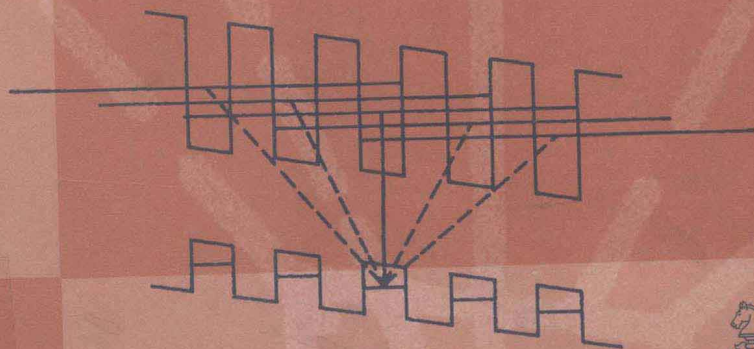# K. Seeger

Semiconductor Physics

An Introduction

8th Edition

Springer

Karlheinz Seeger

Semiconductor Physics

An Introduction

Eighth Edition
With 326 Figures

Professor Karlheinz Seeger

Am Modenapark 13/5
1030 Vienna, Austria
and
Institut für Materialphysik der Universität
Boltzmanngasse 5
1090 Vienna, Austria

Preface

This book, now in its eighth edition, still has the character of a textbook with the emphasis on "Physics". I have refrained from sacrificing topics such as the Seebeck, Thomson, Peltier and quantum Hall effects although they do not have a large-scale practical application. The various kinds of field effect transistors such as HBT, HEMT, MODFET and the chip technology have not been considered because going into details would necessarily end up with referring the interested reader to commercially available computer programs, i.e. to the Big Business world.

What is new compared with the previous edition? In Chap. 1 you will now find a short description of production processes of pure single crystal silicon which is nearly free of lattice defects. In Sect. 4.14 the calculation of current in a sample subject to combined dc and high–frequency ac fields is based on quantum mechanics rather than classical physics (called "photon assisted tunneling"). The calculation is applied to charge transport in quantum wires and dots in Chap.14. This chapter includes production methods for quantum dots in their regular arrangement and a dot laser which operates at room temperature. In Chap. 15 there is a new section on doped Fullerenes in view of a field effect transistor based on this material. There are two more appendices: One provides a calculation of the potentials in quantum wells, and the other is a table of 38 semiconductors and semimetals with their characteristic data.

The system of units preferred here is the SI system. If you want to convert e.g. the magnetic field B to its Gaussian unit you have to divide it by the velocity of light c. Vectors are typed boldface. A vector product is put in [brackets], a scalar product in (parenthesis). Notation of the form (5.7.7, 8) is used to refer to both equation (5.7.7) and equation (5.7.8) while [5.7] refers to reference 5.7.

A basic prerequisite for the existence of this book is (and always was) the excellent cooperation with Hofrat Dr.Wolfgang Kerber and his crew at the library of the Institute of Physics, University of Vienna, Austria. I am grateful to Prof. Dr. H. Stetter, Technical University of Vienna, for providing help with installing LATEX on my computer and for introducing me to this method of doing the type - setting work,(today's general request by publishers), and to Prof. Dr. H. Kühnelt for providing very helpful computer programs. Dr. Hubertus von Riedesel, managing editor at Springer - Verlag, Heidelberg, and his assistant Dr. Angela Lahee deserve my heartfelt thanks for their cooperation. Last but not least my wife suffered from my occasional mental absences during this period of both writing and taming an unwilling computer, with remarkable good humour and I would like to express my appreciation for her support.

Vienna, 2002   K.S.

Contents

1. Elementary Properties of Semiconductors

1.1 Insulator – Semiconductor – Semimetal – Metal

A consequence of the discovery of electricity was the observation that metals are good conductors while nonmetals are poor conductors. The latter were called insulators. Metallic conductivity is typically between 10^6 and 10^4 ($\Omega\,\mathrm{cm})^{-1}$, while typical insulators have conductivities of less than 10^{-10} ($\Omega\,\mathrm{cm})^{-1}$. Some solids with conductivities between 10^4 and 10^{-10} ($\Omega\,\mathrm{cm})^{-1}$ are classified as *semiconductors*. However, substances such as alkali-halides whose conductivity is due to electrolytic decomposition shall be excluded. Also we restrict our discussion to chemically uniform, *homogeneous* substances and prefer those which can be obtained in monocrystalline form. Even then we have to distinguish between semiconductors and semimetals. This distinction is possible only as a result of thorough investigation of optical and electrical properties and how they are influenced by temperature, magnetic field, etc. Without giving further explanations at this stage, the statement is made that semiconductors have an *energy gap* while semimetals and metals have no such gap. However, very impure semiconductors show a more or less metallic behavior and with many substances, the art of purification is not so far advanced that a distinction can easily be made. The transition between semiconductors and insulators is even more gradual and depends on the ratio of the energy gap to the temperature of investigation. Very pure semiconductors may become insulators when the temperature approaches the absolute zero.

Typical elemental semiconductors are germanium and silicon. An inspection of the periodic table of elements reveals that these materials belong to the fourth group while typical metals such as the alkalis are in the first group and typical nonmetals such as the halogens and the noble gases which crystallize at low temperatures are in the seventh and eighth group, respectively. Other semiconducting elements in the fourth group are diamond which is a modification of carbon, while *gray* tin ($\alpha - \mathrm{Sn}$), which is stable only at low temperatures is a semimetal. All fourth-group semiconductors crystallize in a structure known as the diamond structure in which neighboring atoms are arranged in tetrahedral symmetry. In the third group, the lightest element boron, and in the sixth group, the heavy elements selenium and tellurium, are semiconductors. A typical semimetal is the heaviest fifth group element, bismuth, and also the lighter elements of this group, arsenic and antimony, may be classified as such although they are at present less thoroughly investigated.

Typical compound semiconductors are the III–V compounds such as gallium arsenide, GaAs, and indium antimonide, InSb, and the II–VI compounds such as zinc sulfide, ZnS (*zinc blende*). They crystallize in the zinc blende structure which can be obtained from the diamond structure by replacing the carbon atoms alternately by, e.g., zinc and sulfur atoms. These compounds have a stoichiometric composition, just as, e.g., the semiconductor silicon carbide, SiC, while germanium silicon alloys may be obtained as semiconducting mixed crystals for any arbitrary composition. Many metal oxides and sulfides are semiconductors, often with nonstoichiometric composition. Some of them are of technical importance, such as cuprous oxide, Cu_2O (formerly used for rectifiers), lead sulfide, PbS (for infrared detectors) and the ferrites (iron oxides) for their magnetic properties. Today silicon is mainly used for the fabrication of transistors which serve for amplification of electric signals. This is the most important technical application of semiconductors nowadays.

Semiconduction is specified by the following properties:

(a) In a *pure* semiconductor, conductivity rises exponentially with temperature (*thermistor* action). At lower temperatures, smaller concentration of impurities is required in order to ensure this behavior.

(b) In an impure semiconductor, the conductivity depends strongly on the impurity concentration. For example, nickel oxide NiO in a pure condition is an insulator. By doping (which means intentionally adding impurities) with 1% lithium oxide, Li_2O , the conductivity is raised by a factor of 10^{13} . In the heavily doped material, however, the conductivity changes only slightly with temperature, just as in a metal.

(c) The conductivity is changed (in general, raised) by irradiation with light or high-energy electrons or by the *injection* of carriers from a suitable metallic contact (injection will be explained in Sect. 5.1).

(d) Depending on the kind of doping, the charge transport may be either by electrons or by so-called *positive holes*. The electric behavior of positive holes is the same as that of positrons but otherwise there is no similarity. It is possible to dope a single crystal nonuniformly such that in some parts, charge transport is by (negative) electrons and at the same time in others by positive holes. Semiconductor diodes and transistors are single crystals of that kind.

Semiconducting behavior is not restricted to solids. There are liquid semiconductors. However, because of atomic diffusion, regions with different dopings will mix rapidly and a stable device with an inhomogeneous structure is not possible. Recently, attention has been paid to glassy and amorphous semiconductors which may possibly find a technical application for solar cells.

As mentioned before, semiconductors become metallic when heavily doped. Superconductivity, known for some metals and inorganic as well as organic compounds at low temperatures, has also been observed with some heavily doped semiconductors. Transition temperatures are below 150 K. Some aromatic hydrocarbons and even fullerenes, which represent besides coal and diamond a third modification of carbon, show semiconducting behavior They will be reported in Sect. 15.4. [1.1].

Growth of single crystals of silicon in a cylindrical form of over 20 cm in diameter and a meter length with dislocation densities of less than 10^3 cm^{-3}(W.C. Dash[1.2]) and impurity concentrations of less than one part per trillion (10^{12}) is today's industrial standard. How are these crystals produced? In the Czochralski technique [1.3] a small seed crystal is put in contact with the molten originally polycrystalline material (melting point 1415°C) and then pulled up at a speed of a few millimeters per minute. Slow rotation of the seed crystal ensures that the resultant crystal is cylindrical. The crystallographic atomic structure is, of course, the one of the seed crystal. The original material found in nature is quartz sand (chemially contaminated silicon dioxide SiO_2) which is heated with charcoal in order to remove the oxygen. Further processes of getting rid of impurities involve the conversion of the raw silicon to trichlorosilane $SiHCl_3$, a liquid from which by fractional distillation a boron impurity is removed which otherwise in the final product would produce an unwanted metallic electrical conduction. Another semiconductor of technical importance is gallium arsenide GaAs (melting point 1238°C) which in liquid form without protection loses arsenic in the form of vapor in the Czochralski process. This decomposition of the compound is avoided by putting a layer of molten boron oxide on top of the gallium arsenide. The technique is known as *Liquid Encapsulation Czochralski method*, LEC. In the *Bridgman* variation of the Czochralski technique, a temperature gradient along the length of the crucible is produced by heaters. By gradually reducing the energy supplied to the heaters, the crystal grows beginning at the seed crystal.

Floating zone purification of a cylindrically shaped crystal with its axis arranged vertically implies heating part of the cylinder by *high-frequency inductive heating* to the melting temperature and gradually moving the heaters along the crystal axis. The melt is kept in place by surface tension. The recrystallized section is of higher purity than the original section, i.e., the motion of the heated section transports impurities to one end of the sample rod which is finally deleted. This technique avoids contamination of the material by a physical contact to a boat-shaped crucible [1.4].

Further processing steps will finally lead to the various devices as they are produced at present, details of which are beyond the scope of this book. Let us just briefly mention *Molecular Beam Epitaxy*, MBE, as a method of producing a very thin layer of one semiconductor on top of another semiconductor which has the form of a small plate[1.5]. The source material is vaporized in a heated cell with a very small orifice (Knudsen cell) creating a molecular beam in ultrahigh vacuum (10^{-11} torr) thus avoiding collisions with residual gases and allowing epitaxial growth of the layer on the target plate (*epitaxial* means an arrangement of the beam molecules as they get stuck according to the arrangement of the atoms on the target surface). The layer quality is recorded during growth by *Reflection High-Energy Electron Diffraction* (RHEED)[1.6] where a 10 keV electron beam incident on the semiconductor surface and reflected at a very large angle of (*grazing*)incidence produces a diffraction pattern on a phosphor screen. *Quantum wells* are produced by this technique where a layer of about

1 nm thickness is sandwiched between two tin layers of another semiconductor with a larger band gap (Fig. 9.5).

1.2 The Positive Hole

As mentioned in the above section, charge transport may be due to *positive holes*. In this chapter we shall explain this idea qualitatively by considering a lattice of carbon atoms which, for simplicity, is assumed to be 2-dimensional. In Fig. 1.1, the bonds between neighboring atoms are covalent. Each C-atom contributes 4 valence electrons and receives from its 4 neighbors 1 electron each so that a *noble gas configuration* with 8 electrons in the outer shell is obtained. This is similar in a 3-dimensional lattice.

Now imagine an extra electron being transferred somehow into the otherwise perfect diamond crystal (by *perfect* we also mean that the surface with its free bonds is far enough away from the electron. Ideally let us think of an infinitely large crystal). No free bonds will be available for this electron where it could attach itself. The electron will move randomly with a velocity depending on the lattice temperature. However, if we apply an external electric field to the crystal, a drift motion will be superimposed on the random motion which, in the simplest case, will have a direction opposite to the field because of the negative charge of the electron. The extra electron which we call a *conduction electron* makes the crystal n-*type* which means that a *n*egative charge is transported. In practice, the extra electron will come from an impurity atom in the crystal.

We can also take a valence electron away from an electrically neutral diamond. The crystal as a whole is now positively charged. It does not matter which one of the many C-atoms loses the electron. The main point is that this atom will now replace its lost electron by taking one from one of its neighbors. The neighbor in turn will react similarly with one of its neighbors. This process is repeated over and over again with the result that the hole produced by taking away an electron from the crystal moves in a random motion throughout the crystal just as the extra electron did in the n-type crystal. What happens if we now apply an external electric field? Wherever the hole is, a valence electron will fill it by moving in a direction opposite to the electric field with the effect that
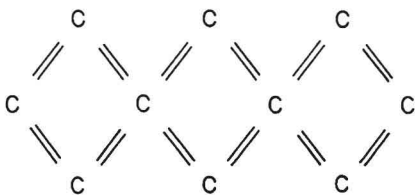

Fig. 1.1: Schematic two-dimensional representation of perfect diamond lattice. Each covalent bond (=) represents two valence electrons of opposite spin

the hole drifts in the direction of the field. This is exactly what one would expect from a positive charge. Since the crystal as a whole is charged positively, we may think of this charge as being localized at the position of the hole. In semi-conductor physics, positive holes are treated as if they were positively charged electrons. Conductivity by *positive* holes is called p-*type*. For a comparison, consider the carbon dioxide bubbles in mineral water. Instead of an electric field, there is the gravitational field and instead of an electric charge, there is the mass of the water molecules. Since the bubbles drift in a direction opposite to the field direction, they can formally be treated like negative mass particles as long as they are in the bulk of the liquid although, of course, carbon dioxide has a positive mass and is subjected to a lift only. Similarly, the assumption of positively charged particles called *holes* in semiconductors is a very simple formal description of an otherwise quite involved process. However, one should keep in mind that the hole is actually a missing valence electron. In Chap. 2 we shall derive the concept of the hole by the more rigorous quantummechanical method.

1.3 Conduction Processes, Compensation, Law of Mass Action

Before becoming involved in wave mechanics, we will continue with the classical model to investigate thermal pair generation and annihilation. Let us call ε_G the binding energy of a valence electron to an atom in the crystal (G stands for *gap* which will be explained in Chap. 2). If the energy ε_G is supplied thermally, a conduction electron may be generated which leaves a hole where the electron has been. The electron and the hole move through the crystal independent of each other. Since we consider the hole as a particle similar to the electron except for the sign of its charge, we have created an electron hole pair. Occasionally a conduction electron will recombine with a hole which actually means that it finds a free bond and *decides* to stay there. The binding energy ε_G is transformed either into electromagnetic radiation (*recombination radiation*) or atomic vibrations (*phonons*). From the particle point of view, the annihilation of the electron hole pair is usually called *recombination*. Denoting electrons by the symbol e^- and holes by e^+, a chemical reaction equation of the form

$$e^- + e^+ \rightleftharpoons \varepsilon_G \tag{1.3.1}$$

will be an adequate description of the process. Assuming that no radiation is incident, the generation energy ε_G is taken from the lattice vibrations. There-fore, with increasing temperature, the equilibrium is shifted towards the lhs of the equation, the number of carriers and therefore the conductivity is increased which is so characteristic of semiconductors. Of course, radiation is incident on

the crystal even if it is in thermal equilibrium with its environment. This *black-body radiation* compensates the recombination radiation of energy ε_G exactly [see the rhs of (1.3.1)].

It is shown in statistical mechanics that a *small system* which is in thermal contact with a *large system* can acquire an energy ε_G at a rate proportional to $\exp(-\varepsilon_G/k_B T)$, where k_B is Boltzmann's constant and T the absolute temperature (at room temperature $k_B T = 25.9\,\text{meV}$). In the present case, the *small system* is a valence electron and the *large system* the crystal. The exponential is multiplied by a power of T; however, the temperature dependence is essentially determined by the exponential, as is well known from, e.g., the law of thermionic emission from metals. For (1.3.1), the power function is T^3 if we apply the law of mass action. Denoting the concentrations of conduction electrons and holes by n and p, respectively, it is

$$n\,p = CT^3 \exp\left(\frac{-\varepsilon_G}{k_B T}\right) . \tag{1.3.2}$$

The value of the constant C depends on the semiconductor material. The form of (1.3.2) is similar to the one describing the concentrations of H^+ and OH^- in water where these concentrations are always small compared with the concentration of the neutral water molecules. In a semiconductor, the electron and hole concentrations will also be small relative to the concentrations of atoms because otherwise, the conductor would have to be classified as a metal. A rigorous derivation of (1.3.2) including a calculation of the constant of proportionality C will be given in Chap. 3 (3.1.14).

In a pure semiconductor, for every conduction electron a hole is also produced and $n = p$. We call this *intrinsic conduction* and add a subscript i to n. Equation (1.3.2) yields

$$n_i = C^{1/2}\,T^{3/2} \exp\left(\frac{-\varepsilon_G}{2k_B T}\right) \tag{1.3.3}$$

In Figs. 1.2a, b, n_i is plotted vs temperature for silicon ($\varepsilon_G = 1.12$ eV at 300 K) and germanium ($\varepsilon_G = 0.665$ eV at 300 K). At temperatures above 250 K (for Si) and 200 K (for Ge), ε_G varies linearly with temperature

$$\varepsilon_G(T) = \varepsilon_G(0) - \alpha T , \tag{1.3.4}$$

the coefficient α being 2.84×10^{-4} eV/K for Si and 3.90×10^{-4} eV/K for Ge. However, this does not change the exponential law (1.3.3), except for a change in the factor C, since

$$\exp\left(\frac{-\varepsilon_G + \alpha T}{2k_B T}\right) = \exp\left(\frac{\alpha}{2k_B}\right)\exp\left(\frac{-\varepsilon_G}{2k_B T}\right)$$

Therefore, in (1.3.3) we use the value of ε_G obtained by extrapolation from the range linear in T to absolute zero. At low temperatures, a T^2 term with the
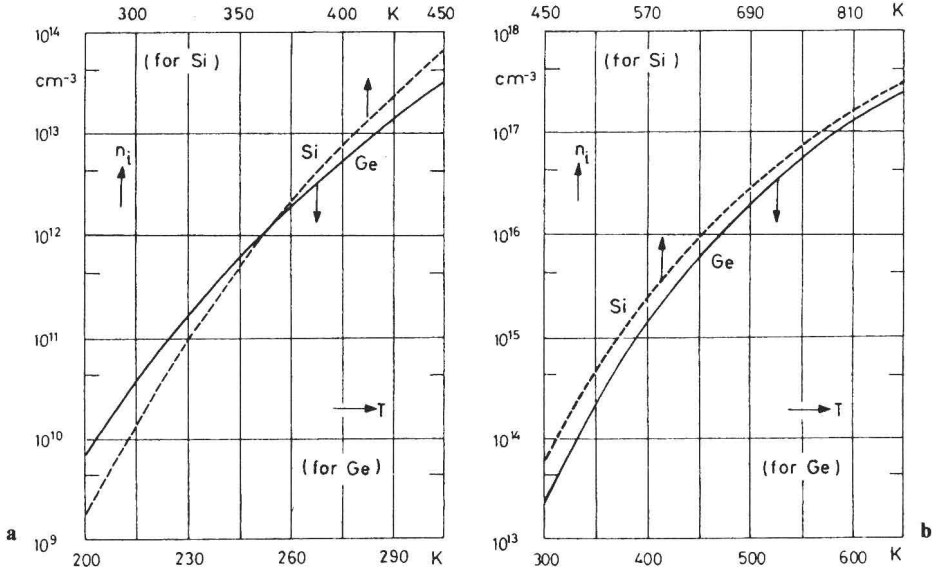
Fig. 1.2: (a) Intrinsic carrier concentration as a function of temperature for silicon and germanium: 200 to 305 K (Ge); 275 to 450 K (Si) (after[1.7]), (b) same as (a) except for temperature range: 300 to 650 K (Ge); 450 to 870 K (Si)

same sign as the αT term is found in ε_G . One obtains these additional terms from optical investigations.

Now we consider a doped semiconductor. Assume one atom to be replaced by an impurity atom, e.g., a 4-valent C atom by a 5-valent phosphorous atom. Only 4 of the 5 valence electrons of the phosphorous are required to bind the 4 neighboring C atoms. The fifth electron is bound very loosely. The binding energy $\Delta\varepsilon_D$ of the fifth electron is considerably lower than the binding energy ε_G of a valence electron to a C atom. An impurity which releases one or several electrons in this way is called a *donor* D. If we denote a neutral donor by D^x and a singly-ionized donor by D^+ , the reaction is

$$e^- + D^+ \rightleftharpoons D^x + \Delta\varepsilon_D \ . \tag{1.3.5}$$

At high temperatures, all donors are thermally ionized and the concentration n of conduction electrons is equal to that of donors, N_D . The concentration p of holes is given by

$$n\,p = n_i^2 \tag{1.3.6}$$

where n_i is given by (1.3.3). Charge carriers which are in the minority are called *minority carriers*. At still higher temperatures, n_i will become larger than N_D and the semiconductor will then be intrinsic. The temperature range where
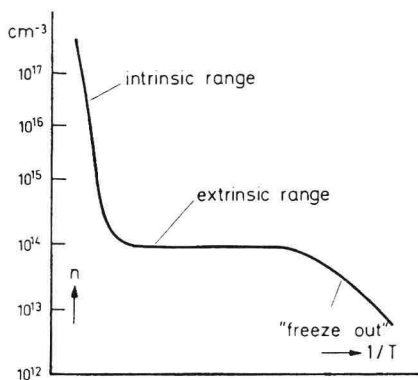
Fig. 1.3: Schematic diagram of the carrier concentration as a function of the reciprocal of the temperature

$n = N_D$ is independent of temperature is called *extrinsic*.[1]   Figure 1.3 gives a schematic diagram of carrier concentration vs the inverse temperature. At the right end of the curve where the temperature is low, the balance of (1.3.5) is shifted towards the right-hand side. There is carrier *freeze-out* at the donors. The application of the law of mass action again yields an essentially exponential temperature dependence with $\Delta\varepsilon_D$ in the exponent. Because $\Delta\varepsilon_D$ is smaller than ε_G , the slope of the curve is less steep than in the intrinsic range.

If in diamond a 4-valent C atom is replaced by a 3-valent boron atom, a valence electron is lacking. Supplying an energy $\Delta\varepsilon_A$, an electron from a C atom is transferred to the B atom and a mobile hole is thus created. The B atom becomes negatively charged. Boron in diamond is an *acceptor A*. Denoting the neutral acceptor by A^x and the singly-ionized acceptor by A^- , the chemical reaction can be written in the form

$$e^+ + A^- \rightleftharpoons A^x + \Delta\varepsilon_A. \tag{1.3.7}$$

The temperature dependence of the hole concentration p is similar to that of n in n-type semiconductors.

If both donors and acceptors are distributed at random in the semiconductor we have a *compensated* semiconductor.[2]   At ideal compensation there would be equal numbers of donors and acceptors. Quite often a high-resistivity material which is supposed to be very pure is compensated. A convenient method of compensation of acceptors is the lithium drift process. The positive lithium ion Li^+ has the helium configuration of the electron shell and diffuses nearly as easily through some solids as helium gas. Diffusion takes place via the interstitial mechanism. The negatively charged acceptors attract the positively charged lithium ions forming ion pairs with the result of ideal compensation. In a p-n junction during diffusion, an external electric field is applied such that ion drift

---

[1] For a *partly compensated* semiconductor N_D should be replaced by $N_D - N_A$
[2] A random distribution of a phosphorous dopant is achieved in silicon by neutron transmutation: Si_{14}^{30}+ thermal neutron $\rightarrow Si_{14}^{31} + \gamma$ (2.62 hr)$\rightarrow P_{15}^{31} + \beta$ [1.7].

produces a p-i-n junction with a large perfectly compensated intrinsic region between the p and the n regions (e.g., a type of γ-ray counter is produced by this method [1.9, 10]).

The normal case is partial compensation of impurities. According to whether $N_D > N_A$ or vice versa, the semiconductor is of an n- or p-type, respectively. Compound semiconductors with nonstoichiometric composition are of an n- or p-type depending on which component of the compound is in excess. Assume that in CdS, e.g., one S^{--} ion is lacking somewhere in the crystal. There is then an excess Cd^{++} ion which will be neutralized by two localized electrons. Therefore, the lattice vacancy acts as an acceptor and makes the semiconductor p-type. In small-band-gap polar semiconductors like PbTe and HgTe, the free carriers are mainly produced by deviations from stoichiometric composition (while in large-gap ionic solids such as sodium chloride, the gap energy is larger than the energy required to generate a vacancy which compensates the ionized impurities and therefore makes the crystal insulating; this is called *auto-compensation*) [1.8].

Problems

1.1. Discuss the temperature dependence of the carrier concentration of a semiconductor (a) in the intrinsic range (b) in the extrinsic range and compare with that of a metal.

1.2. 100 g silicon are homogeneously doped with 4 μg aluminum. What is the hole concentration assuming one hole per aluminum atom? (silicon: density 2.23 g/cm^3, atomic weight 28.086; aluminum: atomic weight 26.98).

1.3. The conductivity is the product of carrier concentration, mobility, and elementary charge. Assuming a mobility of 600 cm^2/Vs, what is the conductivity of the material of Problem 1.2?

1.4. A sample of 3 cm length has been prepared from the material of Problem 1.2. A voltage of 4.5 V is applied between the ends of the sample. How much time is spent by a hole in travelling through the sample? Compare with a copper sample of equal length and for equal voltage (density 8.95 g/cm^3; atomic weight 63.57; resistivity $1.7 \times 10^{-8} \Omega$ cm; assume one electron per atom). Is the time longer than the time it takes to develop enough Joule heat to raise the temperature of the copper sample by 1°C? (apply the well known Dulong–Petit law).

1.5. Four metal pins in a linear array at equal distance D are pressed by springs against a semiconductor sample. Apply a current I through contact No. 1 and 4 and measure the voltage drop V between probes No. 2 and 3. Show that for a sample thickness $\gg$ D the resistivity is given by $\varrho = 2\pi$ D V/I. Modify the calculation for unequal distances. *Hint*: think of the current first going from No. 1 to infinity causing a voltage drop V' and then from infinity to No. 4 causing V''. Potential theory shows $V = V' + V''$.