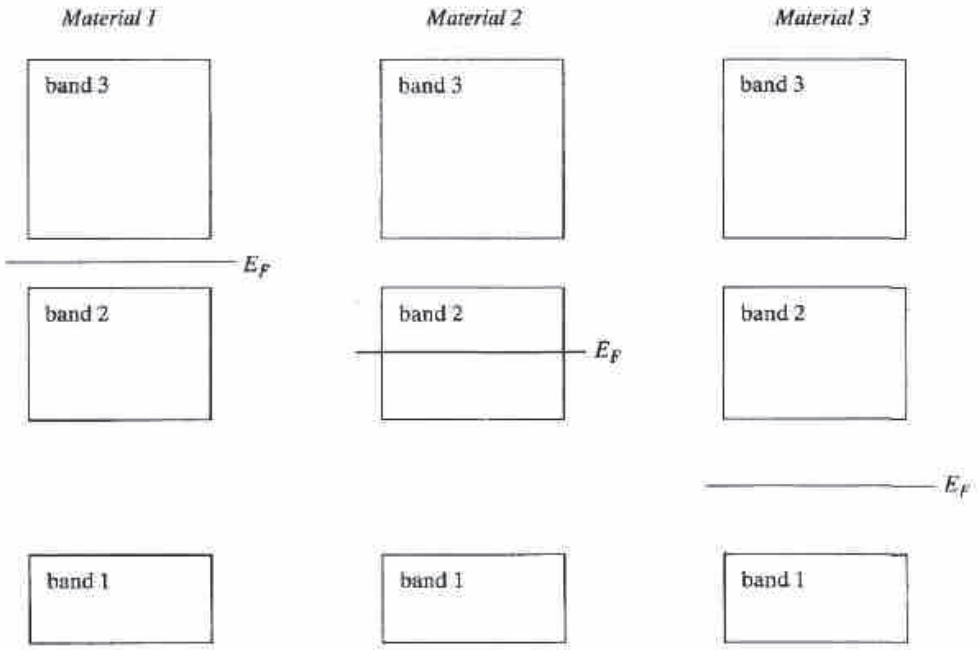
SELF QUIZ

Question 1

(a) The following three diagrams show three different energy bands of some hypothetical crystalline materials (energy varies vertically). The only difference between the three materials is the assumed Fermi level energy E_F. Characterize each material as a ***metal***, an ***insulator***, or a ***semiconductor***.

(b) Assuming you can see through one and only one of the materials of part (a) above, which one would it most likely be? ***Material 1 / Material 2 / Material 3***

CIRRUS EX. 1062 - 65/190

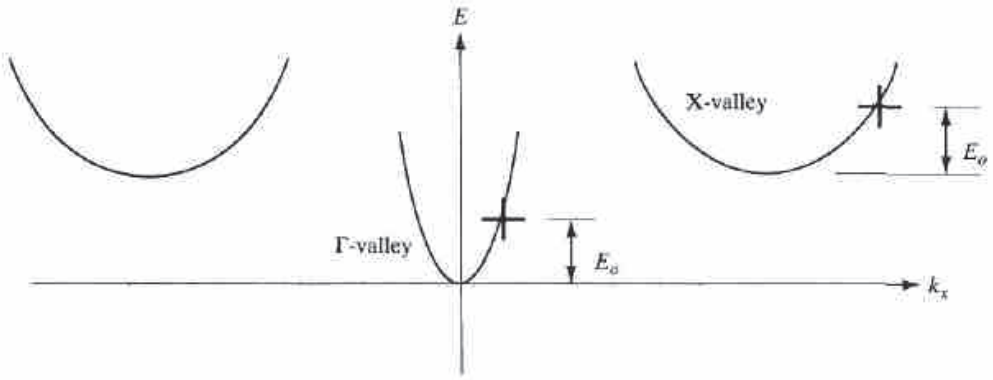


Question 2

Consider the following conduction band energy E vs. wave vector k_x dispersion relation.

(a) Which energy valley has the greater effective mass in the x-direction m_x (circle one)? **Γ-valley / X-valley**

(b) Consider two electrons, one each located at the positions of the heavy crosses. Which has the greater velocity magnitude (circle one)? **The one in the Γ-valley / The one in the X-valley**



CIRRUS EX. 1062 - 66/190

Question 3

These questions refer to the band structures of Si and GaAs shown in Fig. 3–10.

(a) Which material appears to have the *lowest* (conduction band) electron effective mass, Si or GaAs?

(b) Which of these would you expect to produce photons (light) more efficiently through electron–hole recombination?

(c) Consistent with your answer to part (b) and making use of Appendix III, what would you expect the energy of the emitted photons to be? What would be their wavelength in μm? Would these be visible, infrared, or ultraviolet?

(d) How many equivalent conduction band minima do we have for Si? GaAs?

Question 4

Refer to Fig. 3–10, which shows the E vs. $\boldsymbol{k}$ dispersion relations for gallium arsenide (GaAs) and for silicon (Si) along the [111] and [100] directions, showing both valence and conduction bands.

(a) Neglecting differences in electron scattering rates in the two materials, would you expect Si or GaAs to have the greatest electron mobility μ_n?

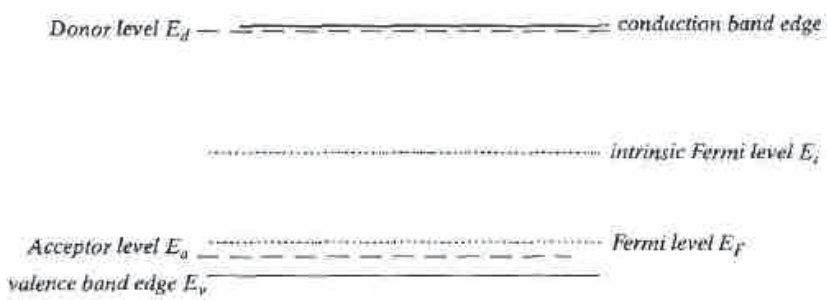
(b) If a constant force were applied in the [100] direction for a short period of time on an electron initially located at the conduction band minimum of each semiconductor and if scattering were neglected, would the magnitude of change in **k** in Si be greater, equal to, or smaller than the magnitude of the change in **k** in GaAs for the same force **F**?

greater/equal/smaller

Question 5

(a) The equilibrium band diagram for a doped direct gap semiconductor is shown below. Is it n-type, p-type, or unknown? Circle one below.

n-type/p-type/not enough information provide



CIRRUS EX. 1062 - 67/190

(b) Based on the band diagram opposite (E_i is in the middle of the gap), would you expect that the conduction band density-of-states effective mass is greater than, equal to, or smaller than the valence band effective mass? Circle one:

Greater than/equal/smaller than

(c) What, if any, of the following conditions by themselves could lead to the above band diagram? Circle each correct answer.

(a) very high temperature

(b) very high acceptor doping

(c) very low acceptor doping

Question 6

A hypothetical semiconductor has an intrinsic carrier concentration of $1.0 \times 10^{10}/cm^3$ at 300 K, it has conduction and valence band effective densities of states N_c and N_v, both equal to $10^{19}/cm^3$.

(a) What is the band gap E_g?

(b) If the semiconductor is doped with $N_d = 1 \times 10^{16}$ donors/cm^3, what are the equilibrium electron and hole concentrations at 300K?

(c) If the same piece of semiconductor, already having $N_d = 1 \times 10^{16}$ donors/cm^3, is also doped with $N_a = 2 \times 10^{16}$ acceptors/cm^3, what are the new equilibrium electron and hole concentrations at 300 K?

(d) Consistent with your answer to part (c), what is the Fermi level position with respect to the intrinsic Fermi level, $E_F - E_i$?

Question 7

What is the difference between density of states and *effective* density of states, and why is the latter such a useful concept?

Question 8

(a) Does mobility have any meaning at very high field? Why?

(b) How do you measure mobility and carrier concentration?

CIRRUS EX. 1062 - 68/190

Chapter 4

Excess Carriers in Semiconductors

OBJECTIVES

1. Understand how photons interact with direct and indirect band gap semiconductors
2. Understand generation–recombination of excess carriers, possibly through trap sites
3. Introduce quasi-Fermi levels in non-equilibrium
4. Calculate diffusion currents from carrier concentration gradients and diffusivity
5. Use the continuity equation to study time dependence of carrier concentrations

Most semiconductor devices operate by the creation of charge carriers in excess of the thermal equilibrium values. These excess carriers can be created by optical excitation or electron bombardment, or as we shall see in Chapter 5, they can be injected across a forward-biased p-n junction. However the excess carriers arise, they can dominate the conduction processes in the semiconductor material. In this chapter we shall investigate the creation of excess carriers by optical absorption and the resulting properties of photoluminescence and photoconductivity. We shall study more closely the mechanism of electron–hole pair recombination and the effects of carrier trapping. Finally, we shall discuss the diffusion of excess carriers due to a carrier gradient, which serves as a basic mechanism of current conduction along with the mechanism of drift in an electric field.

4.1 OPTICAL ABSORPTION[1]

An important technique for measuring the band gap energy of a semiconductor is the absorption of incident photons by the material. In this experiment photons of selected wavelengths are directed at the sample, and relative transmission of the various photons is observed. Since photons with energies greater than the band gap energy are absorbed while photons with

[1]In this context the word "optical" does not necessarily imply that the photons absorbed are in the visible part of the spectrum. Many semiconductors absorb photons in the infrared region, but this is included in the term "optical absorption."

CIRRUS EX. 1062 - 69/190

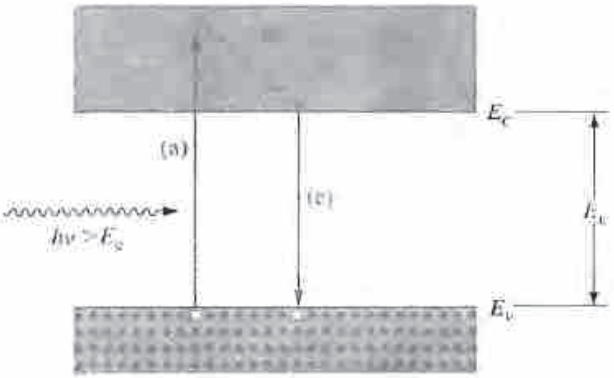


Figure 4–1 Optical absorption of a photon with $h\nu > E_g$: (a) An EHP is created during photon absorption; (b) the excited electron gives up energy to the lattice by scattering events; (c) the electron recombines with a hole in the valence band.

energies less than the band gap are transmitted, this experiment gives an accurate measure of the band gap energy.

It is apparent that a photon with energy $h\nu \geq E_g$ can be absorbed in a semiconductor (Fig. 4–1). Since the valence band contains many electrons and the conduction band has many empty states into which the electrons may be excited, the probability of photon absorption is high. As Fig. 4–1 indicates, an electron excited to the conduction band by optical absorption may initially have more energy than is common for conduction band electrons (almost all electrons are near E_c unless the sample is very heavily doped). Thus the excited electron loses energy to the lattice in scattering events until its velocity reaches the thermal equilibrium velocity of other conduction band electrons. The electron and hole created by this absorption process are *excess carriers;* since they are out of balance with their environment, they must eventually recombine. While the excess carriers exist in their respective bands, however, they are free to contribute to the conductivity of the material.

A photon with energy less than E_g is unable to excite an electron from the valence band to the conduction band. Thus in a pure semiconductor, there is negligible absorption of photons with $h\nu < E_g$. This explains why some materials are transparent in certain wavelength ranges. We are able to "see through" certain insulators, such as a good NaCl crystal, because a large energy gap containing no electron states exists in the material. If the band gap is about 2 eV wide, only long wavelengths (infrared) and the red part of the visible spectrum are transmitted; on the other hand, a band gap of about 3 eV allows infrared and the entire visible spectrum to be transmitted.

If a beam of photons with $h\nu > E_g$ falls on a semiconductor, there will be some predictable amount of absorption, determined by the properties of the material. We would expect the ratio of transmitted to incident light intensity to depend on the photon wavelength and the thickness of the sample. To calculate this dependence, let us assume that a photon beam of intensity $\mathbf{I}_0$ (photons/cm^2-s) is directed at a sample of thickness l (Fig. 4–2). The beam contains only photons of wavelength λ, selected by a monochromator. As the beam passes through the sample, its intensity at a distance x from the surface can be calculated by considering the probability of absorption within

CIRRUS EX. 1062 - 70/190

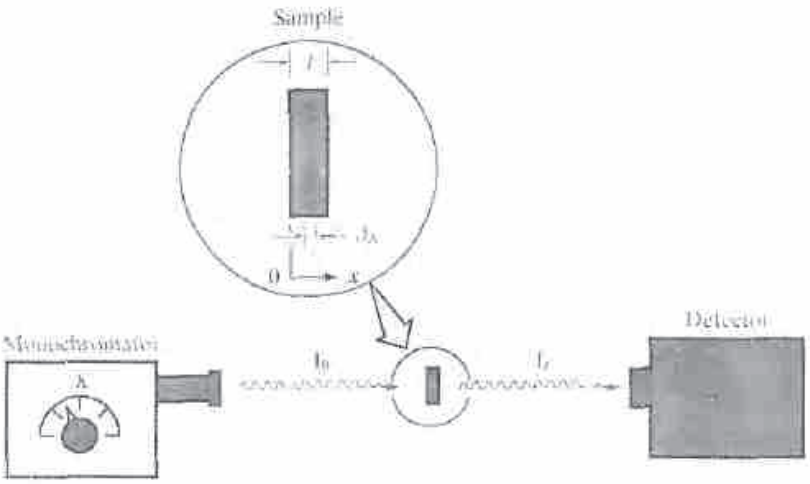


Figure 4–2
Optical absorption experiment.

any increment dx. Since a photon which has survived to x without absorption has no memory of how far it has traveled, its probability of absorption in any dx is constant. Thus the degradation of the intensity $-d\mathbf{I}(x)/dx$ is proportional to the intensity remaining at x:

$$-\frac{d\mathbf{I}(x)}{dx} = \boldsymbol{\alpha}\mathbf{I}(x) \tag{4-1}$$

The solution to this equation is

$$\mathbf{I}(x) = \mathbf{I}_0 e^{-\boldsymbol{\alpha}x} \tag{4-2}$$

and the intensity of light transmitted through the sample thickness l is

$$\mathbf{I}_t = \mathbf{I}_0 e^{-\boldsymbol{\alpha}l} \tag{4-3}$$

The coefficient $\boldsymbol{\alpha}$ is called the *absorption coefficient* and has units of cm^{-1}. This coefficient will of course vary with the photon wavelength and with the material. In a typical plot of $\boldsymbol{\alpha}$ vs. wavelength (Fig. 4-3), there is negligible absorption at long wavelengths ($h\nu$ small) and considerable absorption of photons with energies larger than E_g. According to Eq. (2-2), the relation between photon energy and wavelength is $E = hc/\lambda$. If E is given in electron volts and λ in micrometers, this becomes $E = 1.24/\lambda$.

Figure 4-4 indicates the band gap energies of some of the common semiconductors, relative to the visible, infrared, and ultraviolet portions of the spectrum. We observe that GaAs, Si, Ge, and InSb lie outside the visible region, in the infrared. Other semiconductors, such as GaP and CdS, have band gaps wide enough to pass photons in the visible range. It is important to note here that a semiconductor absorbs photons with energies equal to the band gap, *or larger*. Thus Si absorbs not only band gap light (~1 μm) but also shorter wavelengths, including those in the visible part of the spectrum.

CIRRUS EX. 1062 - 71/190

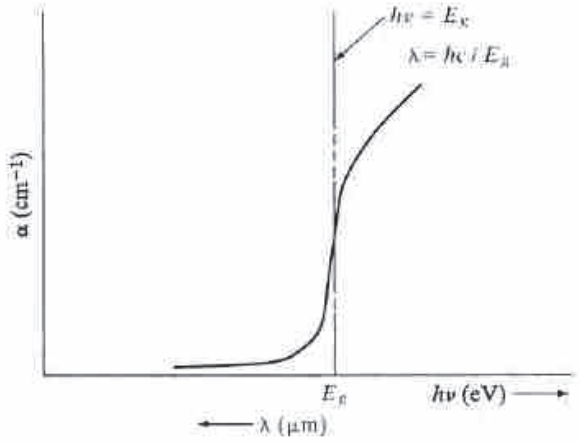


Figure 4–3
Dependence of optical absorption coefficient α for a semiconductor on the wavelength of incident light.

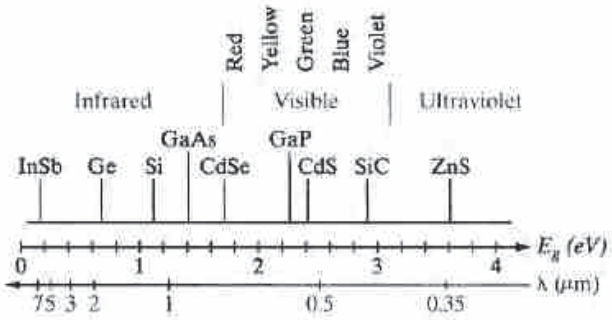


Figure 4–4
Band gaps of some common semiconductors relative to the optical spectrum.

4.2 LUMINESCENCE

When electron-hole pairs are generated in a semiconductor, or when carriers are excited into higher impurity levels from which they fall to their equilibrium states, light can be given off by the material. Many of the semiconductors are well suited for light emission, particularly the compound semiconductors with direct band gaps. The general property of light emission is called *luminescence.*[2] This overall category can be subdivided according to the excitation mechanism: If carriers are excited by photon absorption, the radiation resulting from the recombination of the excited carriers is called *photoluminescence;* if the excited carriers are created by high-energy electron bombardment of the material, the mechanism is called *cathodoluminescence;* if the excitation occurs by the introduction of current into the sample, the resulting luminescence is called *electroluminescence.* Other types of excitation are possible, but these three are the most important for device applications.

[2]The emission processes considered here should not be confused with radiation due to incandescence which occurs in heated materials. The various luminescent mechanisms can be considered "cold" processes as compared to the "hot" process of incandescence, which increases with temperature. In fact, most luminescent processes become more efficient as the temperature is lowered.

CIRRUS EX. 1062 - 72/190

4.2.1 Photoluminescence

The simplest example of light emission from a semiconductor occurs for direct excitation and recombination of an EHP, as depicted in Fig. 3-5a. If the recombination occurs directly rather than via a defect level, band gap light is given off in the process. For steady state excitation, the recombination of EHPs occurs at the same rate as the generation, and one photon is emitted for each photon absorbed. Direct recombination is a fast process; the mean lifetime of the EHP is usually on the order of 10^{-8} s or less. Thus the emission of photons stops within approximately 10^{-8} s after the excitation is turned off. Such fast luminescent processes are often referred to as *fluorescence*. In some materials, however, emission continues for periods up to seconds or minutes after the excitation is removed. These slow processes are called *phosphorescence*, and the materials are called *phosphors*. An example of a slow process is shown in Fig. 4-5. This material contains a defect level (perhaps due to an impurity) in the band gap which has a strong tendency to temporarily capture (*trap*) electrons from the conduction band. The events depicted in the figure are as follows: (a) An incoming photon with $h\nu_1 > E_g$ is absorbed, creating an EHP; (b) the excited electron gives up energy to the lattice by scattering until it nears the bottom of the conduction band; (c) the electron is *trapped* by the impurity level E_t and remains trapped until it can be thermally reexcited to the conduction band (d); (e) finally direct recombination occurs as the electron falls to an empty state in the valence band, giving off a photon ($h\nu_2$) of approximately the band gap energy. The delay time between excitation and recombination can be relatively long if the probability of thermal reexcitation from the trap (d) is small. Even longer delay times result if the electron is retrapped several times before recombination. If the trapping probability is greater than the probability of recombination, an electron may make several trips between the trap and the conduction band before recombination finally occurs. In such material the emission of phosphorescent light persists for a relatively long time after the excitation is removed.

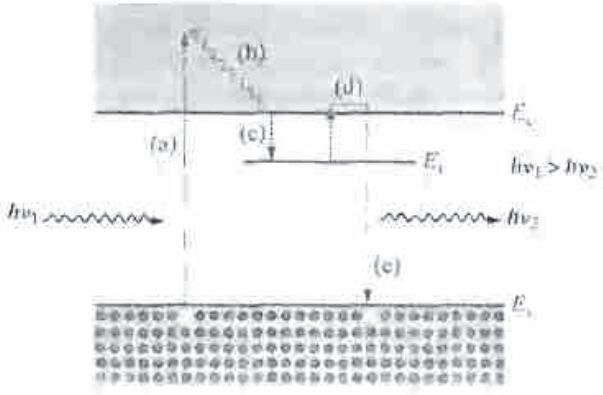


Figure 4–5 Excitation and recombination mechanisms in photoluminescence with a trapping level for electrons.

CIRRUS EX. 1062 - 73/190

The color of light emitted by a phosphor such as ZnS depends primarily on the impurities present, since many radiative transitions involve impurity levels within the band gap. This selection of colors is particularly useful in the fabrication of a color television screen.

One of the most common examples of photoluminescence is the fluorescent lamp. Typically such a lamp is composed of a glass tube filled with gas (e.g., a mixture of argon and mercury), with a fluorescent coating on the inside of the tube. When an electric discharge is induced between electrodes in the tube, the excited atoms of the gas emit photons, largely in the visible and ultraviolet regions of the spectrum. This light is absorbed by the luminescent coating, and the visible photons are emitted. The efficiency of such a lamp is considerably better than that of an incandescent bulb, and the wavelength mixture in light given off can be adjusted by proper selection of the fluorescent material.

EXAMPLE 4–1

A 0.46-μm-thick sample of GaAs is illuminated with monochromatic light of $h\nu = 2$ eV. The absorption coefficient α is 5×10^4 cm^{-1}. The power incident on the sample is 10 mW.

(a) Find the total energy absorbed by the sample per second (J/s).

(b) Find the rate of excess thermal energy given up by the electrons to the lattice before recombination (J/s).

(c) Find the number of photons per second given off from recombination events, assuming perfect quantum efficiency.

SOLUTION

(a) From Eq. (4–3),

$$I_t = I_0 e^{-\alpha t} = 10^{-2} \exp(-5 \times 10^4 \times 0.46 \times 10^{-4})$$
$$= 10^{-2} e^{-2.3} = 10^{-3} \text{ W}$$

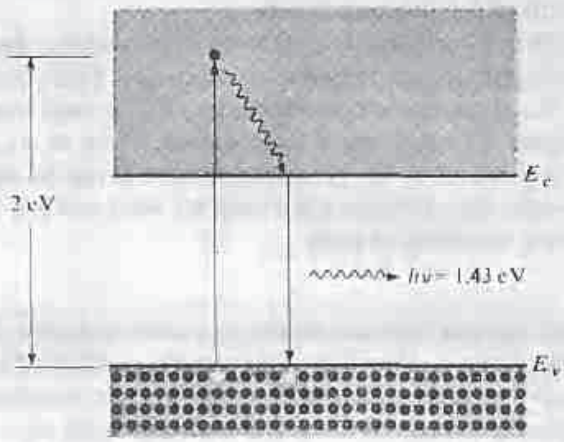


Figure 4–6 Excitation and band-to-band recombination leading to photoluminescence.

CIRRUS EX. 1062 - 74/190

Thus the absorbed power is

$$10 - 1 = 9\ \text{mW} = 9 \times 10^{-3}\ \text{J/s}$$

(b) The fraction of each photon energy unit which is converted to heat is

$$\frac{2 - 1.43}{2} = 0.285$$

Thus the amount of energy converted to heat per second is

$$0.285 \times 9 \times 10^{-3} = 2.57 \times 10^{-3}\ \text{J/s}$$

(c) Assuming one emitted photon for each photon absorbed (perfect quantum efficiency), we have

$$\frac{9 \times 10^{-3}\ \text{J/s}}{1.6 \times 10^{-19}\ \text{J/eV} \times 2\ \text{eV/photon}} = 2.81 \times 10^{16}\ \text{photons/s}$$

Alternative solution: Recombination radiation accounts for 9 - 2.57 = 6.43 mW at 1.43 eV/photon.

$$\frac{6.43 \times 10^{-3}}{1.6 \times 10^{-19} \times 1.43} = 2.81 \times 10^{16}\ \text{photons/s}$$

4.2.2 Electroluminescence

There are many ways by which electrical energy can be used to generate photon emission in a solid. In LEDs an electric current causes the injection of minority carriers into regions of the crystal where they can recombine with majority carriers, resulting in the emission of recombination radiation. This important effect (*injection electroluminescence*) will be discussed in Chapter 8 in terms of p-n junction theory.

The first electroluminescent effect to be observed was the emission of photons by certain phosphors in an alternating electric field (the Destriau effect). In this device, a phosphor powder such as ZnS is held in a binder material (often a plastic) of a high dielectric constant. When an a-c electric field is applied, light is given off by the phosphor. Such cells can be useful as lighting panels, although their efficiency has thus far been too low for most applications and their reliability is poor.

4.3 CARRIER LIFETIME AND PHOTOCONDUCTIVITY

When excess electrons and holes are created in a semiconductor, there is a corresponding increase in the conductivity of the sample as indicated by Eq. (3-43). If the excess carriers arise from optical luminescence, the resulting increase in conductivity is called *photoconductivity*. This is an important effect, with useful applications in the analysis of semiconductor materials and in the operation of

CIRRUS EX. 1062 - 75/190

several types of devices. In this section we shall examine the mechanisms by which excess electrons and holes recombine and apply the recombination kinetics to the analysis of photoconductive devices. However, the importance of recombination is not limited to cases in which the excess carriers are created optically. In fact, virtually every semiconductor device depends in some way on the recombination of excess electrons and holes. Therefore, the concepts developed in this section will be used extensively in the analyses of diodes, transistors, lasers, and other devices in later chapters.

4.3.1 Direct Recombination of Electrons and Holes

It was pointed out in Section 3.1.4 that electrons in the conduction band of a semiconductor may make transitions to the valence band (i.e., recombine with holes in the valence band) either directly or indirectly. In direct recombination, an excess population of electrons and holes decays by electrons falling from the conduction band to empty states (holes) in the valence band. Energy lost by an electron in making the transition is given up as a photon. Direct recombination occurs *spontaneously;* that is, the probability that an electron and a hole will recombine is constant in time. As in the case of carrier scattering, this constant probability leads us to expect an exponential solution for the decay of the excess carriers. In this case the rate of decay of electrons at any time t is proportional to the number of electrons remaining at t and the number of holes, with some constant of proportionality for recombination, α_r. The *net* rate of change in the conduction band electron concentration is the thermal generation rate $\alpha_r n_i^2$ from Eq. (3–7) minus the recombination rate

$$\frac{dn(t)}{dt} = \alpha_r n_i^2 - \alpha_r n(t)p(t) \tag{4–4}$$

Let us assume the excess electron–hole population is created at $t = 0$, for example by a short flash of light, and the initial excess electron and hole concentrations Δn and Δp are equal.[3] Then as the electrons and holes recombine in pairs, the instantaneous concentrations of excess carriers $\delta n(t)$ and $\delta p(t)$ are also equal. Thus we can write the total concentrations of Eq. (4–4) in terms of the equilibrium values n_0 and p_0 and the excess carrier concentrations $\delta n(t) = \delta p(t)$. Using Eq. (3–24) we have

$$\begin{aligned}\frac{d\delta n(t)}{dt} &= \alpha_r n_i^2 - \alpha_r[n_0 + \delta n(t)][p_0 + \delta p(t)] \\ &= -\alpha_r[(n_0 + p_0)\delta n(t) + \delta n^2(t)]\end{aligned} \tag{4–5}$$

[3]We will use $\delta n(t)$ and $\delta p(t)$ to mean instantaneous excess carrier concentrations, and Δn, Δp for their values at $t = 0$. Later we will use similar symbolism for spatial distributions, such as $\delta n(x)$ and $\Delta n(x = 0)$.

CIRRUS EX. 1062 - 76/190

This nonlinear equation would be difficult to solve in its present form. Fortunately, it can be simplified for the case of low-level injection. If the excess carrier concentrations are small, we can neglect the δn^2 term. Furthermore, if the material is extrinsic, we can usually neglect the term representing the equilibrium minority carriers. For example, if the material is p-type ($p_0 \gg n_0$), Eq. (4–5) becomes

$$\frac{d\delta n(t)}{dt} = -\alpha_r p_0 \delta n(t) \tag{4–6}$$

The solution to this equation is an exponential decay from the original excess carrier concentration Δn:

$$\delta n(t) = \Delta n e^{-\alpha_r p_0 t} = \Delta n e^{-t/\tau_n} \tag{4–7}$$

Excess electrons in a p-type semiconductor recombine with a decay constant $\tau_n = (\alpha_r p_0)^{-1}$, called the *recombination lifetime*. Since the calculation is made in terms of the minority carriers, τ_n is often called the *minority carrier lifetime*. The decay of excess holes in n-type material occurs with $\tau_p = (\alpha_r n_0)^{-1}$. In the case of direct recombination, the excess majority carriers decay at exactly the same rate as the minority carriers.

There is a large percentage change in the minority carrier electron concentration in Example 4–2 and a small percentage change in the majority hole concentration. Basically, the approximations of extrinsic material and low-level injection allow us to represent $n(t)$ in Eq. (4–4) by the excess concentration $\delta n(t)$ and $p(t)$ by the equilibrium value p_0. Figure 4–7 indicates that this is a good approximation for the example. A more general expression for the carrier lifetime is

$$\tau_n = \frac{1}{\alpha_r(n_0 + p_0)} \tag{4–8}$$

This expression is valid for n- or p-type material if the injection level is low.

EXAMPLE 4–2 A numerical example may be helpful in visualizing the approximations made in the analysis of direct recombination. Let us assume a sample of GaAs is doped with 10^{15} acceptors/cm^3. The intrinsic carrier concentration of GaAs is approximately 10^6 cm^{-3}; thus the minority electron concentration is $n_0 = n_i^2/p_0 = 10^{-3}$ cm^{-3}. Certainly the approximation of $p_0 \gg n_0$ is valid in this case. Now if 10^{14} EHP/cm^3 are created at $t = 0$, we can calculate the decay of these carriers in time. The approximation of $\delta n \ll p_0$ is reasonable, as Fig. 4–7 indicates. This figure shows the decay in time of the excess populations for a carrier recombination lifetime of $\tau_n = \tau_p = 10^{-8}$ s.

CIRRUS EX. 1062 - 77/190

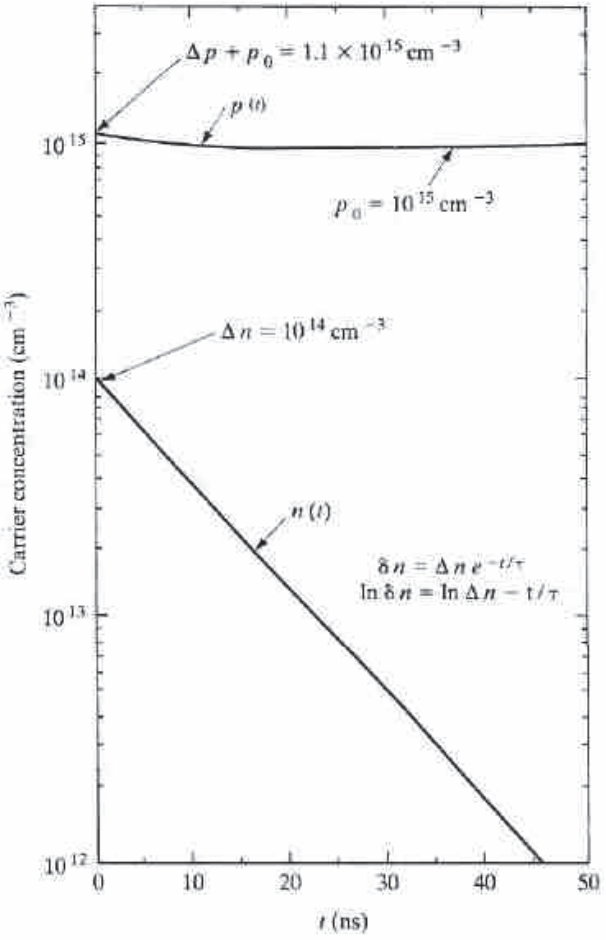


Figure 4–7
Decay of excess electrons and holes by recombination, for $\Delta n = \Delta p = 0.1 p_0$, with n_0 negligible, and $\tau = 10$ ns (Example 4–2). The exponential decay of $\delta n(t)$ is linear on this semilogarithmic graph.

4.3.2 Indirect Recombination; Trapping

In column IV semiconductors and in certain compounds, the probability of direct electron–hole recombination is very small (Appendix III). There is some band gap light given off by materials such as Si and Ge during recombination, but this radiation is very weak and may be detected only by sensitive equipment. The vast majority of the recombination events in indirect materials occur via *recombination levels* within the band gap, and the resulting energy loss by recombining electrons is usually given up to the lattice as heat rather than by the emission of photons. Any impurity or lattice defect can serve as a recombination center if it is capable of receiving a carrier of one type and subsequently capturing the opposite type of carrier, thereby annihilating the pair. For example, Fig. 4–8 illustrates a recombination level E_r which is below E_F at equilibrium and therefore is substantially filled with electrons. When excess electrons and holes are created in this material, each EHP recombines at E_r in two steps: (a) hole capture and (b) electron capture.

CIRRUS EX. 1062 - 78/190

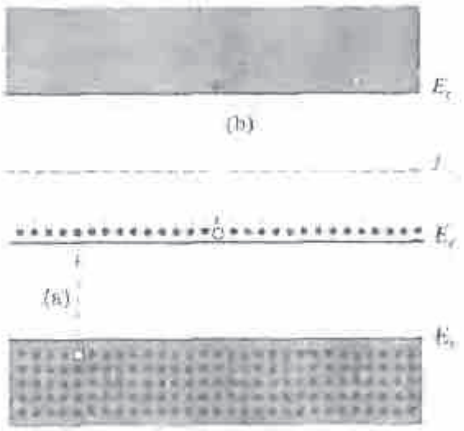


Figure 4–8
Capture processes at a recombination level: (a) hole capture at a filled recombination center; (b) electron capture at an empty center.

Since the recombination centers in Fig. 4–8 are filled at equilibrium, the first event in the recombination process is hole capture. It is important to note that this event is equivalent to an electron at E_t falling to the valence band, leaving behind an empty state in the recombination level. Thus in hole capture, energy is *given up* as heat to the lattice. Similarly, energy is given up when a conduction band electron subsequently falls to the empty state in E_t. When both of these events have occurred, the recombination center is back to its original state (filled with an electron), but an EHP is missing. Thus one EHP recombination has taken place, and the center is ready to participate in another recombination event by capturing a hole.

The carrier lifetime resulting from indirect recombination is somewhat more complicated than is the case for direct recombination, since it is necessary to account for unequal times required for capturing each type of carrier. In particular, recombination is often delayed by the tendency for a captured carrier to be thermally reexcited to its original band before capture of the opposite type of carrier can occur (Section 4.2.1). For example, if electron capture (b) does not follow immediately after hole capture (a) in Fig. 4–8, the hole may be thermally reexcited to the valence band. Energy is required for this process, which is equivalent to a valence band electron being raised to the empty state in the recombination level. This process delays the recombination, since the hole must be captured again before recombination can be completed.

When a carrier is trapped temporarily at a center and then is reexcited without recombination taking place, the process is often called *temporary trapping*. Although the nomenclature varies somewhat, it is common to refer to an impurity or defect center as a *trapping center* (or simply *trap*) if, after capture of one type of carrier, the most probable next event is reexcitation. If the most probable next event is capture of the opposite type of carrier, the center is predominately a recombination center. The recombination can be slow or fast, depending on the average time the first carrier is held before the second carrier is captured. In general, trapping levels located deep in the

CIRRUS EX. 1062 - 79/190

band gap are slower in releasing trapped carriers than are the levels located near one of the bands. This results from the fact that more energy is required, for example, to reexcite a trapped electron from a center near the middle of the gap to the conduction band than is required to reexcite an electron from a level closer to the conduction band.

As an example of impurity levels in semiconductors, Fig. 4-9[4] shows the energy level positions of various impurities in Si. In this diagram a superscript indicates whether the impurity is positive (donor) or negative (acceptor) when ionized. Some impurities introduce multiple levels in the band gap; for example, Zn introduces a level (Zn^-) located 0.31 eV above the valence band and a second level ($Zn^=$) near the middle of the gap. Each Zn impurity atom is capable of accepting two electrons from the semiconductor, one in the lower level and then one in the upper level.

The effects of recombination and trapping can be measured by a *photoconductive decay* experiment. As Fig. 4-7 shows, a population of excess electrons and holes disappears with a decay constant characteristic of the particular recombination process. The conductivity of the sample during the decay is

$$\sigma(t) = q[n(t)\mu_n + p(t)\mu_p] \tag{4-9}$$

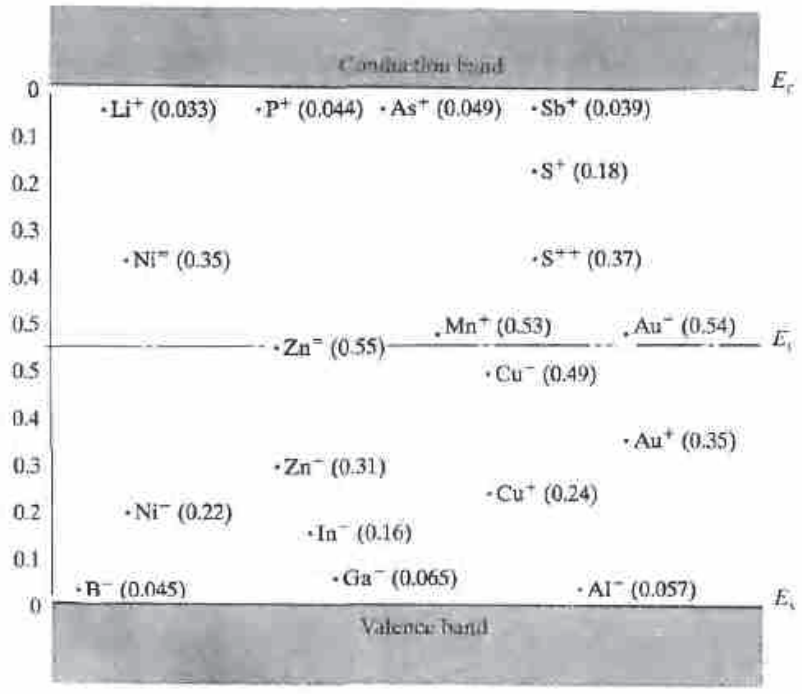


Figure 4-9
Energy levels of impurities in Si. The energies are measured from the nearest band edge (E_v or E_c); donor levels are designated by a plus sign and acceptors by a minus sign.

[4]References: S. M. Sze and J. C. Irvin, "Resistivity, Mobility, and Impurity Levels in GaAs, Ge and Si at 300 K," *Solid State Electronics*, vol. 11, pp. 599–602 (June 1968); E. Schibli and A. G. Milnes, "Deep Impurities in Silicon," *Materials Science and Engineering*, vol. 2, pp. 173–180 (1967).

CIRRUS EX. 1062 - 80/190

Therefore, the time dependence of the carrier concentrations can be monitored by recording the sample resistance as a function of time. A typical experimental arrangement is shown schematically in Fig. 4-10. A source of short pulses of light is required, along with an oscilloscope for displaying the sample voltage as the resistance varies. Microsecond light pulses can be obtained by periodically discharging a capacitor through a flash tube containing a gas such as xenon. For shorter pulses, special techniques such as the use of a pulsed laser must be used.

4.3.3 Steady State Carrier Generation; Quasi-Fermi Levels

In the previous discussion we emphasized the transient decay of an excess EHP population. However, the various recombination mechanisms are also important in a sample at thermal equilibrium or with a steady state EHP generation–recombination balance.[5] For example, a semiconductor at equilibrium experiences thermal generation of EHPs at a rate $g(T) = g_i$ described by Eq. (3-7). This generation is balanced by the recombination rate so that the equilibrium concentrations of carriers n_0 and p_0 are maintained:

$$g(T) = \alpha_r n_i^2 = \alpha_r n_0 p_0 \tag{4-10}$$

This equilibrium rate balance can include generation from defect centers as well as band-to-band generation.

If a steady light is shone on the sample, an optical generation rate g_{op} will be added to the thermal generation, and the carrier concentrations n and p will increase to new steady state values. We can write the balance between generation and recombination in terms of the equilibrium carrier concentrations and the departures from equilibrium δn and δp:

$$g(T) + g_{op} = \alpha_r np = \alpha_r(n_0 + \delta n)(p_0 + \delta p) \tag{4-11}$$

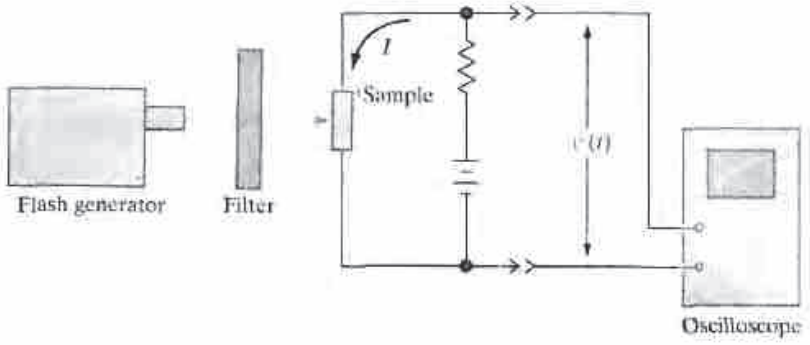


Figure 4-10 Experimental arrangement for photoconductive decay measurements, and a typical oscilloscope trace.

[5]The term *equilibrium* refers to a condition of no external excitation except for temperature, and no net motion of charge (e.g., a sample at a constant temperature, in the dark, with no fields applied). *Steady state* refers to a nonequilibrium condition in which all processes are constant and are balanced by opposing processes (e.g., a sample with a constant current or a constant optical generation of EHPs just balanced by recombination).

CIRRUS EX. 1062 - 81/190

For steady state recombination and no trapping, $\delta n = \delta p$; thus Eq. (4–11) becomes

$$g(T) + g_{op} = \alpha_r n_0 p_0 + \alpha_r[(n_0 + p_0)\delta n + \delta n^2] \tag{4–12}$$

The term $\alpha_r n_0 p_0$ is just equal to the thermal generation rate $g(T)$. Thus, neglecting the δn^2 term for low-level excitation, we can rewrite Eq. (4–12) as

$$g_{op} = \alpha_r(n_0 + p_0)\delta n = \frac{\delta n}{\tau_n} \tag{4–13}$$

The excess carrier concentration can be written as

$$\delta n = \delta p = g_{op}\tau_n \tag{4–14}$$

More general expressions are given in Eq. (4–16), which allow for the case $\tau_p \neq \tau_n$, when trapping is present.

EXAMPLE 4–3

As a numerical example, let us assume that 10^{13} EHP/cm^3 are created optically every microsecond in a Si sample with $n_0 = 10^{14}$ cm^{-3} and $\tau_n = \tau_p = 2$ μsec. The steady state excess electron (or hole) concentration is then 2×10^{13} cm^{-3} from Eq. (4–14). While the percentage change in the majority electron concentration is small, the minority carrier concentration changes from

$$p_0 = n_i^2/n_0 = (2.25 \times 10^{20})/10^{14} = 2.25 \times 10^6 \text{ cm}^{-3} \qquad (\textit{equilibrium})$$

to

$$p = 2 \times 10^{13} \text{ cm}^{-3} \qquad (\textit{steady state})$$

Note that the equilibrium equation $n_0 p_0 = n_i^2$ cannot be used with the subscripts removed; that is, $np \neq n_i^2$ when excess carriers are present.

It is often desirable to refer to the steady state electron and hole concentrations in terms of Fermi levels, which can be included in band diagrams for various devices. The Fermi level E_F used in Eq. (3–25) is meaningful only when no excess carriers are present. However, we can write expressions for the steady state concentrations in the same *form* as the equilibrium expressions by defining separate *quasi-Fermi levels* F_n and F_p for electrons and holes. The resulting carrier concentration equations

$$\boxed{\begin{aligned} n &= n_i e^{(F_n - E_i)/kT} \\ p &= n_i e^{(E_i - F_p)/kT} \end{aligned}} \tag{4–15}$$

can be considered as defining relations for the quasi-Fermi levels.[6]

[6] In some texts the quasi-Fermi level is called *IMREF*, which is Fermi spelled backward.

CIRRUS EX. 1062 - 82/190

EXAMPLE 4–4 In Example 4–3, the steady state electron concentration is

$$n = n_0 + \delta n = 1.2 \times 10^{14} = (1.5 \times 10^{10}) e^{(F_n - E_i)/0.0259}$$

where $kT \simeq 0.0259$ eV at room temperature. Thus the electron quasi-Fermi level position $F_n - E_i$ is found from

$$F_n - E_i = 0.0259 \ln(8 \times 10^3) = 0.233 \text{ eV}$$

and F_n lies 0.233 eV above the intrinsic level. By a similar calculation, the hole quasi-Fermi level lies 0.186 eV below E_i (Fig. 4–11). In this example, the equilibrium Fermi level is $0.0259 \ln(6.67 \times 10^3) = 0.228$ eV above the intrinsic level.

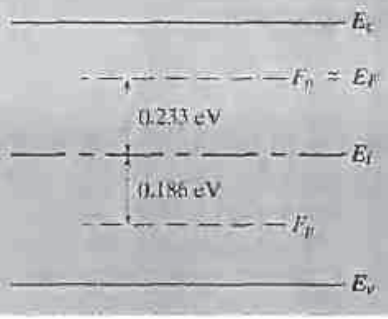


Figure 4–11 Quasi-Fermi levels F_n and F_p for a Si sample with $n_0 = 10^{14}$ cm^{-3}, $\tau_p = 2$ μs, and $g_{op} = 10^{19}$ EHP/cm^3-s (Example 4–4).

The quasi-Fermi levels of Fig. 4–11 illustrate dramatically the deviation from equilibrium caused by the optical excitation; the steady state F_n is only slightly above the equilibrium E_F, whereas F_p is greatly displaced below E_F. From the figure it is obvious that the excitation causes a large percentage change in minority carrier hole concentration and a relatively small change in the electron concentration.

In summary, the quasi-Fermi levels F_n and F_p are the steady state analogues of the equilibrium Fermi level E_F. When excess carriers are present, the deviations of F_n and F_p from E_F indicate how far the electron and hole populations are from the equilibrium values n_0 and p_0. A given concentration of excess EHPs causes a large shift in the minority carrier quasi-Fermi level compared with that for the majority carriers. The separation of the quasi-Fermi levels $F_n - F_p$ is a direct measure of the deviation from equilibrium (at equilibrium $F_n = F_p = E_F$). The concept of quasi-Fermi levels is very useful in visualizing minority and majority carrier concentrations in devices where these quantities vary with position.

4.3.4 Photoconductive Devices

There are a number of applications for devices which change their resistance when exposed to light. For example, such light detectors can be used in the

CIRRUS EX. 1062 - 83/190

home to control automatic night lights which turn on at dusk and turn off at dawn. They can also be used to measure illumination levels, as in exposure meters for cameras. Many systems include a light beam aimed at the photoconductor, which signals the presence of an object between the source and detector. Such systems are useful in moving-object counters, burglar alarms, and many other applications. Detectors are used in optical signaling systems in which information is transmitted by a light beam and is received at a photoconductive cell.

Considerations in choosing a photoconductor for a given application include the sensitive wavelength range, time response, and optical sensitivity of the material. In general, semiconductors are most sensitive to photons with energies equal to the band gap or slightly more energetic than band gap. Less energetic photons are not absorbed, and photons with $h\nu \gg E_g$ are absorbed at the surface and contribute little to the bulk conductivity. Therefore, the table of band gaps (Appendix III) indicates the photon energies to which most semiconductor photodetectors respond. For example, CdS ($E_g = 2.42$ eV) is commonly used as a photoconductor in the visible range, and narrow-gap materials such as Ge (0.67 eV) and InSb (0.18 eV) are useful in the infrared portion of the spectrum. Some photoconductors respond to excitations of carriers from impurity levels within the band gap and therefore are sensitive to photons of less than band gap energy.

The optical sensitivity of a photoconductor can be evaluated by examining the steady state excess carrier concentrations generated by an optical generation rate g_{op}. If the mean time each carrier spends in its respective band before capture is τ_n and τ_p, we have

$$\delta n = \tau_n g_{op} \quad \text{and} \quad \delta p = \tau_p g_{op} \tag{4-16}$$

and the photoconductivity change is

$$\Delta\sigma = q g_{op}(\tau_n \mu_n + \tau_p \mu_p) \tag{4-17}$$

For simple recombination, τ_n and τ_p will be equal. If trapping is present, however, one of the carriers may spend little time in its band before being trapped. From Eq. (4-17) it is obvious that for maximum photoconductive response, we want high mobilities and long lifetimes. Some semiconductors are especially good candidates for photoconductive devices because of their high mobility; for example, InSb has an electron mobility of about 10^5 cm^2/V-s and therefore is used as a sensitive infrared detector in many applications.

The time response of a photoconductive cell is limited by the recombination times, the degree of carrier trapping, and the time required for carriers to drift through the device in an electric field. Often these properties can be adjusted by proper choice of material and device geometry, but in some cases improvements in response time are made at the expense of sensitivity. For example, the drift time can be reduced by making the device short, but this

CIRRUS EX. 1062 - 84/190

substantially reduces the responsive area of the device. In addition, it is often desirable that the device have a large dark resistance, and for this reason, shortening the length may not be practical. There is usually a compromise between sensitivity, response time, dark resistance, and other requirements in choosing a device for a particular application.

4.4 DIFFUSION OF CARRIERS

When excess carriers are created nonuniformly in a semiconductor, the electron and hole concentrations vary with position in the sample. Any such spatial variation (*gradient*) in n and p calls for a net motion of the carriers from regions of high carrier concentration to regions of low carrier concentration. This type of motion is called *diffusion* and represents an important charge transport process in semiconductors. The two basic processes of current conduction are diffusion due to a carrier gradient and drift in an electric field.

4.4.1 Diffusion Processes

When a bottle of perfume is opened in one corner of a closed room, the scent is soon detected throughout the room. If there is no convection or other net motion of air, the scent spreads by diffusion. The diffusion is the natural result of the *random motion* of the individual molecules. Consider, for example, a volume of arbitrary shape with scented air molecules inside and unscented molecules outside the volume. All the molecules undergo random thermal motion and collisions with other molecules. Thus each molecule moves in an arbitrary direction until it collides with another air molecule, after which it moves in a new direction. If the motion is truly random, a molecule at the edge of the volume has equal probabilities of moving into or out of the volume on its next step (assuming the curvature of the surface is negligible on the molecular scale). Therefore, after a mean free time $\bar{\tau}$, half the molecules at the edge will have moved into the volume and half will have moved out of the volume. The net effect is that the volume containing scented molecules has increased. This process will continue until the molecules are uniformly distributed in the room. Only then will a given volume gain as many molecules as it loses in a given time. In other words, net diffusion will continue as long as gradients exist in the distribution of scented molecules.

Carriers in a semiconductor diffuse in a carrier gradient by random thermal motion and scattering from the lattice and impurities. For example, a pulse of excess electrons injected at $x = 0$ at time $t = 0$ will spread out in time as shown in Fig. 4-12. Initially, the excess electrons are concentrated at $x = 0$; as time passes, however, electrons diffuse to regions of low electron concentration until finally $n(x)$ is constant.

We can calculate the rate at which the electrons diffuse in a one-dimensional problem by considering an arbitrary distribution $n(x)$ such as Fig. 4-13a. Since the mean free path $\bar{l}$ between collisions is a small

CIRRUS EX. 1062 - 85/190

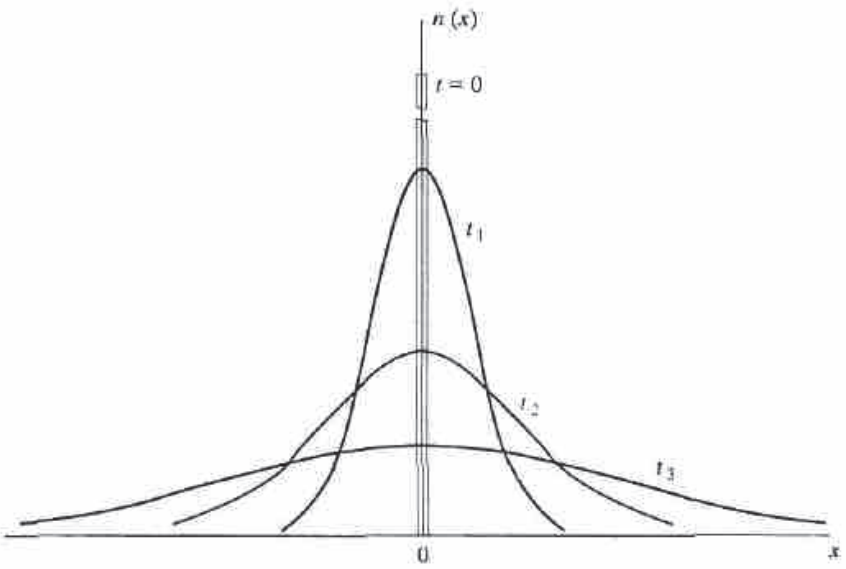


Figure 4–12
Spreading of a pulse of electrons by diffusion.

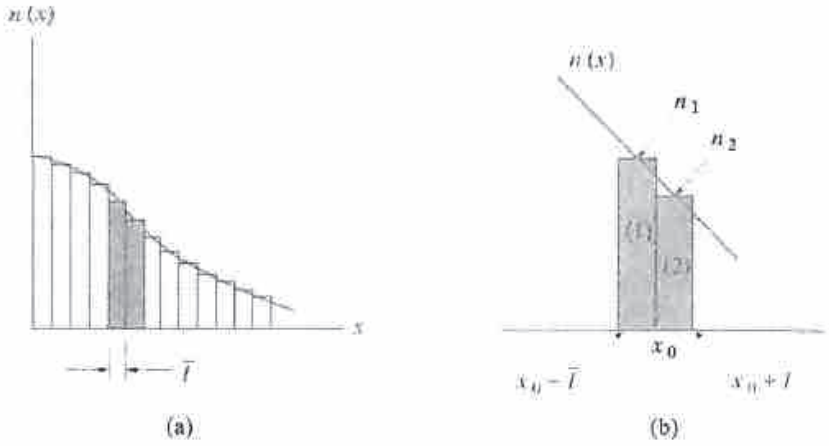


Figure 4–13
An arbitrary electron concentration gradient in one dimension: (a) division of $n(x)$ into segments of length equal to a mean free path for the electrons; (b) expanded view of two of the segments centered at x_0.

incremental distance, we can divide x into segments $\bar{l}$ wide, with $n(x)$ evaluated at the center of each segment (Fig. 4–13b).

The electrons in segment (1) to the left of x_0 in Fig. 4–13b have equal chances of moving left or right, and in a mean free time $\bar{t}$ one-half of them will move into segment (2). The same is true of electrons within one mean free path of x_0 to the right; one-half of these electrons will move through x_0 from right to left in a mean free time. Therefore, the *net* number of electrons passing x_0 from left to right in one mean free time is $\frac{1}{2}(n_1\bar{l}A) - \frac{1}{2}(n_2\bar{l}A)$, where the

CIRRUS EX. 1062 - 86/190

area perpendicular to x is A. The rate of electron flow in the $+x$-direction per unit area (the electron flux density ϕ_n) is given by

$$\phi_n(x_0) = \frac{\bar{l}}{2\bar{t}}(n_1 - n_2) \tag{4-18}$$

Since the mean free path $\bar{l}$ is a small differential length, the difference in electron concentration $(n_1 - n_2)$ can be written as

$$n_1 - n_2 = \frac{n(x) - n(x + \Delta x)}{\Delta x}\bar{l} \tag{4-19}$$

where x is taken at the center of segment (1) and $\Delta x = \bar{l}$. In the limit of small Δx (i.e., small mean free path $\bar{l}$ between scattering collisions), Eq. (4-18) can be written in terms of the carrier gradient $dn(x)/dx$:

$$\phi_n(x) = \frac{\bar{l}^2}{2\bar{t}} \lim_{\Delta x \to 0} \frac{n(x) - n(x + \Delta x)}{\Delta x} = \frac{-\bar{l}^2}{2\bar{t}} \frac{dn(x)}{dx} \tag{4-20}$$

The quantity $\bar{l}^2/2\bar{t}$ is called the *electron diffusion coefficient*[7] D_n, with units cm^2/s. The minus sign in Eq. (4-20) arises from the definition of the derivative; it simply indicates that the net motion of electrons due to diffusion is in the direction of *decreasing* electron concentration. This is the result we expect, since net diffusion occurs from regions of high particle concentration to regions of low particle concentration. By identical arguments, we can show that holes in a hole concentration gradient move with a diffusion coefficient D_p. Thus

$$\phi_n(x) = -D_n \frac{dn(x)}{dx} \tag{4-21a}$$

$$\phi_p(x) = -D_p \frac{dp(x)}{dx} \tag{4-21b}$$

The diffusion current crossing a unit area (the current density) is the particle flux density multiplied by the charge of the carrier:

$$J_n(\text{diff.}) = -(-q)D_n \frac{dn(x)}{dx} = +qD_n \frac{dn(x)}{dx} \tag{4-22a}$$

$$J_p(\text{diff.}) = -(+q)D_p \frac{dp(x)}{dx} = -qD_p \frac{dp(x)}{dx} \tag{4-22b}$$

[7] If motion in three dimensions were included, the diffusion would be smaller in the x-direction. Actually, the diffusion coefficient should be calculated from the true energy distributions and scattering mechanisms. Diffusion coefficients are usually determined experimentally for a particular material, as described in Section 4.4.5.

CIRRUS EX. 1062 - 87/190

It is important to note that electrons and holes move together in a carrier gradient [Eqs. (4–21)], but the resulting currents are in opposite directions [Eqs. (4–22)] because of the opposite charge of electrons and holes.

4.4.2 Diffusion and Drift of Carriers; Built-in Fields

If an electric field is present in addition to the carrier gradient, the current densities will each have a drift component and a diffusion component

$$J_n(x) = q\mu_n n(x)\mathscr{E}(x) + qD_n \frac{dn(x)}{dx} \qquad (4\text{–}23a)$$

drift diffusion

$$J_p(x) = q\mu_p p(x)\mathscr{E}(x) - qD_p \frac{dp(x)}{dx} \qquad (4\text{–}23b)$$

and the total current density is the sum of the contributions due to electrons and holes:

$$J(x) = J_n(x) + J_p(x) \qquad (4\text{–}24)$$

We can best visualize the relation between the particle flow and the current of Eqs. (4–23) by considering a diagram such as shown in Fig. 4–14. In this figure an electric field is assumed to be in the x-direction, along with carrier distributions $n(x)$ and $p(x)$ which decrease with increasing x. Thus the derivatives in Eqs. (4–21) are negative, and diffusion takes place in the $+x$-direction. The resulting electron and hole diffusion currents [J_n (diff.) and J_p (diff.)] are in opposite directions, according to Eqs. (4–22). Holes drift in the direction of the electric field [ϕ_p (drift)], whereas electrons drift in the opposite direction because of their negative charge. The resulting drift current is in the $+x$-direction in each case. Note that the drift and diffusion components of the current are additive for holes when the field is in the direction of decreasing hole concentration, whereas the two components are subtractive for electrons under similar conditions. The total current may be due primarily to the flow of electrons or holes, depending on the relative concentrations and the relative magnitudes and directions of electric field and carrier gradients.

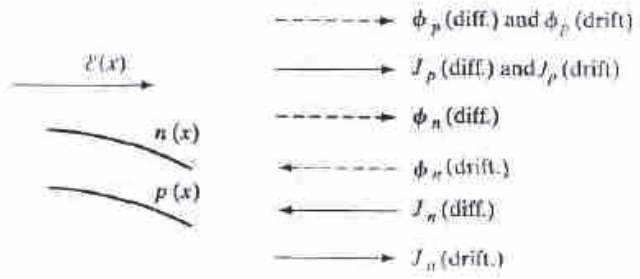


Figure 4–14
Drift and diffusion directions for electrons and holes in a carrier gradient and an electric field. Particle flow directions are indicated by dashed arrows, and the resulting currents are indicated by solid arrows.

CIRRUS EX. 1062 - 88/190

An important result of Eqs. (4–23) is that minority carriers can contribute significantly to the current through diffusion. Since the drift terms are proportional to carrier concentration, minority carriers seldom provide much drift current. On the other hand, diffusion current is proportional to the *gradient* of concentration. For example, in n-type material the minority hole concentration p may be many orders of magnitude smaller than the electron concentration n, but the gradient dp/dx may be significant. As a result, minority carrier currents through diffusion can sometimes be as large as majority carrier currents.

In discussing the motion of carriers in an electric field, we should indicate the influence of the field on the energies of electrons in the band diagrams. Assuming an electric field $\mathscr{E}(x)$ in the x-direction, we can draw the energy bands as in Fig. 4–15, to include the change in potential energy of electrons in the field. Since electrons drift in a direction opposite to the field, we expect the potential energy for electrons to increase in the direction of the field, as in Fig. 4–15. The electrostatic potential $\mathscr{V}(x)$ varies in the opposite direction, since it is defined in terms of positive charges and is therefore related to the electron potential energy $E(x)$ displayed in the figure by $\mathscr{V}(x) = E(x)/(-q)$.

From the definition of electric field,

$$\mathscr{E}(x) = -\frac{d\mathscr{V}(x)}{dx} \tag{4-25}$$

we can relate $\mathscr{E}(x)$ to the electron potential energy in the band diagram by choosing some reference in the band for the electrostatic potential. We are interested only in the spatial variation $\mathscr{V}(x)$ for Eq. (4–25). Choosing E_i as a convenient reference, we can relate the electric field to this reference by

$$\boxed{\mathscr{E}(x) = -\frac{d\mathscr{V}(x)}{dx} = -\frac{d}{dx}\left[\frac{E_i}{(-q)}\right] = \frac{1}{q}\frac{dE_i}{dx}} \tag{4-26}$$

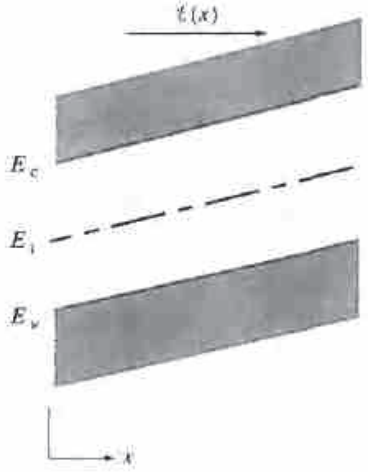


Figure 4–15
Energy band diagram of a semiconductor in an electric field $\mathscr{E}(x)$.

CIRRUS EX. 1062 - 89/190

Therefore, the variation of band energies with $\mathscr{E}(x)$ as drawn in Fig. 4–15 is correct. The direction of the slope in the bands relative to $\mathscr{E}$ is simple to remember: Since the diagram indicates electron energies, we know the slope in the bands must be such that electrons drift "downhill" in the field. Therefore, $\mathscr{E}$ points "uphill" in the band diagram.

At equilibrium, no net current flows in a semiconductor. Thus any fluctuation which would begin a diffusion current also sets up an electric field which redistributes carriers by drift. An examination of the requirements for equilibrium indicates that the diffusion coefficient and mobility must be related. Setting Eq. (4–23b) equal to zero for equilibrium, we have

$$\mathscr{E}(x) = \frac{D_p}{\mu_p} \frac{1}{p(x)} \frac{dp(x)}{dx} \tag{4-27}$$

Using Eq. (3–25b) for $p(x)$,

$$\mathscr{E}(x) = \frac{D_p}{\mu_p} \frac{1}{kT} \left(\frac{dE_i}{dx} - \frac{dE_F}{dx} \right) \tag{4-28}$$

The equilibrium Fermi level does not vary with x, and the derivative of E_i is given by Eq. (4–26). Thus Eq. (4–28) reduces to

$$\boxed{\frac{D}{\mu} = \frac{kT}{q}} \tag{4-29}$$

This result is obtained for either carrier type. This important equation is called the *Einstein relation*. It allows us to calculate either D or μ from a measurement of the other. Table 4–1 lists typical values of D and μ for several semiconductors at room temperature. It is clear from these values that $D/\mu = 0.026$ V.

An important result of the balance of drift and diffusion at equilibrium is that *built-in* fields accompany gradients in E_i [see Eq. (4–26)]. Such gradients in the bands at equilibrium (E_F constant) can arise when the band gap varies due to changes in alloy composition. More commonly, built-in fields result from doping gradients. For example, a donor distribution $N_d(x)$ causes a gradient in $n_0(x)$, which must be balanced by a built-in electric field $\mathscr{E}(x)$.

Table 4–1 Diffusion coefficient and mobility of electrons and holes for intrinsic semiconductors at 300 K. *Note:* Use Fig. 3–23 for doped semiconductors.

	D_n (cm²/s)	D_p (cm²/s)	μ_n (cm²/V-s)	μ_p (cm²/V-s)
Ge	100	50	3900	1900
Si	35	12.5	1350	480
GaAs	220	10	8500	400

CIRRUS EX. 1062 - 90/190

4.4.3 Diffusion and Recombination; The Continuity Equation

In the discussion of diffusion of excess carriers, we have thus far neglected the important effects of recombination. These effects must be included in a description of conduction processes, however, since recombination can cause a variation in the carrier distribution. For example, consider a differential length Δx of a semiconductor sample with area A in the yz-plane (Fig. 4–16). The hole current density leaving the volume, $J_p(x + \Delta x)$, can be larger or smaller than the current density entering, $J_p(x)$, depending on the generation and recombination of carriers taking place within the volume. The net increase in hole concentration per unit time, $\partial p/\partial t$, is the difference between the hole flux per unit volume entering and leaving, minus the recombination rate. We can convert hole current density to hole particle flux density by dividing J_p by q. The current densities are already expressed per unit area; thus dividing $J_p(x)/q$ by Δx gives the number of carriers per unit volume entering $\Delta x A$ per unit time, and $(1/q)J_p(x + \Delta x)/\Delta x$ is the number leaving per unit volume and time:

$$\left.\frac{\partial p}{\partial t}\right|_{x \to x + \Delta x} = \frac{1}{q}\frac{J_p(x) - J_p(x + \Delta x)}{\Delta x} - \frac{\delta p}{\tau_p} \tag{4-30}$$

$$\text{Rate of hole buildup} = \text{increase of hole concentration in } \delta x A \text{ per unit time} - \text{recombination rate}$$

As Δx approaches zero, we can write the current change in derivative form:

$$\frac{\partial p(x, t)}{\partial t} = \frac{\partial \delta p}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{\partial x} - \frac{\delta p}{\tau_p} \tag{4-31a}$$

The expression (4–31a) is called the *continuity equation* for holes. For electrons we can write

$$\frac{\partial \delta n}{\partial t} = \frac{1}{q}\frac{\partial J_n}{\partial x} - \frac{\delta n}{\tau_n} \tag{4-31b}$$

since the electronic charge is negative.

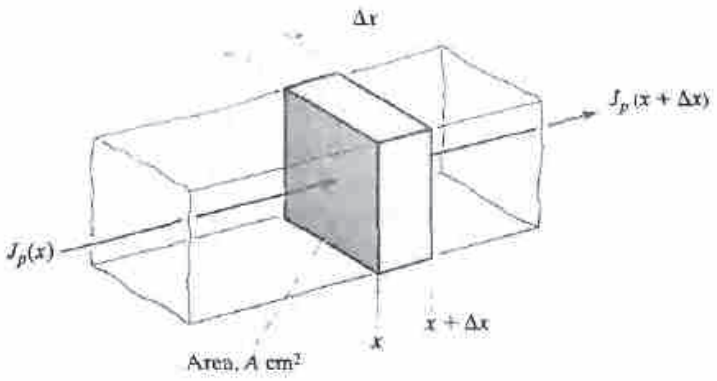


Figure 4–16
Current entering and leaving a volume $\Delta x A$.

CIRRUS EX. 1062 - 91/190

When the current is carried strictly by diffusion (negligible drift), we can replace the currents in Eqs. (4-31) by the expressions for diffusion current; for example, for electron diffusion we have

$$J_n(\text{diff.}) = qD_n \frac{\partial \delta n}{\partial x} \tag{4-32}$$

Substituting this into Eq. (4-31b) we obtain the *diffusion equation* for electrons,

$$\boxed{\frac{\partial \delta n}{\partial t} = D_n \frac{\partial^2 \delta n}{\partial x^2} - \frac{\delta n}{\tau_n}} \tag{4-33a}$$

and similarly for holes,

$$\boxed{\frac{\partial \delta p}{\partial t} = D_p \frac{\partial^2 \delta p}{\partial x^2} - \frac{\delta p}{\tau_n}} \tag{4-33b}$$

These equations are useful in solving transient problems of diffusion with recombination. For example, a pulse of electrons in a semiconductor (Fig. 4-12) spreads out by diffusion and disappears by recombination. To solve for the electron distribution in time, $n(x, t)$, we would begin with the diffusion equation, Eq. (4-33a).

4.4.4 Steady State Carrier Injection; Diffusion Length

In many problems a steady state distribution of excess carriers is maintained, such that the time derivatives in Eqs. (4-33) are zero. In the steady state case the diffusion equations become

$$\boxed{\frac{d^2 \delta n}{dx^2} = \frac{\delta n}{D_n \tau_n} \equiv \frac{\delta n}{L_n^2}} \tag{4-34a}$$

$$\boxed{\frac{d^2 \delta p}{dx^2} = \frac{\delta p}{D_p \tau_p} \equiv \frac{\delta p}{L_p^2}} \tag{4-34b}$$

(steady state)

where $L_n \equiv \sqrt{D_n \tau_n}$ is called the electron *diffusion length* and L_p is the diffusion length for holes. We no longer need partial derivatives, since the time variation is zero for steady state.

The physical significance of the diffusion length can be understood best by an example. Let us assume that excess holes are somehow injected into a semi-infinite semiconductor bar at $x = 0$, and the steady state hole injection maintains a constant excess hole concentration at the injection point $\delta p(x = 0) = \Delta p$. The injected holes diffuse along the bar, recombining with

CIRRUS EX. 1062 - 92/190

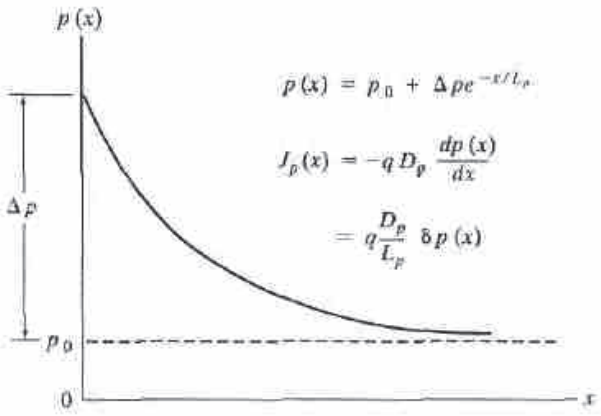


Figure 4–17
Injection of holes at $x = 0$, giving a steady state hole distribution $p(x)$ and a resulting diffusion current density $J_p(x)$.

a characteristic lifetime τ_p. In steady state we expect the distribution of excess holes to decay to zero for large values of x, because of the recombination (Fig. 4–17). For this problem we use the steady state diffusion equation for holes, Eq. (4–34b). The solution to this equation has the form

$$\delta p(x) = C_1 e^{x/L_p} + C_2 e^{-x/L_p} \tag{4–35}$$

We can evaluate C_1 and C_2 from the boundary conditions. Since recombination must reduce $\delta p(x)$ to zero for large values of x, $\delta p = 0$ at $x = \infty$ and therefore $C_1 = 0$. Similarly, the condition $\delta p = \Delta p$ at $x = 0$ gives $C_2 = \Delta p$, and the solution is

$$\boxed{\delta p(x) = \Delta p e^{-x/L_p}} \tag{4–36}$$

The injected excess hole concentration dies out exponentially in x due to recombination, and the diffusion length L_p represents the distance at which the excess hole distribution is reduced to $1/e$ of its value at the point of injection. We can show that *L_p is the average distance a hole diffuses before recombining*. To calculate an average diffusion length, we must obtain an expression for the probability that an injected hole recombines in a particular interval dx. The probability that a hole injected at $x = 0$ *survives* to x without recombination is $\delta p(x)/\Delta p = \exp(-x/L_p)$, the ratio of the steady state concentrations at x and 0. On the other hand, the probability that a hole at x will *recombine* in the subsequent interval dx is

$$\frac{\delta p(x) - \delta p(x + dx)}{\delta p(x)} = \frac{-(d\delta p(x)/dx)dx}{\delta p(x)} = \frac{1}{L_p} dx \tag{4–37}$$

Thus the total probability that a hole injected at $x = 0$ will recombine in a given dx is the product of the two probabilities:

$$(e^{-x/L_p})\left(\frac{1}{L_p} dx\right) = \frac{1}{L_p} e^{-x/L_p} dx \tag{4–38}$$

CIRRUS EX. 1062 - 93/190

Then, using the usual averaging techniques described by Eq. (2-21), the average distance a hole diffuses before recombining is

$$\langle x \rangle = \int_0^\infty x \frac{e^{-x/L_p}}{L_p} dx = L_p \tag{4-39}$$

The steady state distribution of excess holes causes diffusion, and therefore a hole current, in the direction of decreasing concentration. From Eqs. (4-22b) and (4-36) we have

$$J_p(x) = -qD_p \frac{dp}{dx} = -qD_p \frac{d\delta p}{dx} = q \frac{D_p}{L_p} \Delta p e^{-x/L_p} = q \frac{D_p}{L_p} \delta p(x) \tag{4-40}$$

Since $p(x) = p_0 + \delta p(x)$, the space derivative involves only the excess concentration. We notice that since $\delta p(x)$ is proportional to its derivative for an exponential distribution, the diffusion current at any x is just proportional to the excess concentration δp at that position.

Although this example seems rather restricted, its usefulness will become apparent in Chapter 5 in the discussion of p-n junctions. The injection of minority carriers across a junction often leads to exponential distributions as in Eq. (4-36), with the resulting diffusion current of Eq. (4-40).

EXAMPLE 4-5 In a very long p-type Si bar with cross-sectional area = 0.5 cm^2 and $N_a = 10^{17}$ cm^{-3}, we inject holes such that the steady state *excess* hole concentration is 5×10^{16} cm^{-3} at $x = 0$. What is the steady state separation between F_p and E_c at $x = 1000$ Å? What is the hole current there? How much is the excess stored hole charge? Assume that $\mu_p = 500$ cm^2/V-s and $\tau_p = 10^{-10}$ s.

SOLUTION

$$D_p = \frac{kT}{q} \mu_p = 0.0259 \times 500 = 12.95 \text{ cm/s}$$

$$L_p = \sqrt{D_p \tau_p} = \sqrt{12.95 \times 10^{-10}} = 3.6 \times 10^{-5} \text{ cm}$$

$$p = p_0 + \Delta p e^{-\frac{x}{L_p}} = 10^{17} + 5 \times 10^{16} e^{-\frac{10^{-5}}{3.6 \times 10^{-5}}}$$

$$= 1.379 \times 10^{17} = n_i e^{(E_i - F_p)/kT} = (1.5 \times 10^{10} \text{ cm}^{-3}) e^{(E_i - F_p)/kT}$$

$$E_i - F_p = \left(\ln \frac{1.379 \times 10^{17}}{1.5 \times 10^{10}} \right) \cdot 0.0259 = 0.415 \text{ eV}$$

$$E_c - F_p = 1.1/2 \text{ eV} + 0.415 \text{ eV} = \mathbf{0.965 \text{ eV}}$$

CIRRUS EX. 1062 - 94/190

We can calculate the hole current from Eq. (4–40)

$$I_p = -qAD_p\frac{dp}{dx} = qA\frac{D_p}{L_p}(\Delta p)e^{-\frac{x}{L_p}}$$

$$= 1.6 \times 10^{-19} \times 0.5 \times \frac{12.95}{3.6 \times 10^{-5}} \times 5 \times 10^{16}\, e^{-\frac{10^{-3}}{3.6\times10^{-5}}}$$

$$= \mathbf{1.09 \times 10^{3}\ A}$$

$$Q_p = qA(\Delta p)L_p$$

$$= 1.6 \times 10^{-19}(0.5)(5 \times 10^{16})(3.6 \times 10^{-5})$$

$$= \mathbf{1.44 \times 10^{-7}\ C}$$

4.4.5 The Haynes–Shockley Experiment

One of the classic semiconductor experiments is the demonstration of drift and diffusion of minority carriers, first performed by J. R. Haynes and W. Shockley in 1951 at the Bell Telephone Laboratories. The experiment allows independent measurement of the minority carrier mobility μ and diffusion coefficient D. The basic principles of the Haynes–Shockley experiment are as follows: A pulse of holes is created in an n-type bar (for example) that contains an electric field (Fig. 4–18); as the pulse drifts in the field and spreads out by diffusion, the excess hole concentration is monitored at some point down the bar; the time required for the holes to drift a given distance in the field gives a measure of the mobility; and the spreading of the pulse during a given time is used to calculate the diffusion coefficient.

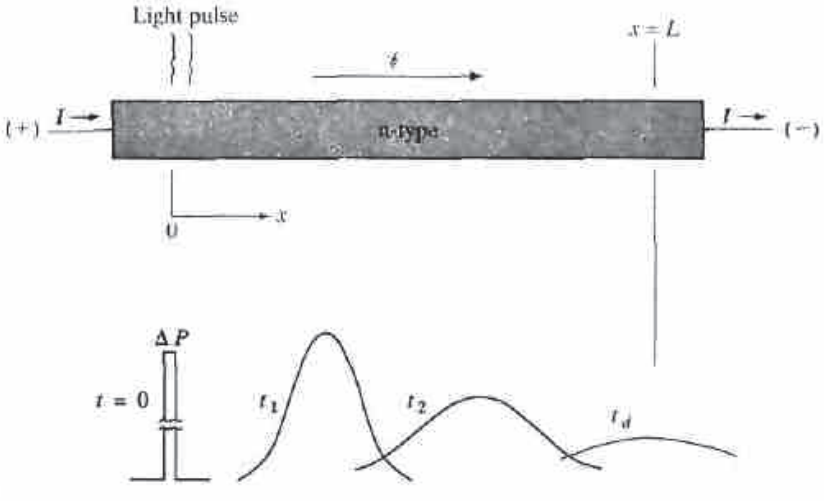


Figure 4–18 Drift and diffusion of a hole pulse in an n-type bar: (a) sample geometry; (b) position and shape of the pulse for several times during its drift down the bar.

CIRRUS EX. 1062 - 95/190

In Fig. 4–18 a pulse of excess carriers is created by a light flash at some point $x = 0$ in an n-type semiconductor ($n_0 \gg p_0$). We assume that the excess carriers have a negligible effect on the electron concentration but change the hole concentration significantly. The excess holes drift in the direction of the electric field and eventually reach the point $x = L$, where they are monitored. By measuring the drift time t_d, we can calculate the drift velocity v_d and, therefore, the hole mobility:

$$v_d = \frac{L}{t_d} \tag{4-41}$$

$$\mu_p = \frac{v_d}{\mathscr{E}} \tag{4-42}$$

Thus the hole mobility can be calculated directly from a measurement of the drift time for the pulse as it moves down the bar. In contrast with the Hall effect (Section 3.4.5), which can be used with resistivity to obtain the *majority* carrier mobility, the Haynes–Shockley experiment is used to measure the *minority* carrier mobility.

As the pulse drifts in the $\mathscr{E}$ field it also spreads out by diffusion. By measuring the spread in the pulse, we can calculate D_p. To predict the distribution of holes in the pulse as a function of time, let us first reexamine the case of diffusion of a pulse *without drift, neglecting recombination* (Fig. 4–12). The equation which the hole distribution must satisfy is the time-dependent diffusion equation, Eq. (4–33b). For the case of negligible recombination (τ_p long compared with the times involved in the diffusion), we can write the diffusion equation as

$$\frac{\partial \delta p(x, t)}{\partial t} = D_p \frac{\partial^2 \delta p(x, t)}{\partial x^2} \tag{4-43}$$

The function which satisfies this equation is called a *gaussian distribution,*

$$\delta p(x, t) = \left[\frac{\Delta P}{2\sqrt{\pi D_p t}}\right] e^{-x^2/4D_p t} \tag{4-44}$$

where ΔP is the number of holes per unit area created over a negligibly small distance at $t = 0$. The factor in brackets indicates that the peak value of the pulse (at $x = 0$) decreases with time, and the exponential factor predicts the spread of the pulse in the positive and negative x-directions (Fig. 4–19). If we designate the peak value of the pulse as $\delta\hat{p}$ at any time (say t_d), we can use Eq. (4–44) to calculate D_p from the value of δp at some point x. The most convenient choice is the point $\Delta x/2$, at which δp is down by $1/e$ of its peak value $\delta\hat{p}$. At this point we can write

$$e^{-1}\delta\hat{p} = \delta\hat{p}e^{-(\Delta x/2)^2/4D_p t_d} \tag{4-45}$$

$$D_p = \frac{(\Delta x)^2}{16 t_d} \tag{4-46}$$

CIRRUS EX. 1062 - 96/190

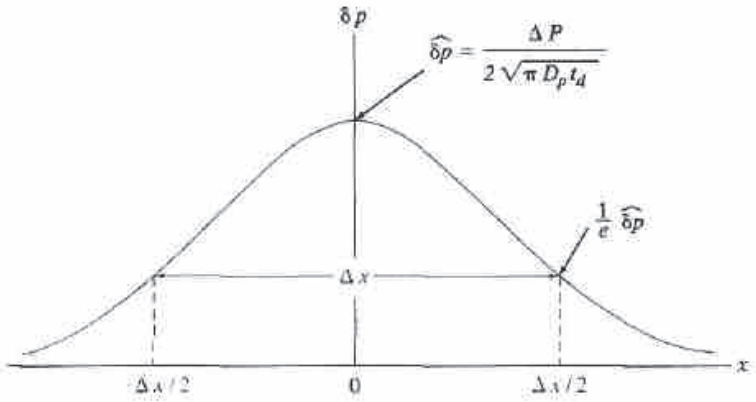


Figure 4–19
Calculation of D_p from the shape of the δp distribution after time t_d. No drift or recombination is included.

Since Δx cannot be measured directly, we use an experimental setup such as Fig. 4–20, which allows us to display the pulse on an oscilloscope as the carriers pass under a detector. As we shall see in Chapter 5, a forward-biased p-n junction serves as an excellent injector of minority carriers, and a reverse-biased junction serves as a detector. The measured quantity in Fig. 4–20 is the pulse width Δt displayed on the oscilloscope in time. It is related to Δx by the drift velocity, as the pulse drifts past the detector point (2)

$$\Delta x = \Delta t v_d = \Delta t \frac{L}{t_d} \tag{4-47}$$

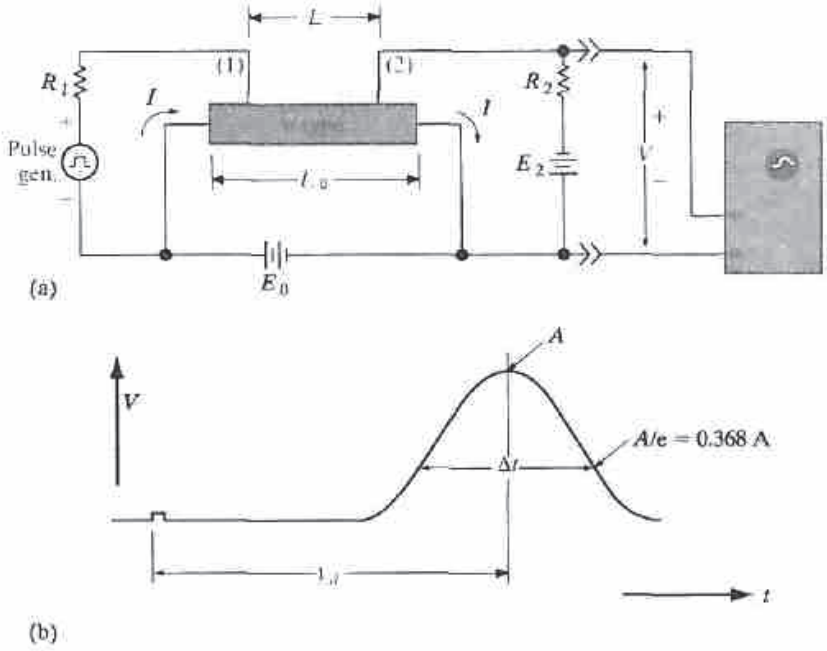


Figure 4–20
The Haynes–Shockley experiment: (a) circuit schematic; (b) typical trace on the oscilloscope screen.

CIRRUS EX. 1062 - 97/190

EXAMPLE 4–6

An n-type Ge sample is used in the Haynes-Shockley experiment shown in Fig. 4–20. The length of the sample is 1 cm, and the probes (1) and (2) are separated by 0.95 cm. The battery voltage E_0 is 2 V. A pulse arrives at point (2) 0.25 ms after injection at (1); the width of the pulse Δt is 117 μs. Calculate the hole mobility and diffusion coefficient, and check the results against the Einstein relation.

SOLUTION

$$\mu_p = \frac{v_d}{\mathscr{E}} = \frac{0.95/(0.25 \times 10^{-3})}{2/1} = 1900 \text{ cm}^2/(\text{V-s})$$

$$D_p = \frac{(\Delta x)^2}{16 t_d} = \frac{(\Delta t L)^2}{16 t_d^3}$$

$$= \frac{(117 \times 0.95)^2 \times 10^{-12}}{16(0.25)^3 \times 10^{-9}} = 49.4 \text{ cm}^2/\text{s}$$

$$\frac{D_p}{\mu_p} = \frac{49.4}{1900} = 0.026 = \frac{kT}{q}$$

4.4.6 Gradients in the Quasi-Fermi Levels

In Section 3.5 we saw that equilibrium implies no gradient in the Fermi level E_F. In contrast, any combination of drift and diffusion implies a gradient in the steady state quasi-Fermi level.

We can use the results of Eqs. (4–23), (4–26), and (4–29) to demonstrate the power of the concept of quasi-Fermi levels in semiconductors [see Eq. (4–15)]. If we take the general case of nonequilibrium electron concentration with drift and diffusion, we must write the total electron current as

$$J_n(x) = q\mu_n n(x)\mathscr{E}(x) + qD_n \frac{dn(x)}{dx} \tag{4–48}$$

where the gradient in electron concentration is

$$\frac{dn(x)}{dx} = \frac{d}{dx}\left[n_i e^{(F_n - E_i)/kT}\right] = \frac{n(x)}{kT}\left(\frac{dF_n}{dx} - \frac{dE_i}{dx}\right) \tag{4–49}$$

Using the Einstein relation, the total electron current becomes

$$J_n(x) = q\mu_n n(x)\mathscr{E}(x) + \mu_n n(x)\left[\frac{dF_n}{dx} - \frac{dE_i}{dx}\right] \tag{4–50}$$

But Eq. (4–26) indicates that the subtractive term in the brackets is just $q\mathscr{E}(x)$, giving a direct cancellation of $q\mu_n n(x)\mathscr{E}(x)$ and leaving

$$J_n(x) = \mu_n n(x)\frac{dF_n}{dx} \tag{4–51}$$

CIRRUS EX. 1062 - 98/190

Thus, the processes of electron drift and diffusion are summed up by the spatial variation of the quasi-Fermi level. The same derivation can be made for holes, and we can write the current due to drift and diffusion in the form of a *modified Ohm's law*

$$J_n(x) = q\mu_n n(x)\frac{d(F_n/q)}{dx} = \sigma_n(x)\frac{d(F_n/q)}{dx} \tag{4-52a}$$

$$J_p(x) = q\mu_p p(x)\frac{d(F_p/q)}{dx} = \sigma_p(x)\frac{d(F_p/q)}{dx} \tag{4-52b}$$

Therefore, any drift, diffusion, or combination of the two in a semiconductor results in currents proportional to the gradients of the two quasi-Fermi levels. Conversely, a lack of current implies constant quasi-Fermi levels. One can use a hydrostatic analogy for quasi-Fermi levels and identify it as water pressure in a system. Just as water flows from a high-pressure region to a low-pressure region, until in equilibrium the water pressure is the same everywhere, similarly electrons flow from a high- to low-electron quasi-Fermi level region, until we get a flat Fermi level in equilibrium. Quasi-Fermi levels are sometimes also known as electrochemical potentials because, as we just saw, the driving force for carriers is governed partly by gradients of electrical potential (or electric field), which determines drift, and partly by gradients of carrier concentration (which is related to a thermodynamic concept called *chemical potential*), giving rise to diffusion.

SUMMARY

4.1 *Excess* carriers, above the equilibrium values contributed by doping, may be created *optically* (or by electrical *biasing* in devices). *Generation–recombination* (G–R) of electron–hole pairs (*EHPs*) can occur by absorption of the photons with energy greater than the band gap, balanced by direct or indirect recombination.

4.2 G–R processes can be mediated by *traps*, especially deep traps near midgap. Band-to-band or trap-assisted G–R processes lead to an average *lifetime* for the excess carriers. Carrier lifetime multiplied by the optical generation rate establishes a steady state excess population of carriers. The square root of carrier lifetime multiplied by the diffusion coefficient determines the diffusion length.

4.3 In *equilibrium*, we have a *constant Fermi level*. In nonequilibrium with excess carriers, Fermi levels are generalized to separate *quasi-Fermi levels* for electrons and holes. The quasi-Fermi level splitting is a measure of the departure from equilibrium. Minority carrier quasi-Fermi levels change more than majority carrier quasi-Fermi levels because the relative change of minority carriers is larger. Gradients in the quasi-Fermi level determine the net drift–diffusion current.

CIRRUS EX. 1062 - 99/190

4.4 *Diffusion flux* measures the flow of carriers from *high- to low-concentration* regions and is given by the *diffusivity* times the concentration *gradient*. The direction of diffusion current is opposite to the flux for the negative electrons, but in the same direction for the positive holes. Carrier diffusivity is related to mobility by the thermal voltage kT/q (*Einstein relation*).

4.5 When carriers move in a semiconductor due to drift or diffusion, the time-dependent carrier concentrations at different points is given by the *carrier continuity* equation, which says that if more carriers flow into a point than flow out, the concentration will increase as a function of time and vice versa. G–R processes also affect carrier concentrations.

PROBLEMS

4.1 With E_F located 0.4 eV above the valence band in a Si sample, what charge state would you expect for most Ga atoms in the sample? What would be the predominant charge state of Zn? Au? *Note:* By charge state we mean neutral, singly positive, doubly negative, etc.

4.2 A Si sample with $10^{16}/cm^3$ donors is optically excited such that $10^{19}/cm^3$ electron–hole pairs are generated per second uniformly in the sample. The laser causes the sample to heat up to 450 K. Find the quasi-Fermi levels and the change in conductivity of the sample upon shining the light. Electron and hole lifetimes are both 10 μs. $D_p = 12\ cm^2/s$; $D_n = 36\ cm^2/s$; $n_i = 10^{14}\ cm^{-3}$ at 450 K. What is the change in conductivity upon shining light?

4.3 Construct a semilogarithmic plot such as Fig. 4–7 for Si doped with 2×10^{15} donors/cm^3 and having 4×10^{14} EHP/cm^3 created uniformly at $t = 0$. Assume that $\tau_n = \tau_p = 5$ μs.

4.4 Calculate the recombination coefficient α_r for the low-level excitation described in Prob. 4.3. Assume that this value of α_r applies when the GaAs sample is uniformly exposed to a steady state optical generation rate $g_{op} = 10^{19}$ EHP/cm^3-s. Find the steady state excess carrier concentration $\Delta n = \Delta p$.

4.5 An intrinsic Si sample is doped with donors from one side such that $N_d = N_0 \exp(-ax)$. (a) Find an expression for the built-in electric field at equilibrium over the range for which $N_d \gg n_i$. (b) Evaluate the field when $a = 1\ (\mu m)^{-1}$. (c) Sketch a band diagram such as in Fig. 4–15 and indicate the direction of the field.

4.6 A Si sample with $10^{15}/cm^3$ donors is uniformly optically excited at room temperature such that $10^{19}/cm^3$ electron–hole pairs are generated per second. Find the separation of the quasi-Fermi levels and the change of conductivity upon shining the light. Electron and hole lifetimes are both 10 μs. $D_p = 12\ cm^2/s$.

4.7 An n-type Si sample with $N_d = 10^{15}\ cm^{-3}$ is steadily illuminated such that $g_{op} = 10^{21}$ EHP/cm^3-s. If $\tau_n = \tau_p = 1\mu s$ for this excitation, calculate the separation in the quasi-Fermi levels, $(F_n - F_p)$. Draw a band diagram such as Fig. 4–11.

CIRRUS EX. 1062 - 100/190

4.8 For a 2-cm-long doped Si bar ($N_d = 10^{16}$ cm^{-3}) with a cross-sectional area = 0.05 cm^2, what is the current if we apply 10V across it? If we generate 10^{20} electron-hole pairs per second per cm^3 uniformly in the bar and the lifetime $\tau_n = \tau_p = 10^{-4}$ s, what is the new current? Assume the low-level α, doesn't change for high-level injection. If the voltage is then increased to 100,000 V, what is the new current? Assume $\mu_p = 500$ cm^2/V-s, but you must choose the appropriate value for electrons.

4.9 Design and sketch a photoconductor using a 5-μm-thick film of CdS, assuming that $\tau_n = \tau_p = 10^{-6}$ s and $N_d = 10^{14}$ cm^{-3}. The dark resistance (with $g_{op} = 0$) should be 10 MΩ, and the device must fit in a square 0.5 cm on a side; therefore, some sort of folded or zigzag pattern is in order. With an excitation of $g_{op} = 10^{21}$ EHP/cm^3-s, what is the resistance change?

4.10 A 100-mW laser beam with wavelength $\lambda = 6328$ Å is focused onto a GaAs sample 100 μm thick. The absorption coefficient at this wavelength is 3×10^4 cm^{-1}. Find the number of photons emitted per second by radiative recombination in the GaAs, assuming perfect quantum efficiency. What power is delivered to the sample as heat?

4.11 Assume that a photoconductor in the shape of a bar of length L and area A has a constant voltage V applied, and it is illuminated such that g_{op} EHP/cm^3-s are generated uniformly throughout. If $\mu_n \gg \mu_p$, we can assume the optically induced change in current ΔI is dominated by the mobility μ_n and lifetime τ_n for electrons. Show that $\Delta I = qAL g_{op} \tau_n / \tau_t$ for this photoconductor, where τ_t is the transit time of electrons drifting down the length of the bar.

4.12 For the steady state minority hole distribution shown in Fig. 4–17, find the expression for the hole quasi-Fermi level position $E_i - F_p(x)$ while $p(x) \gg p_0$ (i.e., while F_p is below E_F). On a band diagram, draw the variation of $F_p(x)$. Be careful—when the minority carriers are few (e.g., when δp is n_i), F_p still has a long way to go to reach E_F.

4.13 In an n-type semiconductor bar, there is an increase in electron concentration from left to right and an electric field pointing to the left. With a suitable sketch, indicate the directions of the electron drift and diffusion current flow and explain why. If we double the electron concentration everywhere, what happens to the diffusion current and the drift current? If we add a constant concentration of electrons everywhere, what happens to the drift and diffusion currents? Explain your answers with appropriate equations.

4.14 The current required to feed the hole injection at $x = 0$ in Fig. 4–17 is obtained by evaluating Eq. (4–40) at $x = 0$. The result is $I_p(x = 0) = qAD_p \Delta p / L_p$. Show that this current can be calculated by integrating the charge stored in the steady state hole distribution $\delta p(x)$ and then dividing by the average hole lifetime τ_p. Explain why this approach gives $I_p(x = 0)$.

4.15 The direction of the built-in electric field can be deduced without math by sketching the result of a doping gradient on the band diagram. Starting with a flat Fermi level at equilibrium, place E_i near or far from E_F as the doping is

CIRRUS EX. 1062 - 101/190

varied for the two cases of a gradient in donor or acceptor doping as in Prob. 4.5. Show the electric field direction in each case, based on Eq. (4–26). If a *minority* carrier is injected into the impurity gradient region, in what direction is it accelerated in the two cases? This is an interesting effect that we will use later in discussing bipolar transistors.

4.16 In Prob. 4.5, the direction of the built-in electric field due to a gradient in doping was determined from Eqs. (4–23) and (4–26). In this problem, you are asked to explain qualitatively why the field must arise and find its direction. (a) Sketch a donor doping distribution as in Prob. 4.5, and explain the field required to keep the mobile electrons from diffusing down the gradient. Repeat for acceptors and holes. (b) Sketch a microscopic region of the doping distribution, showing ionized donors and the resulting mobile electrons. Explain the origin and direction of the field as the electrons attempt to diffuse toward lower concentrations. Repeat for acceptors and holes.

4.17 We wish to use the Haynes–Shockley experiment to calculate the hole lifetime τ_p in an n-type sample. Assume that the peak voltage of the pulse displayed on the oscilloscope screen is proportional to the hole concentration under the collector terminal at time t_d and that the displayed pulse can be approximated as a gaussian, as in Eq. (4–44), which decays due to recombination by e^{-t/τ_p}. The electric field is varied and the following data are taken: For $t_d = 200$ μs, the peak is 20 mV; for $t_d = 50$ μs, the peak is 80 mV. What is τ_p?

READING LIST

Ashcroft, N. W., and N. D. Mermin. *Solid State Physics*. Philadelphia: W.B. Saunders, 1976.

Bhattacharya, P. *Semiconductor Optoelectronic Devices*. Englewood Cliffs, NJ: Prentice Hall, 1994.

Blakemore, J. S. *Semiconductor Statistics*. New York: Dover Publications, 1987.

Neamen, D. A. *Semiconductor Physics and Devices: Basic Principles*. Homewood, IL: Irwin, 2003.

Pankove, J. I. *Optical Processes in Semiconductors*. Englewood Cliffs, NJ: Prentice Hall, 1971.

Pierret, R. F. *Semiconductor Device Fundamentals*. Reading, MA: Addison-Wesley, 1996.

Singh, J. *Semiconductor Devices*. New York: McGraw-Hill, 1994.

Wolfe, C. M., G. E. Stillman, and N. Holonyak, Jr. *Physical Properties of Semiconductors*. Englewood Cliffs, NJ: Prentice Hall, 1989.

SELF QUIZ

Question 1

Consider a p-type semiconductor that has a band gap of 1.0 eV and a minority electron lifetime of 0.1 μs, and is uniformly illuminated by light having photon energy of 2.0 eV.

(a) What rate of uniform excess carrier generation is required to generate a uniform electron concentration of $10^{10}/cm^3$?

CIRRUS EX. 1062 - 102/190

(b) How much optical power per cm^3 must be absorbed in order to create the excess carrier population of part (a)? (You may leave your answer in units of $eV/s\text{-}cm^3$.)

(c) If the carriers recombine via photon emission, approximately how much optical power per cm^3 will be generated? (You may leave your answer in units of $eV/s\text{-}cm^3$.)

Question 2

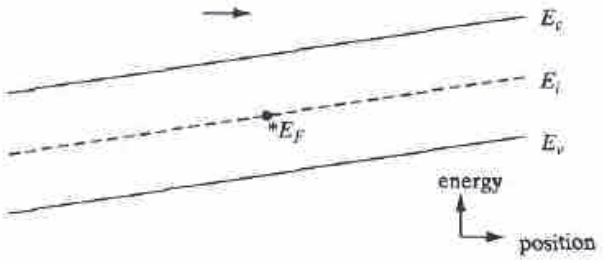
(a) What do we mean by "deep" versus "shallow" traps? Which are more harmful for semiconductor devices and why? What is an example of a deep trap in Si?

(b) Are absorption lengths of slightly above band gap photons longer in Si or GaAs? Why?

(c) Do absorption coefficients of photons increase or decrease with photon energy? Why?

Question 3

Consider the following equilibrium band diagram for a portion of a semiconductor sample with a built-in electric field $\mathscr{E}$:

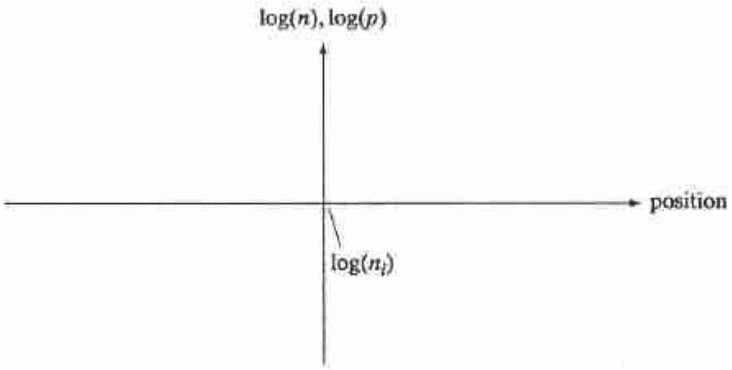


(a) Sketch the Fermi level as a function of position *through* the indicated point, E_F, across the width of the band diagram above.

(b) On the band diagram, sketch the direction of the electric field. Is the field constant or position dependent?

(c) On the following graph, sketch and label both the electron and hole concentrations as a function of position across the full width of the sample. Note that the carrier concentration scale is logarithmic such that exponential variations in the carrier concentration with position appear as straight lines. Note also that the horizontal axis corresponds to the intrinsic carrier concentration of n_i.

CIRRUS EX. 1062 - 103/190



Question 4

(a) Indicate the directions of the hole and electron flux densities ϕ due to diffusion and drift under these *equilibrium* conditions corresponding to the previous Question 3.

(b) Indicate the directions of the hole and electron current densities j due to diffusion and drift under these *equilibrium* conditions.

Question 5

(a) What are the relevant equations that must be solved in general for a semiconductor device problem?

(b) In general how many components of conduction current can you have in a semiconductor device? What are they?

Question 6

(a) Consider a region in a semiconductor with an electric field directed toward the right ($\rightarrow$) and carrier concentrations increasing toward the left ($\leftarrow$). Indicate the directions of particle fluxes ϕ (circle one for each) and charge currents j due to drift and diffusion within that region.

(b) Based on your answers to part (a), indicate the directions of the charge currents j due to drift and diffusion within that region.

CIRRUS EX. 1062 - 104/190