[redacted]

TRANSLATION

THE ELECTRICAL PROPERTIES OF SEMICONDUCTORS

By S. S. Shalyt

October 1959

157 Pages

STAT

PREPARED BY
LIAISON OFFICE
TECHNICAL INFORMATION CENTER
MCLTD
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

STAT

[redacted]

STAT

S. S. Shalyt

Elektricheskiye Svoystva Poluprovodnikov

Vyp. 1

Leningrad 1957

[redacted]

STAT

Contents

PART 1

PART 2

STAT

Shalyt, Simon Solomonovich

Electrical Properties of Semiconductors

Technical Editor: D. P. Freger.

House for Scientific and Technical Propaganda in Leningrad (LDNTP), Nevskiy pr. 58.

STAT

Part 1

INTRODUCTION

One of the foremost among the scientific-technical problems which absorb the attention of research scientists and working engineers at the present time is the problem of semiconductors. This problem consists in study of the special physical properties which appear in the class of substances known as semiconductors, as well as development of technological applications for these materials.

Of the physical properties which distinguish semiconductors as a special class of solids, their electrical properties are the most important. Indeed, it is the utilization of the specific electrical properties of semiconductors which has recently resulted in progress in technical fields such as the rectification of low- and high-voltage alternating currents, the amplification and conversion of electromagnetic signals (radio engineering), the monitoring, regulation and automation of industrial processes, heat engineering, and electric-power engineering on both large and small scales.

The class of semiconductors comprises chemically heterogeneous substances characterized by definite electrical properties peculiar to them. In the majority of instances, these substances represent nothing new from the standpoint of chemical composition.

However, the electrical properties of these materials are of such great interest and importance that the physicists, whose attention was attracted to them several decades ago, prosecuted rapid and thorough investigations of the field,

[redacted]

STAT

assembled a large body of new information, and were soon indicating prospects for numerous technological applications of semiconductors.

It would be wrong to think, however, that the increased interest exhibited by scientists and engineers in semiconductors at the present time is based solely upon the fact that their properties were not previously known. It would be possible to cite a long series of historical examples which demonstrate that the former neglect of semiconductor materials for use in technology was based on reasons other than a lack of knowledge on the part of scientists of some of their important properties. We need refer only to the thermoelectric properties of metallic and nonmetallic conductors, which were investigated as much as a hundred years ago. While the thermoelectric properties of metals have long since become the property of measurement technique (thermocouples), the outstanding thermoelectric properties of semiconductors have been utilized for measuring purposes and for conversion of heat energy into electrical energy only in relatively recent times. Most curious in this particular case is the fact that the materials used in contemporary thermoelectric generators are just those which were investigated in the first half of the last century. A gap of half a century between the discovery date of an effect and the date of its utilization in practice may also be noted in the history of semiconductor rectifier and photocell development.

Such a state of affairs may be explained on the one hand by the fact that the present period is an epoch of tumultuous scientific and technological progress, and on the other as a result of the fact that only contemporary technology has presented needs which can best be satisfied by the use of semiconductors.

It should also be noted that numerous remarkable electrical properties of semiconductors have been discovered only recently, and that they are still far from having been exhaustively exploited for technological purposes.

The majority of semiconductors belong neither to the class of substances rarely encountered in nature nor to the category of products of new and complex

[redacted]

STAT

chemical syntheses. The materials which interest us here are available in practically unlimited quantities, since the majority of solid minerals with which we deal in nature are essentially semiconductors. Oxides, sulfides, tellurides and selenides of numerous metals, as well as certain intermetallic compounds, possess typical semiconductor properties.

Do any of the chemical elements, in the pure state, possess semiconductor properties? No positive answer to this question was available until comparatively recently. This state of affairs has been considerably clarified as a result of a tremendous volume of research, both in the direction of obtaining chemically pure elements, and that of the study of their properties. It has been found, for example, that silicon and tellurium, both of which were formerly considered to be metals, are typical semiconductors in the pure state. As it developed, such insulators as sulphur, phosphorus and iodine must be included in the category of semiconductor substances. Figure 1 presents that part of **Mendeleyev's periodic system** in which elements with semiconducting properties are concentrated. It may be seen from their arrangement in the table that they form a compact group of elements which occupies an intermediate position between the metals and the nonconductors: located to the left of and below the group of semiconductor elements are the highly-conductive metals, for which the weak bond between valence electrons and atoms is characteristic; and to the right of and above this group are those elements which, in the solid state, behave as insulators characterized by considerable affinity to electrons, i.e., a tendency to acquire the latter in the course of chemical reactions with metals.

A similar arrangement may be observed within the isolated group of elements itself: the elements located in the upper right side--C (diamond), P, S and I--are essentially nonconductors, and fall into this group because of their high photoelectric conductivity, which is a characteristic of all semiconductors; the elements in the lower left corner-- Sn and Sb--behave as metals in their stable modifications,

STAT

and semiconductor properties are exhibited only by specially prepared films of these elements. This location of the semiconductors in the periodic table is not accidental. It will be shown below that as is true for all solids, the electrical properties of semiconductors are determined in the last analysis by the structure of the outermost electron shells of the atoms, while the periodic completion of electron shells is the basic factor which determines the chemical properties of the elements, and, consequently, their location in the periodic table. In a large circle to the right of the symbol for each semiconductor element (Fig. 1) appears a value basically characteristic of the element's electric properties; its physical significance will be explained below. (Width of forbidden band; see pages 42- 43). The regularity of the variation of this number with respect to both the columns and lines of the table indicates a close association between the electrical characteristics of semiconductors and their location in the Mendeleyev system. These electrical characteristics are compared in the same table to values which present, in chemistry, a rough measure of the attraction of the valence electrons to some interacting atom, and are called electronegativities of the elements (small circles on the left). There is no doubt as to the marked correlation between these values.

Fig. 1. Elements with semiconductor properties as they are arranged in the periodic table.

However, the fact that the same element may have completely different electrical properties in different crystalline modifications (for instance, carbon in the form of diamond or graphite), indicates that these properties are determined

[redacted]

STAT

not only by the element's position in the table of the elements, but by the nature of the chemical bonds between the atoms of the solid as well.

Let us characterize briefly the 12 elementary semiconductors which have been isolated in Fig. 1. The most typical semiconductors, whose electrical properties have been thoroughly investigated, are germanium, silicon and tellurium. Germanium is the material most important to contemporary semiconductor technology, and has therefore been investigated with the utmost thoroughness--certainly with greater care than any other element; its disadvantage, from a practical point of view, is its high cost, which results from the fact that germanium occurs in nature only in scattered small deposits and from the difficulty encountered in extracting it. Silicon is one of the most widely distributed terrestrial elements (28% of the earth's crust); certain of its physical characteristics, which are of great significance in technological applications, are superior to those of germanium, but since the technique of industrial-scale preparation of highly pure silicon--which manifests its remarkable properties only in this state--is still in the familiarization stage, we can speak at present only of its tremendous possibilities. Tellurium--a semiconductor which has found no application in its pure state--is the basis for the preparation of a number of semiconductor alloys which, in particular, are used in the fabrication of thermal generators and coolers.

Antimony and arsenic exhibit semiconductor properties only when prepared in the form of films by sublimation in a vacuum onto a cool base; at temperatures exceeding 0° C for antimony and 300° C for arsenic, the semiconductor modifications of these elements become unstable, and the specimens are converted to the normal metallic state.

Tin loses its ordinary metallic properties and becomes a semiconductor upon the polymorphous conversion of the white modification to the gray modification (below 13°C).

Still insufficiently investigated are the electrical properties of pure boron.

STAT

However, there exists no doubt as to its classification among semiconductors.

At room temperature, gray selenium, red phosphorus, sulfur and iodine are essentially nonconductors; their semiconductor properties appear in the form of a great increase in conductivity when exposed to light (photoconductivity). Sulfur, for example, which is one of the best insulators at room temperature undergoes a millionfold increase in conductivity under the action of light.

Carbon, finally, is a semiconductor which, depending on the spatial arrangement of the atoms in its crystal lattice, forms solid bodies with either insulating, semiconductor or metallic properties. In the form of diamond, carbon is an insulator sensitive to electromagnetic radiation; in the form of graphite, it is either a semiconductor or a metal, depending on the orientation of the current with respect to the axes of the graphite crystal.

While pointing out that there is nothing new in principle in semiconductor chemistry, we should nevertheless make the reservation that the successful application of semiconductors in contemporary technology involves the solution of difficult problems in physics and chemistry. The manufacture of perfect rectifiers, amplifiers and photocells requires that the material be purified so thoroughly that not even the most skillful chemist would have undertaken it ten years ago. Today the technique of obtaining such high purity--for example, in germanium--has become generally accessible. The problem of utilization of silicon, which, in this particular field of semiconductor application, has considerable advantages over germanium, remains essentially the unsolved problem of its purification. In another important field of semiconductor technology, the goal of which consists in manufacturing economically profitable thermal generators, the problem of the purity of materials is replaced by another, no less difficult problem of manufacturing heat-resistant semiconductor materials with specified electrical and thermal properties.

At the present time, numerous groups of chemists, physicists, metallurgists, radio and electrical engineers and other specialists are at work in most countries of

STAT

[redacted]

the world on the problems of mastering semiconductor techniques and in thorough investigation of the physical properties of semiconductors. The study of semiconductors is of scientific value in that it enables us to establish most completely the general physical laws which govern, to varying degrees, the behavior of the different solids--the metals, semiconductors, and insulators. It is not an exaggeration to state in conclusion that the scientific, technological and economic importance of the semiconductor problem, and the attention which it invites, place this problem on the same level with that of the utilization of nuclear energy.

[redacted]

STAT

CHAPTER I

GENERAL TREATMENT OF CONDUCTIVITY IN SOLIDS. CONCENTRATION AND MOBILITY OF CURRENT CARRIERS

The term "semiconductors" itself implies that the general properties which unify this chemically heterogeneous group of substances are their electrical properties -- or, to be more accurate, the laws which govern the passage of electric currents through them.

The conception according to which the electric current generated in a solid by an external electric field constitutes a directed migration of charged particles, or current carriers, which is superimposed as a "drift" upon the chaotic motion of the particles which prevails in the absence of an external field, has long been in existence. These charged particles are either electrons or the positive or negative ions of the substance. Simultaneous participation in the electric current of all these types of carriers is possible and has already been observed. In the technological application of semiconductors which is of interest to us, the fact that either electrons or ions may participate separately in the electric current is important. In the first instance, we speak of electronic conductivity, and in the second of ionic conductivity. The ionic, or, as it is often called, electrolytic conductivity of solids is a result of their gradual chemical decomposition during passage of an electric current through them. This fact compels us to disregard substances in which conductivity is due to any appreciable extent to the migration of ions, since the utilization of such materials would not satisfy the basic specifications set forth for semiconductor equipment -- that is, operational

[redacted]

STAT

stability and long service life. The experience of several years indicates that if electric current is generated by the motion of electrons only, no chemical effect of the current will appear. Hence at the present time, when we speak of semiconductors, we refer to solid substances in which current is generated solely by the motion of electrons, despite the fact that formerly, when semiconductors had been investigated only very poorly, this term applied to all solid substances with small but detectable conductivities, irrespective of the nature of the current carriers.

It is known that since the electric current is subject to Ohm's law, the specific electrical resistance ρ, or the uniquely-related value referred to as the specific conductivity σ $\left(\sigma = \frac{1}{\rho(\text{ohm cm})} \text{ ohm}^{-1} \text{ cm}^{-1}\right)$ may be taken as a quantitative measure of the conductivity of a solid.

The specific conductivities of good conductor-metals vary within a range of two orders of magnitude between 10^4 and 10^6 ohm^{-1} cm^{-1}. In the other extreme case -- that of the very poor conductors, or insulators -- this value is usually smaller than 10^{-10} ohm^{-1} cm^{-1}. The specific conductivities of the bodies referred to as semiconductors occupy the entire extensive range of values from 10^3 to 10^{-10} ohm^{-1} cm^{-1}, thus varying through a vast region which encompasses 13 orders of magnitude.

It is not the quantitative distinction between semiconductors and metals, however, which leads us to single them out as a special group of substances and permits their utilization in various specific technical fields. A basic and characteristic property of a semiconductor is that the value of its specific conductivity may vary greatly under the influence of various factors, such as temperature, light, pressure, etc., and, what is of great importance, that the conductivity of semiconductors may be controlled by adding very small quantities of impurities to them. It should also be pointed out that the numerous and varied applications of semiconductors in contemporary technology are due not only to the above characteristic effects of impurities, temperature, illumination, etc., on conductivity, but also to a number of other specific physical properties, inherent to one degree or another in the different groups of semiconductors. These properties include the ideal rectifying and amplifying properties of semiconductor junctions, high thermo-and photoelectromotive forces, the high electronic emission of semiconductor cathodes, outstanding luminescent and catalytic properties, a remarkable combination of high electric resistance and ferromagnetism, and many other properties upon which the various uses of semiconductors in technology are based.

Proceeding from absolutely general concepts of the mechanism of the electric current, it may be indicated that the specific conductivity of a solid is determined by the expression

$$\sigma = \underline{enu},$$

where $\underline{e} = 1.6 \cdot 10^{-19}$ coulomb is the electronic charge, $\underline{n}$ is the concentration of current carriers (i.e., their number per unit volume), and $\underline{u}$ is a quantity which is termed the mobility of the current carriers.

Let us clarify the physical significance of this expression. It has already been noted above that electric current is created by the directed motion of free electrons, which exist in a solid irrespective of the presence of an external field, and move in a random manner similar to that of molecules of a gas. However, while the trajectories of the gas molecules are determined by the collisions that occur between them, the trajectories of the electrons, based on a number of physical considerations which we shall consider later, are determined not by collisions, even with the quiescent atoms of the substance, but by collisions with thermal vibrations of atoms and with lattice defects, a more detailed discussion of which will also follow below.

We assume that the electrons move freely in the intervals between collisions, and, consequently, rectilinearly and uniformly. Let $\underline{l}_1$ denote the length of free path and τ the corresponding time of free path. These two quantities stand in the obvious relationship $\underline{l}_1 = \underline{v}\tau$, where $\underline{v}$ represents the velocity of motion which, in the general case, differs with different electrons. The chaotic motion of electrons in the absence of an external electric field does not create an electric current; when a field is applied, the electrons experience an acceleration in the direction opposed to that of the field under the influence of the electrical forces. The secondary mean directed velocity $\Delta \underline{v}$ which the electrons attain over the free path will be just the velocity which determines the magnitude of the electric current. Figure 2 illustrates the considerations outlined here: the solid arrows pointing in different directions represent the velocities of the random motion of the electrons in the absence of a field; the solid arrows pointing from right to left represent the mean secondary directed velocities attained by electrons over the free path when an external electrical field of intensity E is applied; the dotted arrows represent the resultant velocities. In this simplified diagram, all random velocities are represented as being equal, while the secondary directed velocities attained by such electrons will not be equal due to the disparity between the lengths of free path for random motion. Such is the pattern of the motion of electrons in metals.

In semiconductors, the situation is much the same as in gases: the velocities of the electrons involved in conduction have all values from zero to infinity, and are distributed according to the Maxwell law. The case which occurs in metals is characterized by nearly absolute temperature-independence of the electron velocities and is therefore termed the degenerate case. Just as in the case of gas molecules, the electron velocities in semiconductors increase with rising temperature, and there are few instances in which the electron gas in semiconductors approaches the degenerate state in which electron velocities depend only slightly on the temperature of the substance. (The conditions in which degeneration of electron gas occurs will be discussed below on page 58.) Hence, the classical Maxwell-Boltzmann statistics can nearly always be applied to semiconductors. The more general Fermi-Dirac quantum statistics, which includes the classical statistics as a particular and, mathematically speaking, more simple case applies for the

STAT

degenerate electron gas.

Let us turn to the diagram in Fig. 2, and, applying the known laws of elementary physics, calculate the average supplementary velocity $\overline{\Delta v}$. The acceleration of an electron $a = \frac{eE}{m_0}$, and the average supplementary velocity of one electron over the length of free path $\Delta v = 1/2 \frac{eE}{m_0} \tau$, where m_0 is the electronic mass and τ is the free time for this electron.

Fig. 2. Diagram illustrating creation of electric current in a solid.

Taking the statistical distribution of time intervals τ into account in the normal manner, the average additional velocity of the entire ensemble of electrons in the solid will be

$$\overline{\Delta v} = \frac{eE}{m_0} \bar{\tau} , \qquad (1)$$

where $\bar{\tau}$ is the mean free time of the electrons.

The electric-current density is determined by the quantity of charge which the electrons carry across a unit cross section of the conductor per unit time:

$$j = en \, \overline{\Delta v} = en \, \frac{eE}{m_0} \bar{\tau} \cdot$$

By definition, the specific conductivity σ is given by the ratio of the current density j to the field intensity E:

$$\sigma = \frac{j}{E} = \frac{e^2 n}{m_0} \bar{\tau} = e\,n\,u . \qquad (2)$$

The value $u = \frac{e\bar{\tau}}{m_0} = \frac{\overline{\Delta v}}{E}$ is referred to as the mobility of the current carriers, since it determines the average supplementary velocity $\overline{\Delta v}$ attained by the electrons in a field of unit intensity (for example, 1 volt/cm).

STAT

If we express the charge in coulombs ($\underline{e} = 1.6 \cdot 10^{-19}$), $\underline{n}$ in cm^{-3} and $\underline{u}$ in cm^2/ volt·sec, then σ will be obtained in $ohm^{-1}\ cm^{-1}$.

Expression (2) is universal, since it applies both for the mechanism of electron conduction and the mechanism of ionic conduction (in the latter instance the numerical values of the quantities which characterize migrating charges will, of course, be different). Nor does this expression, obtained with the aid of simple classical considerations, change its form when we subject the phenomena of conductivity to rigorous investigation in terms of quantum mechanics; in the quantum theory, however, certain of the values in Expression (2) assume interpretations significantly different from those of the classical theory. This applies, in the first place, to the free time τ, which is defined in both the classical and the quantum theories as the ratio $\tau = \frac{l}{\underline{v}}$. In the classical theory of conductivity, the free path $\underline{l}$ was determined by interatomic distances, since it was assumed that electrons are scattered upon collision with the atoms of the body which obstruct the directed motion of the electrons even when at rest; as to the velocity $\underline{v}$, it is assumed in the classical theory that this quantity is determined by the simple condition of uniform distribution of thermal energy among the electrons and the vibrating atoms in the lattice. The quantum theory has essentially changed our concepts regarding both the free path of electrons in solids, and the conditions by which their velocities are determined. It has become evident that the length of free path in an ideally ordered crystal lattice is equal to infinity (and that the electrical resistance equals zero), so that the quiescent atoms of a strictly periodic lattice do not present obstacles to the electrons which participate in the electric current. In accordance with the quantum theory, electrical resistance in a crystal occurs only when the ideal ordered state of the atoms (or ions) is disturbed by thermal agitation of the atoms (or ions) themselves or lattice defects. Lattice defects include lattice points replaced by foreign atoms, lattice vacancies, atoms (or ions) implanted in interstices, shear between crystal layers, cracks,

[redacted]

STAT

grain boundaries, etc. Thus in perfect monocrystals the free path of an electron can reach tens, hundreds and -- at low temperatures -- even tens of thousands of interatomic distances. The velocity of the random motion of free electrons in solids is determined in the general case from specific quantum conditions which, by the way, lead to the same classic results in the case of the degenerate semiconductors under consideration.

In the above elementary derivation of the formula for the conductivity of a solid, an essential simplification has been made: we have assumed that an electron moves freely along its free path. In actual fact, we know very well that at the distances which separate free electrons from atoms in a solid, very strong electric fields are active, and that their forces exceed by far those of the external fields in which we usually place the solid. These fields must necessarily exert great influence on the motion of the electrons which, for this reason, may not be considered as a free motion even in rough approximation. The quantum theory shows that the effect of strong atomic fields may be taken into account by formally ascribing to the electron a certain mass different from the mass of the freely moving electron. It develops that such an operation enables us to study electron motion in solids without taking the effect of strong internal fields into account. The mass which we ascribe to an electron of a solid is called the effective mass $\underline{m}$. This value is determined by the structure of the medium in which the motion of the electron occurs. In different bodies it may be either larger or smaller than the mass of the free electron $\underline{m}_o = 9 \cdot 10^{-28}$ g. We shall repeatedly be confronted by this problem in the material which follows.

It will be seen from Formula (2) that the direct proportionality between the current and the field known as Ohm's law will occur only in the case when the mobility $\underline{u}$ and the concentration of current carriers $\underline{n}$ do not depend on the field E; in the contrary case, the specific conductivity σ will be a function of the field E, and Ohm's law will not apply. In metals, Ohm's law remains valid under all

STAT

conditions; only in the presence of tremendous current densities which are practically never used can infinitesimal deviations be observed; the reality of these, however, is still a matter of dispute. Deviations from Ohm's law may be observed in semiconductors under practically realizable conditions, and even find applications in technology. The range of electric-field strengths in which Ohm's law is satisfied is termed the weak-field range; the deviations from Ohm's law will be related to the influence exerted by strong electric fields on mobility and concentration. Due to its importance for semiconductors, this phenomenon will be discussed in detail in a separate chapter.

Thus, according to (2), the conductivity of a solid is determined by the concentration $\underline{n}$ and the mobility $\underline{u}$ of current carriers. Let us now establish the reason for the lower conductivities of semiconductors as compared with metals: is it due to the fact that concentration in semiconductors is lower, or does it result from lower mobility of their current carriers? To solve this problem experimentally, it is necessary to determine, in addition to the readily measured value σ, either $\underline{n}$ or $\underline{u}$ by an independent method, and then each of these quantities separately with the aid of Formula (2).

Fig. 3. Scheme of experimental Hall-effect measurement.

There exist various independent methods for determination of the concentration and mobility of electrons on the basis of measurements of the various galvanomagnetic and thermomagnetic effects -- i.e., of physical phenomena which arise in semiconductors in a magnetic field when free electrons acquire a directed motion in this field under the influence of potential or temperature differences. We shall show below why the Hall effect, among eight of these effects, is the most

STAT

suitable for this purpose.

Since we shall refer repeatedly to the Hall effect as the experimental basis for numerous important conclusions regarding the properties of semiconductors, let us first recall the experimental procedure used for measurement of this effect.

It is convenient to use a test conductor having the shape of an elongated plate of rectangular cross section. If we place this conductor in a magnetic field H and pass through it a current I, as indicated in Fig. 3, the potential difference ΔV_H appears across the conductor under the influence of the magnetic field; it disappears with the elimination of the magnetic field. The occurrence of this potential difference is called the Hall effect. This effect is the result of the deflecting action of the magnetic field on the current carriers moving along the plate. Theory and experiment indicate that the magnitude of Hall's potential difference in not too strong magnetic fields is determined by the expression

$$\Delta V_H = R_H \frac{I}{ab} Hb = R_H \frac{IH}{a},$$

where $\frac{I}{ab}$ is the current density j.

The sign of this potential difference -- i.e., the direction of the Hall electric field -- is uniquely related to the sign of the current carriers, and may be determined according to the left-hand rule, which is known to us from high-school physics. With directions as noted in Fig. 3, the current will be deflected downward; consequently, electrons moving against the external field $\vec{E}$ must be deflected downward, thereby imparting a negative charge to the lower edge of the plate. Should positive charges moving along the electric field $\vec{E}$ participate in the electric current, the lower edge of the plate would become positively charged, and the direction of the Hall potential difference would be reversed. Thus the qualitative investigation of this effect enables us to verify the sign of the current carriers. The value of the concentration of current carriers n which is of interest to us can be determined from the value of the Hall coefficient R_H computed from experimental data, since it is possible to show by simple deductions based on the elementary laws of electro-dynamics that in the case of a single species of current carrier

$$R_H = \frac{1}{ne} \text{ cm}^3/\text{coulomb.}$$

Indeed, it is possible to determine the Hall field from the condition of equality of the force evH exerted upon the electron by the magnetic field to the force of the magnetic field eE_H. Taking into account that the current density $j = evn$, and $E_H = R_H jH$, we obtain $R_H = \frac{1}{en}$.

Upon absolutely rigorous derivation, taking the statistical distribution of electron velocities into account, this expression takes the form

$$R_H = \frac{A}{ne} \text{ cm}^3/\text{coulomb,} \qquad (3)$$

where A is some numerical coefficient dependent upon the degree of degeneration of the electron gas and on the mechanism of current-carrier dispersion. The advantage offered by the Hall effect in the choice of a method to determine n and u is due to the fact that the numerical value of the coefficient A, which depends in a complex manner on the above-mentioned factors, varies in the relatively narrow range from 1 to 2. In the other methods, we are also confronted by numerical coefficients which also depend in a complex manner on the same factors; in these instances, however, the intervals of variation of the respective coefficients are much larger. Since the exact determination of the degree of degeneration of the electrons and the mechanism of their dispersion constitutes a separate and difficult problem, we shall incur the smallest error from a none too reliable solution if we use the Hall effect to determine n and u. The fact that the Hall effect is nearly independent of the anisotropy of the crystal when measured for a weak magnetic field is a further advantage of this method.

Experiments designed for measurement of the Hall effect in metals and the

[redacted]

STAT

semiconductors have indicated that the lower conductivity of semiconductors may be explained by smaller concentrations of current carriers. Concerning the mobility u, this quantity can be both larger and smaller in semiconductors than in metals. Whereas the mobility in metals will normally amount to hundreds of $\frac{\text{cm}^2}{\text{v}\cdot\text{sec}}$ at room temperature, this quantity varies in semiconductors from unit numbers to tens of thousands of $\frac{\text{cm}^2}{\text{v}\cdot\text{sec}}$ (InSb - 80,000; InAs - 30,000).

It is therefore of interest to note that high mobility of current carriers can be observed in solids despite small conductivity values. If, for instance, a diamond -- normally a good insulator -- is subjected to electromagnetic or corpuscular radiation, free electrons may be created in it and the passage of a weak current through it may also be observed; in this case, it develops that the mobility of the electrons participating in the current attains a value of 1000 $\frac{\text{cm}^2}{\text{v}\cdot\text{sec}}$, i.e., exceeds that of the electrons in typical metals. The same can be observed in semiconductors at low temperatures, when the mobility attains hundreds of thousands of $\frac{\text{cm}^2}{\text{v}\cdot\text{sec}}$ with very low conductivity. We shall return to the problem of the mobility of current carriers and of its physical significance after having familiarized ourselves with a number of important aspects of the electronic theory of crystals.

Experiments in the measurement of the Hall effect in semiconductors have shown that the strong influence of temperature, light and impurities on their conductivity is basically the result of a change in the concentration of current carriers, although it should be noted that appreciable though much more weakly manifested changes in mobility occur upon a change in temperature and upon the introduction of impurities. Similar experiments made with metals indicate that their current-carrier concentrations are practically constant while the effects of temperature and impurities are asserted in the length of free path only, or, which is the same thing, in the mobility value (which, however, is not always used in describing the electric properties of metals).

The very sharp drop in the concentration of current carriers which is often observed in semiconductors upon a decline in temperature suggests that the conducting state of such materials is a state of excitation, and that semiconductors must become nonconductors at absolute zero due to the total "condensation" of the electron gas, which begins to "evaporate" as the temperature rises above absolute zero, and to a greater extent as the temperature of the body increases. The strong influence of light also becomes understandable from this point of view: electromagnetic quanta (photons) which are absorbed by the solid provide the current carriers with the energy necessary for their "evaporation," or, better said, their activation.

This conclusion, combined with what we have said above concerning the nature of the current carriers, enables us to establish points of resemblance and dissimilarity between the three classes of solid conductors: metals, semiconductors and electrolytes. Considered current carriers, semiconductors are similar to metals and not to electrolytes; from the viewpoint of the energy aspect of the conductive state itself, semiconductors are similar not to metals but to electrolytes, in which the state of conductivity is also one of excitation. The concentration of current carriers is practically independent of the effects of temperature and light; in semiconductors these factors exert decisive influence on concentration, although it is necessary to make the reservation that this conclusion is not universally valid. In every semiconductor there exists a temperature range in which the concentration remains constant, and the influence of light, even when absorbed by the semiconductor, is not necessarily asserted in its conductivity.

The concentration of current carriers in some semiconductors remains constant over a very wide range extending to the lowest temperatures, as in the case of metals. Such substances are not normal metals, since, first, their concentration of current carriers is relatively low and, second, it increases rapidly with rising temperature as soon as the latter reaches a certain value which is quite definite for each of these substances (see page 87). Such materials, which possess metallic properties at low and medium temperatures and the properties of semiconductors at elevated and high temperatures are called semimetals. Semimetals are used in technology at the present time because their electrical and thermal properties still come closest to satisfying the requirements made of materials for thermogenerator construction.

Similarly, even our conclusion concerning the strong effect of light on the conductivity of semiconductors is not universal. The influence of light varies from a degree barely perceptible even with the most precise measurement to a tremendous effect in which the conductivity under illumination is increased thousands and even millions of times as compared to the dark conductivity. This phenomenon, which has been named photoconductivity, is highly selective: this means that it is not observed under the influence of light of any wave length, but only under that of electromagnetic radiation at wave lengths confined to a relatively narrow range which is frequently situated outside the visible band of the spectrum.

Continuing our reservations in regard to the universal prevalence of wide variation in the concentration of current carriers in semiconductors, it should be pointed out that even the presence of impurities does not always exert any great influence. In certain semiconductors, the introduction of trace (< 0.01%) quantities of impurities increases conductivity thousands or even millions of times; the same procedure in other semiconductors interferes with their conductivity to a nearly imperceptible degree. Trace impurities of one chemical nature may prove highly effective with a given semiconductor, while other chemical impurities introduced in much larger quantities have but a weak effect on conductivity; here it sometimes happens that conductivity decreases, instead of increasing, as a result of the introduction of impurities.

CHAPTER II

THE SIGN OF THE CURRENT-CARRIER CHARGE IN SEMICONDUCTORS: ELECTRONS AND HOLES

It was pointed out above in explaining the the nature of the Hall effect that the direction of the transverse Hall electric field is uniquely related to the sign of the current-carrier charge; it was noted further that the mechanism of conductivity in the semiconductor materials which interest us here is electronic. It might be assumed that these two conditions should make it possible to draw a definite conclusion as to the direction of the Hall electric field which should be obtained in investigations of such materials. However, it was observed long ago--and long remained an enigma--that the "correct" direction of the Hall electric field occurs as often as the "incorrect" direction--i.e., the reverse direction. The correct directions, which correspond to the negative sign of the electronic charge, came to be referred to as cases of negative or n-conductivity, and the "incorrect" cases are said to be characterized by positive or p-conductivity. Experiments have indicated and theory has confirmed that the sign of the conductivity observed in the Hall effect also appears in certain other galvanomagnetic and thermomagnetic effects, as well as in the thermoelectric effect. The latter provides the simplest method by which to determine the sign of conductivity, since all that is necessary for this purpose is to determine the direction of the thermoelectromotive force which arises upon nonuniform heating of the conductor; this method does not require external electric- or magnetic-field sources.

Experiment indicates further that the sign of conductivity in a given

STAT

semiconductor may change either upon variation of temperature or the introduction of one or another impurity. Thus, for instance, for tellurium of average purity the sign of conductivity is always positive at the temperature of liquid air, irrespective of the nature of the impurities: as the temperature is elevated to room temperature, the sign of conductivity becomes negative, and with further elevation of temperature to several hundred degrees above room temperature, it becomes positive again. In some semiconductors, including germanium, silicon, sulfide of lead, telluride of lead and many others, the sign of conductivity changes in dependence upon the chemical nature of the impurities which they contain. In others, the sign of conductivity is independent of the chemical nature of the impurities; thus, for example, we deal constantly with cuprous oxide, copper sulfide, selenium, manganese [sesqui-] oxide and other similar semiconductors whose conductivity is always positive; on the other hand, the oxides of zinc, aluminum and titanium and the sulfides of silver, cadmium and mercury are examples of semiconductors whose conductivity is always negative.

Since the possibility of participation of positively charged ions in the electric current is excluded by the experimentally established fact of the absence of electrolytic decomposition in the substances of interest to us, the thought of possible participation of positrons in the electric current occurs to us first when we turn to consideration of positive conductivity in electron conductors. Yet this hypothesis will be rejected immediately when we consider that positrons are not stable in a medium containing electrons, and that the formation of positrons is associated with the consumption of a tremendous energy (1 Mev) which cannot possibly occur in a nonradioactive material; the average thermal energy of atoms in a solid at room temperature amounts to several hundredths of an ev, while the energy increase in current carriers in an external electric field, does not even exceed a hundredth of this value. Thus there remains only to search for some other explanation of the fact that electronic conductivity may manifest itself in different ways--either in the form of a simple motion of negatively charged electrons which is readily understood from elementary concepts, or in the form of some other, more complex motion of these particles which is equivalent to the motion of positive charges which, as experiments show, have approximately the same mass as the electrons participating in negative conductivity. A valid and rigorous explanation of this paradoxical phenomenon has been found in the course of the theoretical study of electron behavior in the crystal lattice of a solid. In essence, this explanation is associated with the particular wave properties of electrons and with the periodic structure of the crystal lattice. The rigorous treatment of this problem is carried out with the aid of the complex mathematical apparatus of contemporary wave mechanics, which does not permit exposition of the physical essence of this phenomenon on the basis of the graphic concepts of classical physics. Nevertheless, and accurate, though not rigorous interpretation of the experimentally established fact that two types of conductivity exist may be presented without reference to the difficult mathematics involved in the quantum theory, using as a basis the general concepts of atomic structure which are known to us from general physics courses. By generalizing these concepts, it is possible to construct a descriptive model of a solid which demonstrates, firstly, the cause of the experimentally observed differences in the electrical properties of metals, semiconductors and insulators, and, secondly, explains the observed effects of temperature, light and impurities on the conductivity of semiconductors which form the basis for their technological applications.

Let us proceed from the familiar assumption that the electrons of free atoms may not possess arbitrary energy values, but only quite definite or, so to speak, discrete, quantized values which are separated from one another by wide forbidden intervals. The energy levels of electrons in an isolated atom, as is known, may be represented schematically in the form of a vertical series of lines; their positions in this series correspond to the energies of a given state (see the right

[redacted]

side of Fig. 4). The intervals between the lines correspond to the intervals of forbidden energy values. A second circumstance which we must take into account in constructing the general scheme of the motion of electrons in a solid consists in the fact that according to the Pauli principle, which is the basis of the quantum theory of the atom, no more than two electrons can exist in one quantum state, which is determined by three quantum numbers. In the course of formation of a solid through the union of individual atoms, the effect of the above restrictive laws begins to apply to the entire microscopic volume of the solid which, in this instance, should be considered as a gigantic molecule consisting of a great number of atoms. Quantum theory indicates that as identical atoms draw together and atomic interaction begins to influence the energy states of electrons, levels which were identical in the case of the free atoms are now displaced along the energy scale by an amount which increases with increasing interaction between atoms--i.e., with their approach to one another. There forms in the solid, from N identical levels of the widely separated atoms, an energy band consisting of N closely juxtaposed different levels. It will be recalled that 1 cm^3 of a solid contains approximately 10^{22} - 10^{23} atoms. The entire tremendous number of electrons in the section of the solid with which we are experimenting form a collective and present a single system which is subject to the same quantum laws as the limited electron system of the individual atom. Yet instead of the restricted level of an individual atom, which contains no more than two electrons (if the level is not degenerate), wide energy bands containing as many levels as there are atoms in the section under consideration form in the solid. If an energy level in a free atom is g-fold degenerate, the energy band formed from it may also be degenerate, and in this case $2g$ electrons may be distributed in each level of the band; but instances are also possible when, under the influence of the internal electric field of the crystal, this energy band is decomposed into several bands (the maximum number of which may not, of course, exceed g). It will be recalled that the integer $g = 1, 2, 3, \ldots$,

[redacted]

STAT

which characterizes the multiplicity of degeneration of an atomic energy level, is the number of different states of the given atom which possess equal energy.

Moreover, it is necessary to note that the order of band distribution in the energy diagram of a solid may not correspond to the order of energy-level distribution in the free atom; in this case, electrons are redistributed according to the Pauli principle in the course of formation of the solid, in such a way as to occupy all lower levels of the energy spectrum. The left side of Fig. 4 presents schematically the genesis of energy bands in a solid as atoms draw together. The bands, or, better said, the intervals of allowed energy values are separated from one another by intervals of forbidden energy values; for the external, so-called valence electrons, the latter have a width of the same order as that of the allowed energy intervals (see Fig. 5). These intervals (bands or ranges) of allowed and forbidden energy values are referred to as energy bands. As seen from Fig. 5, the width of the allowed bands increases with increasing energy of state, while the width of the forbidden bands decreases in the same direction. As in the case of an individual atom, in which a quantum level may be either filled or unfilled (or partially filled), the energy bands in a solid may be filled to different degrees (in extreme instances, they can be either completely filled or completely empty). And just as in an

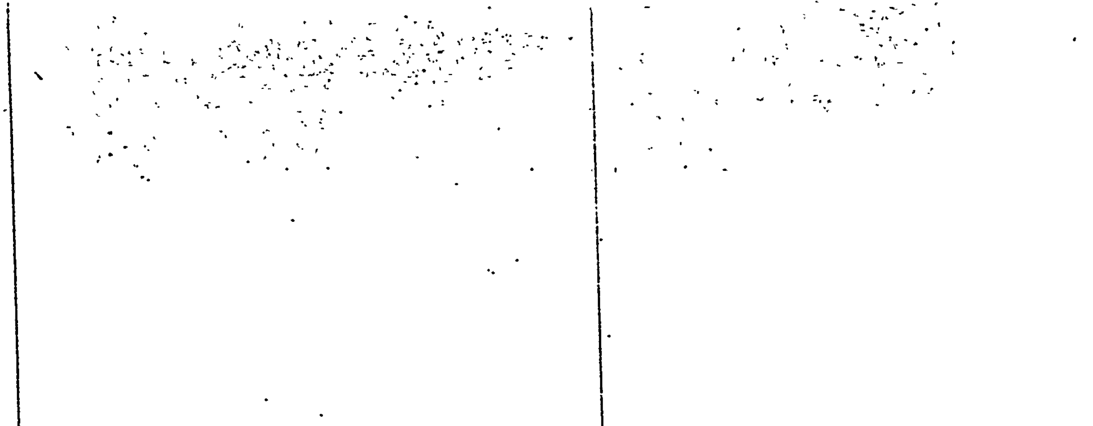

Fig. 4. Splitting of levels occuring on the rapprochement of atoms, and the formation of the energy spectrum in a solid. Distances between energy levels are not given to scale; r_0 stands for interatomic distance in a crystal.

[redacted]

individual atom, where the transition of an electron from a lower quantum level to a higher one requires an external energy source equal to the difference in energy between these levels, it is necessary for the displacement of an electron from a lower to a higher energy band in a solid to consume an energy equal to the width of the forbidden band which lies between them. The same condition must also be observed in the case of electron displacement within the boundaries of a single allowed band. An energy equal to the difference between the respective levels is liberated in the reverse transitions.

The density of the levels in the allowed bands is very large, since the width of the energy bands in a solid does not exceed several electron-volts, while the number of levels in them is equal to the number of atoms in the volume of the solid. With 10^{22} atoms--a number corresponding to a solid volume of approximately 1 cm^3 -- the difference in energy between the neighboring levels amounts to 10^{-22} ev. This condition permits us to assume that the energy levels in a band form a practically continuous spectrum, and that the displacement of electrons within the boundaries of one band occurs quite readily. In particular, the energy which an electron acquires over its free path (10^{-4} to 10^{-8} ev) greatly exceeds the energy which separates adjacent levels in a band. It should be noted here, incidentally, that the average thermal energy of a vibrating atom in a solid at room temperature amounts to 0.05 ev, and that the thermal energy of the crystal lattice may be transferred to electrons, which corresponds in the energy diagram to the transition of electrons to higher levels. It follows from these definitions that

Fig. 5. Scheme of energy bands in a solid.

either an external electric field or thermal vibrations of the atoms in a crystal lattice may bring about displacement of electrons within the limits of a given band. Turning to the possibility of transitions between adjacent bands, which, for valence electrons, are separated by forbidden zones the width of which is of the order of one to several electron-volts, we find that the mechanisms of electrical and thermal excitation are to be contrasted. An ordinary (not very strong) electric field cannot impart to an electron, over its mean free path, an energy any greater than thousandths of an electron-volt, while the mechanism of thermal excitation, due to the statistical distribution of the thermal energy of the entire crystal among its individual atoms and the possible resulting fluctuations, is subject to no upper limit for the energy which can be transferred to an electron. Thus the thermal mechanism of excitation can bring about both intraband and interband transitions. Statistical fluctuations in the free path of the electrons in which the energy imparted to an electron by the external electric field proves sufficient for an interband transition play a negligible part by comparison to the thermal fluctuations indicated above. It should be noted here, however, that an electric field can also bring about electron transitions from one band to another if the allowed bands overlap in the same manner as the 3s and 3p bands in Fig. 4. It will be necessary for us to return repeatedly to these problems. The extent to which light quanta (photons) acting upon electrons can give rise to energy transitions may be judged from the fact that the energy of one photon of visible light entirely absorbed by one electron is approximately equal to 2-3 ev. This question, too, will be discussed in greater detail. Let us return to the effect of the external electric field created by an electric current.

Let us emphasize that for the present purpose we draw no distinction between semiconductors, metals and insulators, and refer to the most general case of a solid consisting of atoms (or ions) distributed in an ordered manner.

We already know that an electric current is a directed motion of charged

STAT

particles, produced by the application of an external electric field to the substance. Under the influence of the field, the charged particles (the motion of which in the absence of a field is chaotic) are accelerated in the direction of the field (electrons are accelerated from the cathode to the anode--i.e., against the field) and experience an increase in energy over the region of their free path. On collision with the scattering centers at the end of the free-path region, the electrons transfer the energy which they have accumulated in the electric field to the lattice in the form of heat. Such is the microscopic pattern of the heat effect of an electric current. Should the solid represent an ideal crystal, the wave properties of electrons, according to quantum mechanics, would assure their unobstructed motion within such a crystal. Only deviations from strict periodicity as a result of thermal vibrations of atoms, the various lattice defects, and impurities cause the electrons to move in a zig-zag fashion in a solid, abandoning their rectilinear course as do the molecules of a gas. It follows from this physical scheme that the existence of an electric current is a result of a process of acceleration of electrons over their free path by the external field, to which the displacement of electrons across adjacent levels of an allowed band corresponds in the energy-band diagram; the reason for the inability of an electron to transfer into the next superior band in an ordinary moderate field(<100v/cm) has already been indicated above. These energy considerations determine a necessary, but insufficient condition for the actual transition of an electron. Since each quantum level accommodates only two electrons, the electron energy transition further requires that the energy levels to which the electron can be transferred by the field be vacant or not fully occupied. This second condition relates both to transitions within the boundaries of one band corresponding to the participation of electrons in the electric current, and to the energy transition of electrons across a forbidden band from one allowed band to another. The last type of transition, which can be brought about by the action of heat, light and strong electric fields, plays a basic and decisive role

STAT

in determining the electrical properties of the group of solids called semiconductors.

Thus the question of whether an electric current will occur in a solid under the influence of an electric field is resolved depending on the structure of the energy bands and on the extent to which these are occupied by electrons. In the same manner in which a quantum shell of an atom may be filled (for example, in the case of inert gases) or unfilled (for example, in the case of the alkali metals and halogens), and indeed, as a result of this, an allowed energy band of a solid may be entirely filled or partially filled. If a band is not completely filled, its electrons are able to participate in electric current; in the opposite case, they cannot do so.

In the majority of cases, but not always, the situation is as follows: electrons which form completed groups in the free atoms create, during the process of the formation of a solid from these atoms, energy bands all levels of which are filled. Examples of this are the atomic lattices of helium, neon, argon and others. In a solid, the zones of excitation into which electrons may be ejected from adjacent zones correspond to the excited levels of the free atom which are not filled in the normal state.

As an example of a substance with an unfilled upper normal band let us consider the alkali metal sodium. We know that in a sodium atom, the first quantum shell, comprising 2 electrons, and the second quantum shell, comprising 8 electrons, are completely filled; the outer-shell eleventh valence electron of the sodium atom occupies the 3s level, which it only half fills. Upon formation of solid sodium, the valence electrons are "collectivized"--i.e., they now no longer belong to individual atoms of solid sodium but to the entire lattice, and form an s-band which is likewise only half-filled. In solid sodium, the completely occupied bands which unify all electrons of the same atomic state correspond to the completed first and second quantum shells of the free sodium atom. According to what has been said

STAT

above, an external electric field can cause electron displacements only in the unfilled valence band, and one electron per atom will participate in the current created in the sodium; the remaining electrons of the sodium atoms cannot participate in the current, since all levels in the bands in which they are situated are occupied. Naturally, these electrons may be ejected into a superior, unfilled band under the action of heat and light and, having attained it, are able to participate in the electric current. As indicated by calculation, however, the concentration of these excited electrons is always very small compared to normal electron concentration in the valence band of sodium ($\sim 10^{22}$ cm^{-3}). From the standpoint of the general band theory relating to all solid bodies, this is why solid sodium (and the remaining alkali crystals) is a good conductor of electric current--i.e., is a metal in which the concentration of current carriers is independent of temperature.

As an example of the contrary case--i.e., that of a solid in which no partially filled bands exist, let us consider the rock-salt crystal NaCl. In the solid rock salt, the valence s-electron of the sodium atom transfers to the chlorine atom, in which the external 3p-subshell, which contains five electrons, is one short of the number required for completion. The transition of one electron from sodium to chlorine is conditioned by an energy advantage, since the energy of the crystal diminishes therewith. With the transition of one electron from sodium to chlorine, the outer shell of the latter, which has now become a Cl^- ion, is entirely filled, just as is the case with argon--the closest neighbor of chlorine. Another consequence of this transition is the loss of the outer sodium shell, as a consequence of which the sodium becomes a Na^+ ion which henceforth possesses the same outer shell as the inert neon. When a crystal is formed from Na^+ and Cl^- ions bound together by normal electrostatic forces of attraction, the energy band which embraces all external electrons of the Cl^- ions is completely filled. All bands below it are also entirely filled, while the 3sNa band of states above it is completely void in this case (see Fig. 6).

The conditions set forth above for the occurrence of an electric current make

STAT

it clear that if no transitions from the filled to the void zone are excited, such a crystal will be a nonconductor. Experiments show that conductivity may be observed in crystals of this type under certain conditions.

It follows from the theory which we have presented that the ejection of electrons into the unfilled band which, for this reason, is called the conduction band, is the condition for the occurrence of conductivity. The band of the neutral state of sodium atoms--3sNa--is this band in the NaCl crystal. To permit an electron to transfer from the valence band 3pCl into this band it is necessary to impart to it an energy equal to the width of the forbidden region located between them (~6 ev).

This electron transition, and the occurrence of conductivity which results from it, will admittedly be more comprehensible if presented in the following way: In a lattice consisting of Na^+ and Cl^- ions, the electron returns to its original place, making the transfer from Cl^- to Na^+, under the influence of an appropriate exciting agency (in this case, ultraviolet light).

Fig. 6. Scheme of upper energy bands in NaCl crystal.

Consequently, two neutral Na and Cl atoms appear simultaneously in the lattice. Yet the neutral state of these atoms, which are surrounded by ions, does not remain localized; instead it migrates in a random manner through the crystal, transferring from neutral atoms to ions. These neutral-state migrations are accomplished by transition of electrons from neutral sodium atoms to any adjacent sodium ions, and by transition of electrons from any chlorine ions adjacent to neutral chlorine to these neutral chlorine atoms. The physical principles which facilitate the electron transitions between atoms or ions of a dielectric medium will be discussed specifically below. Let us turn our attention to the fact that the initial transition of one electron created two chains of states which migrated through space in a random manner: one along a line of Na^+ ions, and another along a line of Cl^- ions. Upon the application of an external electric field, these random migrations acquire a certain directional character, since the electrons, which effect these migrations by jumping from neutral atoms to ions in a manner comparable to a relay race, begin to move preferentially in the direction opposed to that of the field. The neutral state of sodium brought about by transition of an electron from some neutral atom to the closest Na^+ ion moves preferentially from the anode toward the cathode, by which it is absorbed. Similarly, the neutral state of chlorine, which is brought about by the above-described "relay" transition of an electron from any chlorine ion among those grouped around a neutral atom to the latter, moves preferentially from the cathode to the anode, and is absorbed by the latter. In the latter case, instead of speaking of the transition of an electron from any one of the chlorine ions surrounding it to some neutral chlorine atom which contains one electron less than the surrounding chlorine ions, one might speak of the transition of an electron deficiency--which may be referred to as a hole--from the neutral chlorine atom in question to an adjacent chlorine ion. In this case, the transition process along the chain of chlorine ions will be the same as that along the sodium chain, with the sole difference that the electrons which bring about the migration of neutral sodium states move from the cathode toward the anode, while the holes, which accomplish the migration of the neutral chlorine states, move in the opposite direction. This circumstance enables us to describe the hole-migration effect as the motion of positive charges. Both processes brought about by electron migration in a weak electric field will be registered in a closed electrical circuit containing a light-excited fragment of rock salt as a weak electric current which does not result in electrolytic decomposition of the substance.

Thus, a relay motion of numerous electrons occurs in a NaCl crystal upon the excitation of a single electron and its transfer from a Cl^- ion to a Na^+ ion. This motion may be described as simultaneous motions of one negatively charged electron

[redacted]

and one positively charged hole, with the motion of each of these charges occurring only along routes which pass through identical lattice points, and not in random directions within the crystal. The conducting electron moves along the points occupied by sodium ions, and the hole along those occupied by chlorine ions. We shall return to these significant conclusions in our further discussion. Let us note for now that it is directly related to the problem of the nature of the positive conductivity observed in experimental studies of the Hall effect, electromotive force, etc., although the comparatively simple classical pattern of crystal conductivity presented can give no definitive explanation of the existence of two types of conductivity--negative and positive.

For better clarification of the general physical pattern of the electrical properties of solids, one should explain these matters for the case of crystals of elements of the alkaline earth group: Be, Mg, Ca, etc., whose free atoms possess entirely filled valence s-shells. On the basis of the general qualitative relationships set forth above, we should have expected solid substances composed of such atoms to behave as nonconductors, since it would seem that the energy bands of such substances should be entirely filled. In actual fact, however, these substances are typical metals. The band theory of solids explains this instance in the following manner.

It was shown above that the excited energy levels of free atoms, which are vacant when the atoms are in the normal state, create corresponding energy bands in a solid consisting of such atoms. However, while the excited levels are necessarily separated from the normal levels by forbidden bands in an isolated atom and located above the normal levels in the energy diagram, the band created in the solid due to the splitting off of the first excited level can superpose itself upon the valence band in such a way that a forbidden region no longer exists between them. The situation thus described is represented in Fig. 4, which shows the imbrication of the 2s-2p and 3s-3p bands. The figure shows that certain s-band levels prove

STAT

(17)

to be located lower than some of the p-band levels, so that in the normal, unexcited state the valence electrons will distribute themselves in the lower section of the expanded, combined (s + p)-band without filling it entirely. But once this upper hybrid band, which contains valence electrons, is partially occupied, an external electric field may cause displacement of electrons across the vacant levels of this band, and an electric current may arise in the solid body. Thus the metallic properties of alkaline-earth crystals are explained by the overlapping of energy bands.

However, the order of energy band distribution in a solid substance does not always correspond to the order of distribution of the corresponding energy levels in the free atom. A band created by the splitting of a lower atomic level may appear as a higher one in the energy spectrum of the solid, so that a vacant band in the solid may correspond to levels filled by electrons in the free atom. Tellurium (see Fig. 7), in whose free atoms four external electrons occupy the degenerate 5p level, may serve to provide an example of such a situation. The full capacity of the p-level in the atoms is 6 electrons, as we already know. Should this level, remaining degenerate as the association of tellurium atoms progresses, form one band in the tellurium crystal, the latter would be a metal with an unfilled 5p-band. However, conditions in solid tellurium are such that the degeneracy of the energy bands is eliminated by the influence of the electric field of the crystal: the threefold-degenerate 5p-band is split into three bands, one of which moves sharply upward in the energy diagram and, with the nominal parameters of the tellurium crystal lattice, assumes a position even higher than that of the excited--(i.e., conducting) state band which proceeds from the 5d-level. Thus this band (cross-hatched in Fig. 7) proves to be entirely vacant in solid tellurium. The remaining two bands, both of which are genetically related to the 5p-level, are populated by a full complement of electrons--four per atom--and separated from the nearest free band--the 5d conduction band--by a forbidden interval $\sim$0.32 ev in width. Thus does theory explain the experimentally established fact that solid tellurium is a

STAT

semiconductor and not a metal.

In some instances, it proves quite impossible to obtain positive identification of an energy band of a solid with any definite atomic level. This is the case, for instance, with the diamond lattice (see Fig. 8). In free carbon atoms, four valence electrons are distributed in pairs at the 2s and 2p levels, which have a total capacity of 8 electrons. With the association of carbon atoms in the configuration which corresponds to the diamond lattice, the 2s and 2p atomic levels first expand into individual bands with capacities of 2 and 6 electrons; as the atoms draw still closer together, a single united band with a capacity of 8 electrons per atom is formed, and finally, when the interatomic distances have diminished to the point at which their values now correspond to the nominal parameters of the diamond lattice, this united band splits into two bands, each with a capacity of 4 electrons per atom. Since these two bands are separated by a wide forbidden region (~6 ev) in the diamond lattice, the four valence electrons of carbon distribute themselves in the lower band, which they complete, and the diamond crystal proves to be nonconductive. A pattern similar to the one described occurs in the germanium and silicon lattices, but since the energy gap separating the split-off (s + p) bands in germanium and silicon is of the order of 1 ev, these substances prove to be semiconductors.

Thus we see that the valence bonds between the atoms in atomic lattices are strong enough to divest the energy spectrum of the valence electrons of the crystal of all resemblance to the spectrum of the same electrons in the free atoms.

The examples which we have discussed enable us to summarize the basic concepts of the band theory which explain, from a consistent standpoint, the experimentally observed wide diversity in the electrical properties of solid substances.

1. A solid substance represents a single system in which the energy states of electrons form practically continuous bands separated from one another by intervals of forbidden energy values--the forbidden bands. Only the valence band and the

STAT

first excited-state band, which is situated above it, participate in the process of conduction in solids. The energy ranges of these bands are of the order of one to several electron-volts. It will be recalled that the energy unit 1 ev is the basic measure of energy in the fields of atomic and electronic phenomena, and is equal to $1.6 \cdot 10^{-12}$ erg.

Fig. 7. Formation of energy bands from atomic levels in tellurium.

The energy gap--i.e., the width of the forbidden bands which we encounter in investigating the conductivities of solid substances, varies from 8 ev to zero, and may even be negative. This last possibility relates to the case in which the allowed energy bands overlap in such a manner as to place the lower boundary of the upper band (bottom of the band) below the upper boundary of the underlying band.

2. In a solid substance, an electric current may arise at any temperature provided that the valence band or the

Fig. 8. Formation of energy bands from atomic levels in diamond.

combined zone formed by the overlapping bands is not entirely filled with electrons. In this case, energy transitions occur freely as a result of acceleration of electrons in the external electric field, and the conductivity of such a solid is metallic in nature.

3. If the valence band of a solid with an ideally ordered lattice is entirely filled and the overlying vacant band of excited states is separated from it by a forbidden region of finite width, such a solid will be a perfect insulator at absolute zero, in absence of irradiation and in a none too strong electric field. With rising temperature or under luminous radiation, electrons will be ejected from the valence band into the overlying band of excited states, and, having attained the latter, will be in a position to participate in conductivity; hence this band is called the conduction band. As distinguished from a metal, for which the state of conductivity is its normal state and does not require any electron transitions into the overlying band, the state of conductivity of a solid substance with an entirely filled band is an excited state.

4. The width of the forbidden band varies with temperature for two reasons: 1) the amplitude of the thermal vibrations of an atomic lattice varies with temperature, and 2) a change in temperature causes a change in the volume of the substance, or, what is the same thing, changes in its interatomic distances. The first effect always results in a narrowing of the forbidden band with rising temperature; the second may lead either to narrowing or widening of this zone with rising temperature, depending on the shape of curves in the specific energy-band diagram. Figure 9 illustrates the influence of the second effect --the variation in the width of the forbidden band of a crystal with variation in the crystal-lattice constant: the curves which correspond to the limits of the valence band of conductivity and thereby determine the width of the forbidden band may converge as the interatomic distances diminish and then diverge as these distances diminish further; if the lattice constant in the normal state of a crystal is $\underline{a}_2$, a decline in interatomic distances

causes the width of the forbidden band ΔE_2 to diminish; and if the normal state of a crystal is characterized by the lattice constant $\underline{a}_1$, the drawing together of the atoms results in widening of the forbidden band.

As a result of superposition of the two above-mentioned effects the width of the forbidden band increases with rising temperature in some semiconductors, while in others it declines. Thus the influence of pressure, which affects only the interatomic distances, also results on occasion in widening of the forbidden band and sometimes in its becoming narrower.

5. When the mechanism which excites conductivity is of a thermal nature, the number of electrons ejected from the valence band into the conduction band increases very rapidly with rising temperature. On the other hand, the reverse processes of electron recombination, which consist in the return of electrons from the conduction band into the valence band, also occur in a solid simultaneously with the processes of their excitation. This means that the state of conductivity in each current carrier is preserved only for a certain period of time, which is called its lifetime τ_0. The maximum value of this quantity (which depends to a great degree on the perfection of the crystal lattice) has been observed in Ge and Si to be 10^{-3} sec,

Fig. 9. Effect of changes in interatomic distances on width of forbidden band.

and is much smaller in other semiconductors: for PbS, 10^{-5} sec; for InSb, 10^{-7} sec. The intensity of the recombination processes is proportional both to the number of excited electrons and to the number of vacancies left behind them, and, when these numbers are equal, increases as the square of the number of excited electrons. A state of equilibrium is established at any given temperature of the substance as a result of the competition between the excitation and recombination processes. The concentration of excited electrons in this state is equal to:

$$\underline{n}_1 = \frac{2(2\pi\sqrt{\underline{m}_n \underline{m}_p}\,\underline{k}\underline{T})^{3.2}}{\underline{h}^3}\,\underline{e}^{-\frac{\Delta \underline{E}_0}{2\underline{k}\underline{T}}} \qquad (4)$$

In this formula, ΔE_0 is the width of the forbidden band, $\underline{T}$ is the absolute temperature of the substance, $\underline{k}$ is the Boltzmann constant $= 1.38 \cdot 10^{-16}$ erg/degree, $\underline{h}$ is the Planck constant $= 6.62 \cdot 10^{-27}$ erg sec, and $\underline{m}$ is the effective mass of the current carriers. The significance of the subscripts $\underline{n}$ and $\underline{p}$ to the effective-mass symbols will be explained below in the discussion of the nature of positive conductivity (see page 44).

According to Formula (4), the ejection of electrons into the band of conductivity, which requires a definite expenditure of energy, may occur at any temperature of the substance--even extremely low ones. In this connection, it might not be amiss to clarify this circumstance in more detail, and incidentally, to bring out the reason for the exponential dependency of concentration on temperature. The average thermal energy of a vibrating lattice atom clearly depends on the temperature of the substance, and, at a certain, sufficiently low temperature, may become much smaller than the energy $\Delta \underline{E}_0$ which must necessarily be consumed for a single electron-excitation event. In this case, why raise the question whether excitation processes occur at any temperature? The explanation of this circumstance is related to the statistical nature of the distribution of thermal energy among the atoms of the substance. Energy fluctuations--i.e., deviations from the average energy value--are a consequence of this statistical character of the energy distribution. This manifests itself in the fact that the thermal energy of individual atoms may, in the course of a certain period of time, be any number of times larger or smaller than its average value. According to classical statistics, the number of atoms whose energy at a given moment exceeds the value ΔE_0 or, in other words, the number of atoms of the solid which are capable of ejecting an electron into the conduction band at the expense of their own thermal energy, is proportional to $e^{-E_0/kT}$ (when $E_0 > kT$). This relationship indicates that at any temperature there exists a certain number of atoms with energies sufficient for one electron-excitation event, and that the number of these atoms, and consequently the number of such unit excitations as well, must increase exponentially with rising temperature.

The process of light-induced excitation, or, in more general terms, electromagnetic excitation, is of a different nature. In this instance, electrons appear in the conduction band only when the energy of an individual light quantum (photon) h is equal to or exceeds the excitation energy ΔE_0 --i.e., $h\nu \geq \Delta E_0$. The same condition applies to electron excitation by various forms of corpuscular radiation such as electrons, protons, neutrons and ions. If the energy of the electromagnetic quantum or bombarding particle is smaller than the width of the forbidden band of a crystal, the transfer of electrons from the valence band to the conduction band will not occur at all.

6. From the viewpoint of the band theory of solids, the classification of nonmetallic substances as semiconductors or insulators is to be considered as purely conditional, since it is not based on any qualitative physical characteristics, and the theory regards the only difference between them as the width of the forbidden band. It is not possible to establish a limiting value for the width of a forbidden band by which the class of semiconductors is separated from that of insulators; we can do no more than indicate that the width of the forbidden band does not exceed 2 ev in the solid substances considered as semiconductors in modern practice. The state of conductivity is the normal unexcited state in metals, and these (and the

(4.

semimetals) are qualitatively different from the semiconductors and insulators.

The classification of non-metallic solid substances as semiconductors or insulators is sometimes based on their specific resistance at room temperature, with $\rho = 10^{10}$ ohm·cm taken as a conditional boundary for this classification.

The difference between the electrical properties of metals and non-metals should become particularly apparent with the approach to absolute zero, while the difference between the two nonmetallic groups--semiconductors and insulators--should become smaller under the same conditions.

Thus a consistent theoretical interpretation is achieved for the electrical properties of all solids: the excitation energy of the current carriers, or the so-called activation energy of conductivity, equals zero in metals, and increases continuously in the series of semiconductors which, as this energy increases, is continuous at some arbitrary point with the series of insulators; the good conductor metals and the dielectrics with good insulating properties are the extreme cases of this continuous series, into which all solids may be placed on the basis of this criterion.

7. Finally, the last important conclusion of the contemporary theory of conductivity of solids which we shall note here is that due to the wave nature of the electron, its motion in an ideal crystal lattice is entirely unimpeded.

Such a conclusion is astonishing from the viewpoint of the classical theory, since a charged particle moving within the strong electric field of a crystal should, according to this theory, experience deflections so strong in its passage by each atom as to render its average velocity equal to zero, even over microscopic segments of its path. Here, a question may arise as to whether the unimpeded motion of electrons within an ideal crystal should not, according to the quantum theory, give rise to a state of current in this crystal, even when an external electric field does not exist. The quantum theory answers this question in the negative. In this instance, the absence of an electric current may be explained by symmetrical motions of the

STAT

current carriers in opposite directions. The quantum theory leads to the conclusion that disturbance of this symmetrical motion by an external field is possible only in cases in which an energy band is not entirely filled. The current carriers, accelerated by the external field, transfer into the higher, unoccupied levels of this band, and the electric current would pass through the crystal without experiencing any ohmic resistance. The latter arises as a consequence of the scattering of current carriers at the disturbances of the strict periodicity of the crystal which are created by thermal vibrations of the atoms, by all types of lattice defects and by admixtures of foreign elements (contamination). The scattering events are inelastic, and the energy accumulated by a current carrier along its free path in the external electric field is hereby transmitted to the lattice in the form of heat.

Despite the fact that ideally ordered crystals do not exist, and, consequently, the condition indicated above in which the specific resistance of the solid declines to zero is practically never encountered, some metals exhibit a state of superconductivity in which $\rho = 0$ at sufficiently low temperatures despite the fact that the crystalline structure of many of them is obviously disordered. The nature of this physical phenomenon, which was first observed half a century ago, is not yet explained.

Since the activation energy of conductivity in all metals equals zero, the differences in their electrical properties are determined by the value of specific resistance, which is a characteristic constant of the materials and undergoes relatively slight changes upon variation of temperature, under radiation, and upon the introduction of small amounts of impurities. The influence of temperature, irradiation and slight contamination on the electrical properties of semiconductor materials is so strong that the specific electrical resistance cannot serve as a unique characteristic of a real semiconductor. Therefore it is not the value of the specific resistance ρ, but the activation energy of its intrinsic conductivity,

--i.e., the width of the forbidden band ΔE_0, which is mentioned first when describing the electrical properties of practically pure semiconductors of specific chemical composition. Numerical values of this quantity are given in the large circles in the table presented in Fig. 1 for the various elementary semiconductors. Here we should make the reservation that the electrical properties of a real and not very pure semiconductor are determined in a certain temperature range by the excitation energy of the foreign current carriers and not by the activation energy of its intrinsic conductivity. However, to avoid digression from the order of this exposition, we shall postpone treatment of this question until our formal discussion of the roles of intrinsic and impurity conductivity. In the meantime we shall attempt to clarify another important question concerning the nature of the positive electronic conductivity which is observed in experiments.

We already know that electrons cannot participate in the conductivity of a substance if all levels of the valence band are filled. We know further that under the influence of heat and radiant energy, these electrons of the valence band may be ejected into the conduction band, and since the latter is usually occupied by them only to a slight degree, these electrons, migrating freely across the unoccupied levels of the conduction band, are able to participate in conductivity. This electronic motion normally manifests itself in experiments in the form of negative conductivity, which may be inferred from the signs of both the Hall effect and the thermoelectromotive force.

Let us now turn our attention to what occurs in the valence band when some of the electrons occupying it transfer into the conduction band, thereby releasing energy levels in the valence band. The occurrence of unoccupied levels in this previously filled band eliminates the reason for the nonparticipation of this band in conductivity, permitting electrons in the valence band which have not transferred to the conduction band to participate in the electric current too by migrating through the levels rendered vacant in the upper valence-band region. It is appropriate to recall here that in our graphic presentation of the results of transition of <u>one</u> electron from a Cl^- ion to a Na^+ ion, (see page 30) we came to the conclusion that two electron-migration chains are formed in this process in the NaCl crystal, and that these chains may be described as the simultaneous motion of <u>two</u> charges, one of which is negative and the other positive. The quantum theory of solids indicates that, while the behavior of electrons close to the bottom of the conduction band can be described approximately as the motion of negatively charged particles of an ideal gas, the mass of which is positive and, in the general case, different from the mass of a free electron, the behavior of the electrons straying through the levels in the upper region of the band (any band, and not the valence band exclusively) is fundamentally different from that of electrons at the bottom of the band. This fundamental difference consists in the fact that electrons migrating in the upper part of the band are accelerated by the electric field in the direction opposite to the direction of acceleration of the electrons migrating in the lower part of the band. In attempting to interpret this situation purely logically with the aid of the concepts of classical physics, we should be forced to ascribe a negative mass to these anomalous electrons. No such paradox occurs in the contemporary electronic theory of crystals, since the negative sign is ascribed not to the true electron mass but to its effective mass which, by its definition in the theory, may differ from the true mass both in its magnitude and in sign. The introduction of the concept of the effective mass of a current carrier enables us to describe its motion within a crystal as the free motion of a charged particle without taking the periodic field of the crystal into account in the equation of motion.

In order to provide a descriptive explanation of the method used by the theory in this particular case, let us consider the motion in a gravitational field of a material body occupying an ideal, nonviscous fluid. Simplifying the complex pattern of the motion of the fluid which arises in this case, we may write the equation of motion of the body, which has a mass <u>m</u>, as follows:

$$m \frac{d^2 s}{dt^2} = mg - V\rho_{fl}\, g,$$

where the first term on the right side of the equation represents the force of gravity, and the second the Archimedean buoyant force.

In order to simplify this equation, we may eliminate the Archimedean force from it on the assumption that the value of the effective mass m^* of a solid is different from m. Then the initial expression may be written

$$m \frac{d^2 s}{dt^2} = mg - \frac{m}{\rho_s}\rho_{fl}\, g = mg \left(1 - \frac{\rho_{fl}}{\rho_s}\right)$$

$$\frac{m}{1 - \frac{\rho_{fl}}{\rho_s}} \frac{d^2 s}{dt^2} = mg,$$

and the equation of motion of a solid in a liquid expressed in its most simple form, which does not explicitly account for the Archimedean force:

$$m^* \frac{d^2 s}{dt^2} = mg,$$

where

$$m^* = \frac{m}{1 - \frac{\rho_{fl}}{\rho_s}}$$

It is seen from the last expression that the effective mass m^* which we have thus introduced may differ from the true mass m not only in magnitude but in sign as well. Indeed, with $\rho_{fl} > \rho_s$, the effective mass m^* will be negative. The physical significance of this result is clear: in consequence of the fact that the Archimedean buoyant force--which is accounted for in the effective mass in this particular instance--exceeds the force of gravity, the direction of acceleration of the solid body (↑) is opposed to that of the single force which figures in the simplified equation--that is, the force of gravity (↓).

On comparison of the motion of a solid body in a fluid with the motion which interests us--that of electrons in a solid--it is necessary to assume that the part

[redacted]

STAT

corresponding in the crystal to that of the Archimedean force is played by the internal electrical field of the atoms, whose influence on the current carriers may significantly exceed that of the external field applied to the conductor. The fact that a solid floats in a liquid with $\rho_{fl} > \rho_s$ when the direction of acceleration of the solid and that of the force of gravity are opposed to one another, has as its counterpart in the crystal the anomalous behavior of the electrons in the upper part of a band, which are accelerated by the external electric field, in a direction opposite to that in which the electric forces act.

Naturally, it should be kept in mind that the example considered here is no more than a descriptive presentation of the method used in describing the motion to permit the effective mass of a solid to assume negative values, and that it is not possible to draw any physical conclusions concerning the behavior of electrons in a crystal lattice by simple extensions of the analysis of this mechanical analogy.

Theory indicates that on the appearance of a single unoccupied level in the upper part of an almost filled energy band, a collective motion of electrons anomalously accelerated by the external field arises within this band. This motion is equivalent to that of one positive charge accelerated by the external field in the normal manner i.e, in the direction of the field. The magnitude of this positive charge is exactly equal in absolute value to the electronic charge; the positive effective mass of this charge, which determines its acceleration in electric and magnetic fields, differs, as experiment indicates, from the mass of a free electron--generally in the same manner as the effective mass of an electron situated in a thinly populated conduction band. This positive charge, which determines the effects of formation of one unoccupied level in an almost filled band, is called a hole. Upon the application of a magnetic field, these holes, which move in the direction opposed to that of the electrons situated near the bottom of the band under the influence of an electric field, will be deflected by the magnetic field in

STAT

the same direction as the electrons, but will nevertheless show an abnormal Hall potential difference corresponding to positive conductivity in the specimen.

The same conclusion applies to the motion in a thermal field as well: the sign of the electromotive force will differ in these two cases. If the electrons situated in the lower part of an energy band participate in a thermal current from the hot end of a solid toward its cold end, the latter will acquire a negative charge as a result of their accumulation there; these electrons will also create an electric field directed from the hot to the cold end. The migration of holes situated in the upper region of a band in a thermal current in a nonuniformly heated solid gives rise to an electric field in the opposite direction.

It should be clear from the above that the collectivized electrons of a crystal, which give rise to electronic conductivity within it, may not be considered capable of changing the sign of their charge or of their true mass depending on their positions in the energy band. Such a paradoxical conclusion would contradict our customary concepts of electrons; if, nevertheless, one speaks of a negative effective mass or a positive electrical conductivity, these terms should be understood merely as part of a conventional description, in terms of the customary classical concepts, of the extremely complex motion of electrons in the strong periodic field of a crystal. It is not the electrons themselves that are anomalous, but their behavior which seems to be so when we endeavor to present it as analogous to the behavior of free atoms in a vacuum. We should be surprised not at our helplessness in the attempt to represent highly complicated phenomena with the help of simple customary concepts, but at the potency of the mathematical apparatus of the quantum theory which enables us to explain, on the basis of the known general properties of the electron, experimentally established facts which are incomprehensible within the framework of classical physics.

Thus the explanation of the existence of two types of electron conductivity proceeds from those qualitative characteristics of the motion of electrons in the periodic field of a crystal which manifest themselves in different ways depending on the position of the energy levels of the electrons in the band. The immediate physical causes of the paradoxical behavior of electrons in a solid are the wave nature of electrons and the enormous magnitude of the periodic electric fields which exist in the vicinity of the atoms--fields millions of times stronger than the external fields which we ordinarily apply to conductors. Actually, even the fact that the energy spectrum of the current carriers in a solid takes the form of alternating ranges of allowed energy values separated by forbidden bands is also a consequence of the wave nature of electrons and the periodic structure of the electric field of the crystal.

The above rule according to which the experimentally established sign of the current-carrier charge depends on the position of their energy levels in a band applies equally both to semiconductors and to metals, in which the Hall effect also reveals conduction of both signs. When energy bands overlap, the conduction electrons of a metal may occupy both the upper levels of one band, in which case their collective motion is equivalent to the motion of positively charged holes, and the lower levels of the next band, in which case they will behave as normal negative electrons.

Thus we have ascertained that upon the ejection of one electron from the valence band to the conduction band under the influence of an external field, a collective electron displacement begins within the migrating vacated level of the former. This collective motion in the valence band of negatively charged electrons accelerated in an anomalous manner by the external field is equivalent to the motion in this external field of one positive charge with a positive mass. It is the effective mass of this positive charge migrating in the valence band which is the $\underline{m}_p$ appearing in Formula (4). The conductivity due to the motion of these hole-charges is positive in sign, referred to as the positive or hole conductivity and denoted by the symbol $\underline{p}$ in contrast to the negative or electron conductivity which is denoted by the symbol n.

As follows from the physical pattern presented here, each electron which is transferred by excitation into the conduction band leaves behind it a hole in the valence band, so that the number of positive holes in the valence band in this excitation scheme will be exactly equal to the number of electrons in the conduction band--i.e., it should also be possible to determine the concentration p of the positive current carriers by means of Formula (4). This circumstance, which we have already encountered on page 32 in discussing the electron conductivity of the NaCl crystal, explains why the two masses m_n and m_p, which in the majority of cases are unequal, appear in Formula (4).

Also of essential importance is another conclusion which was noted in the same place: namely, that if the particles which form the crystal lattice of a solid, whether they differ as to their chemical nature or are identical in this respect, do not occupy identical positions in the lattice, the motion of any species of current carrier belonging to a given energy band occurs not in random crystallographic directions but only along lines which pass through structurally equivalent lattice points which are occupied by identical particles. This conclusion is linked to the fact that the free motion of current in a crystal occurs as a consequence of the wave-mechanical tunnel effect which enables electrons to cross the high potential barrier which separates the atoms in a lattice without a change in energy, which is contrary to classical theory. Such a passage is accomplished only when the adjacent atoms have isoenergetic systems of energy levels. This condition requires that the atoms (or ions) be not only chemically identical but also that they occupy structurally identical positions in the lattice.

Since two current carriers are created simultaneously in each event of excitation from the valence band to the conduction band, the total conductivity of such a substance should be given by the formula

$$\sigma = en_1u_n + ep_1u_p \qquad (5)$$

Here, u_n and u_p are the mobilities of the electrons and the holes, which differ in their values due to the difference in their effective masses and free times (see page 13). In most, though not all cases, the effective mass of the electrons is smaller than that of the holes, so that the electronic mobility will in these cases exceed the mobility of the holes. The explanation of this circumstance in the theory is that the effective mass of a current carrier is inversely proportional to the width of the band in which the current carrier is situated, and since the width of the allowed bands generally increases with increasing energy, as shown in Fig. 5, the conduction band is usually wider than the valence band. Deviations from this rule are due to the splitting-off and overlapping phenomena which are possible for all bands. It follows from Formula (5) that in the general case, despite the equality of n_i to p_i, the positive and negative components of the electric current will not be equal in the excitation mechanism described, and that the conductivity will have a negative sign when electron mobility is high. The presence of current carriers of two signs complicates the theory of all the electric phenomena, and, in particular, the formula for the Hall effect (3), which applies when the current carriers have but one sign, is replaced in the general case by the more complex compression

$$R_H = A \frac{p_i u_p^2 - n_i u_n^2}{(p_i u_p + n_i u_n)^2}, \qquad (6)$$

where A retains its significance of Formula (3). Despite the equality $n_i = p_i$, the concentration of current carriers cannot be determined from the Hall effect alone. We can, of course, resort again to the equality (5), but it becomes apparent that this, too, is insufficient to determine the three unknowns n_i, u_n and u_p; it is necessary to measure, in addition, some other effect -- galvanomagnetic, thermomagnetic or thermoelectric.

The question may arise as to whether conditions in which the charges of all current carriers in a conductor have the same sign may exist at all. In metals, this occurs when no overlapping of bands occurs; in semiconductors, these conditions are realized for a mechanism of conductivity which differs from the one described above, and which is called the extrinsic conductivity. It is to this question that we will turn now.

CHAPTER III

INTRINSIC AND IMPURITY CONDUCTIVITIES OF SEMICONDUCTORS

In the case discussed above, we assumed a chemically absolutely pure semiconductor whose electrical conductivity can be created only by electrons of intrinsic atoms of the semiconductor substance. However, in order to permit these electrons to participate in conduction, it is necessary to excite them, expending for each excitation event an energy whose magnitude is quite definite for every semiconductor. This quantity of energy, which is equal to the width of the forbidden band ΔE_o, is, indeed, the basic characteristic of the electric properties of a semiconductor. Since two current carriers, the signs of which are opposite, are simultaneously formed in each excitation event in such a substance, the total number of current carriers will be twice as large as the number of electrons in the conduction band -- i.e., $n = n_i + p_i = 2n_i$. In the particular case of a thermal excitation mechanism, n may be determined by the use of Formula (4) for n_i. In the case of a mechanism of excitation by light, the total number of current carriers will be determined by the number of photons absorbed by the substance and individually capable of ejecting an electron from the valence band into the conduction band. The conductivity of such a semiconductor, no matter what the excitation mechanism is, should be determined by Formula (5). The conductivity in a case of this type is called intrinsic conductivity, and the substance in which this conductivity occurs an intrinsic semiconductor. From the following exposition it will become clear that intrinsic conductivity can be observed in a semiconductor only if the quantity of impurities in it is sufficiently small or the temperature is

STAT

sufficiently high, or, finally, in certain instances in which its state of conductivity is excited not by high temperature but by some other source of energy such as light.

The case of intrinsic conductivity considered here is of basic interest to theory, since it permits us to ascertain the basic characteristic physical properties of semiconductors. The performance of most (but not all) types of semiconductor devices is disturbed by the assertion of the intrinsic conductivity of the substance. Hence, the semiconductors most suitable for the fabrication of such devices are those in which the intrinsic conductivity appears at the highest possible temperature -- i.e., semiconductors with wide forbidden bands. In the operating regimes of these semiconductors, their conductivity is effected by the impurity mechanism of conductivity.

As the term suggests, this mechanism of conductivity arises from the presence of impurities in the semiconductor -- the word "impurities" including not only atoms of foreign elements, but also excess (in terms of the stoichiometric composition) atoms of the very same elements which determine the chemical formula of the semiconductor compound; moreover, the effect of impurities is also produced by all types of crystal-lattice defects, including vacancies, atoms and ions implanted in lattice interstices, shears which arise on plastic deformation of the crystal, cracks, etc. In view of this looseness with which the term "impurity" is applied, we often speak not of impurity atoms but of impurity centers.

Exactly what is the role of impurities and how do they effect the electric properties of semiconductors? We already know that the excitation of conductivity in an intrinsic semiconductor requires the transition of an electron from the valence to the conduction band. If the basic crystal lattice of a substance contains impurity atoms, the valence electrons of these atoms, irrespective of their positions in the lattice, often (but not always) distribute themselves not in the energy bands of the basic substance, but on isolated narrow energy levels formed in the forbidden band of the basic diagram. Figs. 10 and 11 present the two most interesting instances.

1. The impurity level is disposed close to the bottom of the conduction band (Fig. 10). In the absence of exciting agencies (heat, light), the valence electrons of all similar impurity atoms are to be found on this narrow level. The physical significance of the "narrowness" of this level consists in location of the impurity atoms at such great distances from one another in the lattice as not to interact to any observable extent, while the electronic energy levels remain narrow as in individual free atoms and the electrons themselves remain localized in their individual atoms -- i.e., they cannot migrate through the crystal and participate in conductivity. If the impurity concentration is so high that the interaction between impurity atoms becomes appreciable, the narrow impurity level expands into an impurity band, and the electrons acquire the ability to migrate through the crystal between impurity atoms and participate in the conductivity of the substance. We shall not dwell here, however, on this insufficiently investigated, interesting case of extrinsic conductivity. In the presence of an energy source (heat, light), the valence electrons of the impurity atoms may be ejected from this level into the conduction band, or, in other words, break away from the impurity atoms and migrate freely through the crystal. The energy ΔE_D necessary for this, which, as seen from Fig. 10, is equal to the distance between the impurity level and the conduction band, may be much smaller than the excitation energy of intrinsic conductivity which is defined in this figure as the distance ΔE_Q between the filled and unoccupied bands of electron conductivity. As a consequence, the number of electrons ejected by thermal excitation from the impurity level into the conduction band will, under certain conditions, greatly exceed the number of intrinsic electrons ejected into the conduction band from the valence band.

4

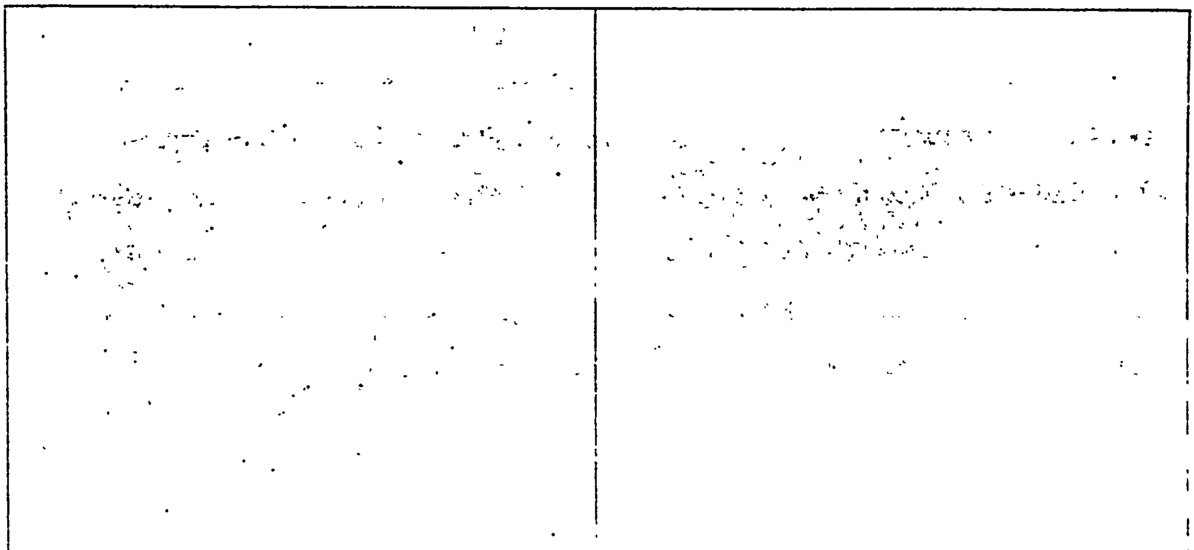

Fig. 10. Energy levels of an electronic semiconductor.

This can be presented more explicitly in the following manner: since the electrons of impurity atoms are bound to their respective atoms less strongly than the electrons of the basic atoms of the lattice, the excitation of impurity electrons due to heating of the substance will exceed the excitation of the basic lattice electrons. The holes which form at impurity atoms located at great distances from one another will remain localized after the removal of electrons -- i.e., they are unable to migrate through the crystal and therefore cannot participate in conductivity.

If the temperature is such that the ejection of electrons from the valence band does not make an appreciable contribution, the electronic conductivity of a semiconductor with such impurities will be negative. In this instance we may use the simple formula for the Hall effect (3). Impurities which contribute electrons to the conduction band are called donors and the energy level of these impurities the donor level.

2. Figure 11 presents a semiconductor in which a local unfilled impurity level lies close to the valence band. In contrast to the previous case, the primary effect of thermal excitation will be to eject electrons from the valence band into this unoccupied impurity level. If $\Delta E_A < \Delta E_0$, and the temperature is not

high enough to cause appreciable ejection of electrons from the valence band into the conduction band, the only current carriers in such a substance will be the holes of the valence band, since the electrons ejected into the impurity level "stick" to it -- i.e., attach themselves to impurity atoms and do not participate in the electric current. The conductivity of such a semiconductor will be purely that of the hole, or positive, type. Impurities which "capture" electrons from the valence band are called acceptors, and their energy level the acceptor level.

Fig. 11. Energy levels of a hole-type semiconductor.

It should be clear from the above that at absolute zero, in total darkness, and with a weak external electric field, an extrinsic semiconductor, just as the intrinsic type, must be an insulator, provided, naturally, that the impurity levels are located neither in the valence band nor the conduction band, and do not themselves create an impurity band. In contradistinction to intrinsic conductivity, extrinsic conductivity is brought about by charges of one sign only: the negative sign in the case of donor impurities, and the positive sign in the case of acceptor impurities. If both donor and acceptor impurities are introduced into the semiconductor, there occurs an impurity compensation which consists in the fact that donor electrons transfer spontaneously to the acceptor level, and the sign of conductivity is determined by the impurity component which is present in excess. In more simple terms, the loosely bound electrons of the donor atoms transfer to

[redacted]

STAT

the acceptor atoms, to which they are more strongly attracted. The necessary character of these electron transitions follows from the general rule that in a state of equilibrium -- i.e., in the absence of external exciting agencies -- electrons must necessarily fill all lower levels of the energy spectrum (in such a way, of course, that no more than two electrons occupy each level), and since the acceptor levels are located below the donor levels, the donor electrons will descend onto the unoccupied acceptor levels.

Complementary energy levels in a forbidden band may be created not only by foreign atoms but by any of the possible lattice defects; the latter may produce levels of either the donor or acceptor type. If the semiconductor is a chemical compound, stoichiometrically superfluous atoms of one of the elements of which the basic lattice is composed, occupying "wrong" positions by virtue of their nature as a stoichiometric excess, may also create complementary levels in a forbidden band. These levels are either of the donor or acceptor type, depending on which chemical component of the compound is present in excess. This explains why any factor which gives rise to complementary levels in a forbidden band is called an impurity, and why the conductivity governed by these levels is referred to as impurity conductivity.

If acceptor impurities are injected into a semiconductor which contains donor impurities, balancing will occur, as a result of which the number of current carriers in the semiconductor will diminish. An analogous process will take place when donor impurities are added to a hole-type semiconductor.

The general expression which determines the number of current carriers in an impurity semiconductor $\underline{n}_I$ is of a very complex form and we cannot present it here, principally because this would involve the explanation of significant new concepts of the quantum theory of conductivity in solids. However, in the two extreme cases which are often realized in practice, this expression assumes a similar form

1) If the width of the forbidden band is sufficiently large that intrinsic

[redacted]

conductivity may be neglected in the entire temperature range which is of interest to us, and if, further, the activation energy of impurities $\Delta E_I > kT$ and the concentration N_I of similar impurities is not too small, the following formula will apply:

$$n_I = \sqrt{N_I}\,\sqrt{\frac{2(2\pi m kT)^{3/2}}{h^3}}\,e^{-\frac{\Delta E_I}{2kT}}\,. \qquad (7)$$

2) If $\Delta E_I \ll kT$, or if, as before, $\Delta E_I > kT$ but the concentration of impurities is small enough,

$$n_I = N_I \qquad (8)$$

-- i.e., in these cases, all of the similar impurity atoms to which these conditions apply are ionized, or, as we say, depleted.

Yet another reservation must be made in connection with Formula (7); this relates to Formula (4) for the intrinsic conductivity. This arises from the degeneracy of the electron gas at higher concentrations, which we mentioned briefly on pp. 10-11. Formulas (4) and (7) relate to nondegenerate states. The critical concentration of current carriers n_c at which degeneration sets in can be determined from the formula

$$n_c = 1.35 \cdot 10^{56}(mT)^{3/2}\,.$$

Assuming $m = m_o = 9 \cdot 10^{-28}$ g, we obtain

$$n_c = 3.6 \cdot 10^{15} T^{3/2}\,.$$

At room temperature ($T = 290^\circ$ K) this gives

$$n_c = 2 \cdot 10^{19} \text{ cm}^{-3}.$$

The concentration of current carriers in those semiconductors with which we deal in practice seldom attains such large values. However, if the concentration approaches the critical concentration, and if the semiconductor is a good conductor being studied at low temperatures, at which the critical concentration diminishes, Formulas (4) and (7) may not be used, and it will be necessary to perform the calculations by a more cumbersome approximate method, which, nevertheless, enables us to obtain a very accurate result.

It follows from the numerical calculations which can be carried out by means of Formulas (4) and (7) that even if the concentration of impurities in the semiconductor is small in the usual chemical sense, but $\Delta E_0 > \Delta E_I$, the bulk of current carriers will be supplied at normal temperature by the impurities and the semiconductor will be of the extrinsic type. With a further purification by modern techniques which, in some instances, enable us to reduce impurities to a millionth part of a per cent, its conductivity will diminish greatly, and the semiconductor will come to be an intrinsic one. By adding the appropriate impurities to such a material, it is possible to produce a semiconductor of any previously specified conductivity, and -- a fact of greatest importance in some cases -- to create in the pre-purified semiconductor regions whose conductivities are of different signs. It becomes clear on comparison of Formulas (4) and (7) why the injection of trace amounts of impurities into a semiconductor might at times alter its conductivity by a factor of several million. Clearly, this will take place in a semiconductor in which $\Delta E_0 \gg \Delta E_I$.

What is the physical reason for the ease with which electrons break away from impurity atoms? After all, the first ionization potential of a free atom of any element does not fall lower than 4 v, and the activation energy of the same atoms in the crystal lattice is often less than 0.1 ev. The physical factor which facilitates the breaking away of an electron is the polarizability of the material medium into which the impurity atom is introduced. This property of the physical medium is characterized by the dielectric constant ϵ. Dielectric polarization of an atomic or ionic medium occurs as a result of the redistribution of electric charges caused by the introduction of an impurity atom into this medium. In the crystal, the polarization process results in a reduction in total energy, while in the impurity atom it leads to the loosening of bonds between the electrons and the nucleus. Therefore the electron orbits of the impurity atom grow larger, and the ionization energy -- i.e., the work required for the removal of an electron from an atom -- diminishes significantly.

Enlargement of the electron orbits of the impurity atoms leads to the formation of the impurity conduction band described on page 54 , even at comparatively small impurity concentrations.

Upon the abrupt enlargement of a valence-electron orbit, the impurity atom may be regarded as a hydrogen-like atom, and the theory of the hydrogen atom may be applied to the determination of the radius of the enlarged orbit as well as of the ionization energy. According to this theory, the radius of the normal orbit of the electron of a hydrogen atom in a dielectric medium is $r = 0.53 \cdot 10^{-8} \epsilon \frac{m_0}{m}$ cm, while the ionization energy is given by

$$E_1 = \frac{13.53}{\epsilon^2} \frac{m}{m_0} \text{ ev} . \qquad (11)$$

For germanium, for which $\epsilon = 16$ and $m = 0.3\, m_0$, we obtain $r = 28 \cdot 10^{-8}$ cm and $E_1 = 0.016$ ev. The latter value may be confirmed directly by experiment since $E_i = \Delta E_I$, and the impurity-activation energy ΔE_I can be measured by several independent methods which we shall discuss below. In numerous instances, the experimental value for germanium is close to 0.015 ev, which is in close agreement with theory. Formula (11) is often used for the determination of the effective mass m of the current carriers, from experimentally known values for ΔE_I and ϵ.

Polarization of a semiconductor medium also directly affects the excitation energy of intrinsic conductivity by facilitating the break-away of an electron from the basic atom of the lattice. In most instances, experimental data confirm that the width of the forbidden band ΔE_0 of a semiconductor diminishes with increasing dielectric constant ϵ. The product $\Delta E_0 \epsilon^2$ is constant for a number of semiconductors, in agreement with Formula (11).

As a concrete example illustrating the above, let us consider germanium, whose forbidden-band width $\Delta E_0 = 0.75$ ev. At room temperature, according to Formula (4), the number of its intrinsic current carriers, the electrons and holes, is $n = 2 \cdot 10^{13} \text{cm}^{-3}$. Since the activation energy of many chemical impurities in germanium does not exceed 0.015 ev, at room temperature, when the value of $2kT = 0.05$ ev, all impurity atoms are ionized at the expense of the thermal energy of the substance, and the concentration of impurity current carriers -- according to Formula (8) -- amounts to $n_I = N_I$. Thus at a concentration of impurity atoms of only 0.001% -- i.e., with $N_I = 5 \cdot 10^{17} \text{cm}^{-3}$, the concentration of impurity current carriers $n_I = 5 \cdot 10^{17} \text{cm}^{-3}$ exceeds the concentration of the intrinsic current carriers by a factor of 25,000! Another conclusion follows from this: namely, that in order for the intrinsic conductivity of germanium to become apparent at all at room temperature, it must be purified to such an extent that its concentration of impurities does not exceed 10^{13}cm^{-3}, which amounts to one ten millionth of a per cent! This condition may be expressed in other terms: no more than one foreign atom may be present per one billion germanium atoms! The most astonishing fact is that such a degree of germanium purification has been accomplished at the present time and that, moreover, this method of purification has been adopted by industry which produces germanium of an even higher degree of purity. At room temperature, specimens of such germanium show intrinsic conductivity, and a specific resistance $\rho = 50$ to 60 ohm-cm. The concentration of intrinsic current carriers decreases rapidly in germanium with falling temperature, in accordance with Formula (4), while the concentration of impurity current carriers n_I remains constant down to very low temperatures, as a result of the small value of $\Delta E_I = 0.015$ ev. Therefore, a specimen of pure germanium which exhibits intrinsic conductivity at room temperature becomes at low temperatures an impurity semiconductor whose current carriers are of one sign. The temperature of transition from intrinsic to extrinsic conductivity depends on impurity concentration: the lower this concentration, the lower the transition temperature (an illustration of this rule is shown in Fig. 18). Naturally, the instances of purely extrinsic and purely intrinsic conductivity are nothing but ideal extreme cases, since every real pure semiconductor always contains a certain amount of impurities, and the excitation of impurity current carriers in an extrinsic semiconductor is always accompanied by simultaneous excitation of its intrinsic current carriers. This fact is not essential for numerous problems which arise in practice, and the influence of a small fraction of intrinsic current carriers in an extrinsic semiconductor may be neglected. In some instances, however, it may be of importance to take into account not only the majority of current carriers with the sign of charge characteristic of the impurity semiconductor, but also the small number of current carriers with the opposite sign of charge which occurs in an extrinsic semiconductor by virtue of its intrinsic conductivity. In this connection it is of interest to note that if the intrinsic conductivity is characterized by the condition $n_i = p_i$, the following simple relationship is valid for a nondegenerate extrinsic semiconductor in which a majority of current carriers of one sign is furnished by the impurities while its intrinsic atoms supply a complementary number of current carriers of both signs:

$$n \cdot p = n_i^2. \tag{9}$$

In this formula n and p represent the total numbers of electrons and holes actually present in the given extrinsic semiconductor, and n [sic] is the concentration of current carriers of one sign which would occur in the given semiconductor if it were completely pure -- i.e., the concentration determined by means of Formula (4). Since n_i depends upon temperature only for a given semiconductor, $n \cdot p$ = const at a given temperature. This means that one of these two

STAT

Fig. 12. Diagram illustrating the appearance of current carriers in a semiconductor in the course of heating. a - intrinsic conductivity; b, c - extrinsic conductivity (b - electron conductivity, c - hole-type conductivity).

concentrations increases and the other diminishes, not only in the relative but also in the absolute sense, as the quantity of impurities is increased.

Fig. 12 illustrates the physical pattern of the appearance of intrinsic and extrinsic conductivity in a semiconductor in the course of heating as described in the present paragraph. a) relates to the case of intrinsic conductivity, b) and c) to cases of extrinsic conductivity -- the electron- and hole-type conductivities, respectively.

For practical purposes it is often necessary to know the influence exerted by impurities not on concentration but directly on specific resistance. For this purpose it is convenient to use the graphs of the type shown in Fig. 13 for germanium and silicon. The difference in resistance of the n and p specimens can be explained by the differing mobilities of electrons and holes (see table of the principal semiconductors at the end of Chapter IX). This diagram shows graphically the tremendous influence exercised by impurities on the conductivity of a semiconductor. Germanium shows a millionfold range of variation of resistance, while the same property for silicon ranges through a billionfold interval.

Fig. 13. Effect of impurities on electrical resistance of germanium and silicon at room temperature.

Impurity atoms may be distributed in the crystal lattice in two ways: either at lattice points, or by implantation in the interstices. In the first case we speak of substitutional and in the second of interstitial impurities. Let us subject the effect caused by a substitution atom in the germanium (or silicon) lattice to more detailed examination. For this purpose, it is essential to know that germanium is a tetravalent element of group IV of the Mendeleyev periodic system, and that it has a crystalline structure of the diamond type (see Fig. 14).

In the structure of this lattice, each atom is surrounded by its four closest neighbors, with which it interacts through covalent bonding forces. A characteristic of these forces is their saturation -- i.e., the closed state of the bonds, which prevents a given atom from interacting with any additional fifth atom.

To clarify the problem with which we are concerned here -- the effect caused by a substitution atom in the germanium lattice -- let us present the actual three-dimensional lattice as shown in Fig. 14 in the form of a plane network in which each atom is surrounded by its four closest neighbors of the same chemical nature (see Fig. 15). The outer electrons of these atoms, by which the valence bonds are formed, cannot break away from their atoms, migrate through the crystal, and thus participate in the electric current without expending a considerable amount of energy. But if a foreign atom of a different valence and chemical nature comes to replace one of the atoms of the basic substance at any lattice point, the system of the valence bonds in the vicinity of this substitution atom is disturbed, and one of two things happens -- see Fig. 16. If the valence of this impurity atom

Fig. 14. Structure of diamond (presented here to illustrate the formation by each atom of four bonds with its closest neighbors).

is higher than that of the atoms of the basic lattice -- i.e., if, for example, an atom of phosphorus or arsenic, which has five valence electrons, finds its way into a point of the germanium lattice, the fifth valence electron of the impurity atom, which is superfluous in the bond system of the germanium lattice, can break away from the impurity atom with relative ease (but with the expenditure of a certain small energy), begin to migrate through the crystal, and, in the case of application of an external electric field, this electron can participate in the electric current. In the case considered here, an atom of group V of the periodic table is a donor.

We had spoken previously of the event in which an electron breaks away from an atom as an event of excitation of the crystal in which the electron jumps from an impurity level to the conduction band (Fig. 12, b). The first explanation is more graphic, while the second, associated with Fig. 12, is a more convenient and shorter expression of the same physical process. The mechanism of NaCl conductivity, which was described in graphic terms on pages 29-33, can also be described briefly as the transfer of an electron under the influence of electromagnetic radiation from the valence band 3$\underline{p}$Cl into the conduction band 3$\underline{s}$Na (Fig. 6).

Fig. 15. Scheme of electronic bonds in the structure of diamond.

16

Fig. 16. Distortion of tetravalent germanium lattice, caused a) by an atom of trivalent boron, b) by an atom of pentavalent phosphorus.

A different result will be obtained in the case of replacement of a germanium atom by a trivalent boron, aluminum, or indium atom. In order to maintain the bond system characteristic of a diamond-type lattice, such an impurity atom may capture an electron from an adjacent germanium atom. As in the preceeding case, a certain energy must be consumed in order to accomplish this event. The value of this energy, which we referred to above as the energy of excitation or impurity activation, may differ between the event of removal of a surplus electron in the previous case and the borrowing of an electron as in the present case; it depends on the concentration of impurities, and, in most cases, declines as this concentration increases. For example, the energy of activation of donors in silicon amounts in many cases to 0.08 ev, and the activation energy of acceptors to 0.06 ev. According to experimental data, the dependency of the energy of activation of impurities ΔE_I upon their concentration N_I in *n*-type silicon takes the form

$$\Delta E_I = \Delta E_I^{o} - a \sqrt[3]{N_I}, \qquad (10)$$

where ΔE_I^{o} is the activation energy of donor atoms present in vanishingly small numbers, which is equal to 0.08 ev, and $\alpha = 4.3 \cdot 10^{-8}$. According to this empirical formula, when $N_I = 6 \cdot 10^{18} cm^{-3}$ -- i.e., when the concentration of impurities is only 0.01% -- $\Delta E_I = 0$, and silicon becomes a semimetal whose electron concentration is independent of temperature in the impurity region of its conductivity. The physical significance of Formula (10) becomes quite clear if one considers that the activation energy of impurity atoms may depend upon their interaction, and that the latter must be determined by the distance d between them, which, on the basis of simple considerations, changes according to the law

$$d = 1/\sqrt[3]{N_I} .$$

In germanium with small concentrations of group III and V impurities, the activation energies of the donors and acceptors are equal at 0.015 ev.

Let us return, however, to consideration of the case in which a substitutional impurity atom of low valence borrows one electron from a tetravalent lattice (see Fig. 16). In this case there appears in the lattice an electron hole which does not localize at any lattice point but migrates at random through the crystal. If such a crystal is subjected to an electric field, the motion of this hole becomes directional, manifesting itself in experiments as positive, hole-type conductivity. Actually, only electrons move in this case, but their successive jumps from one atom to another may be described formally as the motion of a single hole in the direction opposite to that of the electrons. It may be said that the relay-like translocation of many electrons is in this case replaced by the equivalent motion of one hole. The question as to why the motion of this hole is equivalent to the motion of a single positive charge cannot be satisfactorily answered by graphic presentations, since this theoretical conclusion proceeds from the non-objective properties of electrons (see pages 43-49). Thus the influence of substitutional impurity atoms on the atomic lattice of a diamond is determined by their valence, and we can nearly always predict the sign of conductivity which substitutional impurities of one type or another will create in such a lattice.

STAT

If impurity atoms are implanted in the lattice interstices of semiconductors belonging to group IV of the periodic table, the sign of the impurity conductivity is determined by the dimensions of the implanted atoms and by their electronegativity (see page 4). It is known from experimental data, for example, that contrary to the above simple rule of valency, the behavior by lithium (group I) in the germanium lattice is that of a donor, while that of oxygen (group VI) is that of an acceptor. In order to explain these factors it is assumed that lithium and oxygen are implanted in the interstices of the germanium lattice. The implantation of the dimensionally large lithium atom in the tight interstices of the germanium lattice is possible by virtue of the loose bonds of its valence electron which, in a medium with a high dielectric constant, breaks away from its atom with relative ease (see page 60). The small dimensions of the lithium ion permit its implantation in the tight interstices of the lattice, while the released electron imparts electron-type conductivity to the germanium crystal. On the other hand, the implantation in the interstices of an oxygen atom, which is of relatively small dimensions and high electronegativity, is accompanied by capture of electrons of the basic lattice with the result that the latter acquires hole-type conductivity.

If a Ge or Si atom is transferred (for instance, by thermal agitation) into an interstice, two impurity levels are created: The implanted atom acts as a donor, and the vacant lattice point as an acceptor.

The effect of impurities on the conductivity of <u>intermetallic semiconductors</u> may be judged from experimental data in regard to the most thoroughly investigated group of compounds of elements of groups III and V of the table of the elements (InSb, GaSb, InAs, etc). The sign of conductivity is determined in these semiconductors by the same simple valency rule as in the case of the elementary semiconductors of group IV (Ge, Si), namely: substitutional impurity atoms of Group II (Mg, Zn, etc.), the valency of which is lower, create acceptor levels, while impurity atoms of group IV (Te, Se), the valency of which is higher, form donor levels. If we introduce impurity atoms of group IV into these intermetallic compounds, the result will depend on which atom of the compound is replaced by the impurity atom. If, for example, a tetravalent impurity atom replaces a trivalent In atom in the InAs lattice, it will act as a donor; if the same impurity atom replaces a pentavalent As atom in the same lattice, its behavior is that of an acceptor. Which of the two atoms is substituted in each particular instance apparently depends on the relationship between the dimensions of the atoms. Thus, for instance, on finding its way into the InSb lattice, a Pb impurity atom, which is characterized by relatively large dimensions, replaces an In atom therein and acts as a donor; entering an AlSb lattice, the same impurity atom assumes the position of an Sb atom and acts as an acceptor. Substitution of the components of an intermetallic compound by other elements of the same groups of the periodic table (III and V) has no distinct effect on the conductivity of the given intermetallic semiconductors. Very little is as yet known concerning the effects of interstitial impurities on the lattices of these materials.

For ionic lattices (PbS, CdS and oxides) such simple and general rules cannot yet be formulated; therefore the development of new semiconductors of this group proceeds principally by trial and error. It is still not clear why a given impurity, when injected into these semiconductors, may prove utterly ineffective on one occasion -- i.e., to have no influence on the concentration of current carriers -- and on another occasion confer conductivity of the sign opposite to that which one might expect by analogy. This may result from the fact that in semiconductors of the ionic type, the bonds between the particles are maintained in actuality not only by the electrostatic attraction of unlike ions but by a combination of electrovalent and covalent bonds as well. General considerations justify the conclusion that a metallic atom may be implanted in the tight interstices of the lattice only in cases where its dimensions are greatly diminished by the loss of outer electrons, which, migrating through the crystal, impart electron-type conductivity to it. It also follows from general considerations that ions of many nonmetals cannot be implanted in interstices due to their large ionic radii.

If a certain number of negative lattice points are vacant in an ionic lattice, as may be the case with an excess of the metal, it follows from the consideration of the electrical neutrality of the body taken as a whole that the same number of positive ions must be neutralized by acquisition of electrons. This state of electrical neutrality attained by acquisition of the electrons can migrate through the crystal, thus creating electron-type conductivity. By the same reasoning, an excess of the nonmetal realized at the expense of vacant positive lattice points must lead either to the neutralization of adjacent negative ions, or to further ionization (if this is more favorable in terms of energy) of the metallic ions. In either case, the vacant lattice point binds an electron of the lattice -- i.e., creates in the lattice an electron hole which, migrating through the crystal, imparts hole-type conductivity to it. In both of the above cases, the conducting state of the ionic crystal is a state of excitation. The physical significance of this statement consists in the fact that the breaking away of an electron in the first case and the formation of a hole in the crystal in the second require the expenditure of activation energy. Therefore at the absolute zero of temperature, in total darkness, and in the presence of a weak external electric field, such an ionic crystal should be a perfect insulator.

From experimental data on ionic semiconductors we may conclude that the following is usually the case with oxides and sulfides. If the semiconductor can attain conductivity of either sign, as, for instance, PbS, a stoichiometric excess of sulphur or the admixture of oxygen imparts hole-type conductivity to this substance, and an excess in metal a corresponding electron-type conductivity. In semiconductors with impurity conductivities of one sign, an increase in the number of holes is obtained in a hole-type semiconductor through an excess of oxygen or sulfur; an increase in the number of electrons in an electron-type

STAT

semiconductor may be obtained by reducing the concentrations of these elements. It is well known from experience, for example, that the introduction of hole-type cuprous oxide (Cu_2O) into an atmosphere of elevated oxygen concentration results in a rise in conductivity, while the same experiment with zinc oxide has the reverse effect.

For practical purposes it is important that the sign of the impurity conductivity of a semiconductor change either as a function of the impurities which are added to it, or generally in dependence on the treatment to which it is subjected. As was noted above, by no means all semiconductors comply with this condition. For instance, the conductivity of Cu_2O is always of the hole type, while that of ZnO is only of the electron type. Nobody has ever succeeded in producing tellurium or selenium of negative impurity conductivity. The impurity conductivity of many semiconductors changes sign depending on the chemical nature of the impurities injected, but sometimes this depends on their thermal treatment or the exposure to radiation by electrons, neutrons, etc., as well. These are called amphoteric semiconductors. They include germanium, silicon, lead sulfide, silicon carbide, lead telluride and many others.

Of great importance is the circumstance that not only may different fragments of an amphoteric semiconductor possess conductivities of different signs, but regions of different types of conductivity may also be formed at different places in the same specimen of such a semiconductor. For example, by taking a pure germanium or silicon lamina with electron-type conductivity and building up a thin layer of boron or indium on one of its sides, we may, after appropriate heating of this lamina, obtain two sections with different conductivity types. That part of the lamina to which the atoms of boron or indium have found access by diffusion will have hole-type conductivity, while the remainder will have electron-type conductivity. For the practical use of such semiconductors as rectifiers, amplifiers and phototubes, it is of great importance that the transition layer between these two parts, which is referred to as the electron-hole or *p-n* junction, be of the order of 10 cm* in width. The electronic processes which unfold in this transition region are the physical basis of the technological applications mentioned above, and form a separate important division of semiconductor physics. (This problem is discussed in Chapter X.)

It was pointed out above that simple lattice defects, such as cracks and shears, may act as impurities in regard to their effect on the concentration of current carriers. Indeed, it has recently been demonstrated that, for example, plastic shears in a germanium lattice act as acceptors which furnish hole-type current carriers.

It is not possible in all cases to explain the conductivity experimentally observed in solid substances on the basis of the band theory of the electronic energy spectrum. Thus, for instance, the reason why such transition-metal oxides as NiO, CoO, MnO and others do not possess metallic conductivity cannot be explained within the framework of this simple theory. After all, according to this theory, the NiO crystal should include, in addition to the filled $2\underline{p}O^{--}$ band, an unfilled $3\underline{d}Ni^{++}$ band which contains only eight electrons for ten vacant places. In the CoO crystal, only 7 electrons occupy the 10 places of the $3\underline{d}Co^{++}$ band, etc. Why, then, are such crystals insulators and not metals at the stoichiometric composition? The explanation of this fact apparently involves taking into account the interaction of atoms, which the band theory does not consider. Impurity conductivity of the *p*-type can be observed in these crystals after heating in an atmosphere of oxygen, when acceptor levels appear in the energy spectra of these crystals.

In concluding this section, which has been devoted to the explanation of the nature of intrinsic and extrinsic conductivity in semiconductors, let us consider briefly one more specific variety of intrinsic conductivity which we encounter in the semiconductor group called ferrites. This semiconductor group has recently

---

*Translator's Note: illegible in original. 10^{-5}?

acquired great technological importance by virtue of its highly favorable combination of those physical properties which are required for the cores of high-frequency transformers -- namely, ferromagnetism and high ohmic resistance. Two circumstances are significant for the mechanism of conductivity in this group of ionic semiconductors: 1) the position of metallic ions in the crystal lattice, and 2) the variable valency of these ions. Ferrites are oxides of which magnetite, Fe_3O_4, is the chief representative. Upon partial replacement of the iron atoms of this compound by divalent atoms (Ni, Co, Zn, Cd, etc.), we obtain materials with a wide variety of magnetic and electrical properties. These substances crystallize in a close-packed cubic lattice of the inverse spinel type -- see Fig. 17. The unit cell of this structure contains 32 oxygen ions, 16 iron ions which are implanted between these in octahedral positions, and the 8 remaining ions of the metal in tetrahedral positions. Electron exchanges may occur between the doubly-charged and triply-charged ions, which are distributed statistically among the octahedral vacancies of the cell. The transition of an electron from a doubly-charged to a triply-charged ion of the same substance actually amounts to a simple exchange of places between these ions. Since these ions occupy similar positions in the crystal lattice, the energy of the initial and final states of the crystal cell remains unchanged even though these states are separated by an energy barrier. Thus the motion of electrons between closely situated identical doubly- and triply-charged ions assumes an entirely disordered nature in the absence of a field, and is accomplished by transition of electrons across an energy barrier; therefore its intensity -- i.e., the frequency of electron transitions -- increases with rising temperature. On the application of an external electric field, these disordered electron transitions acquire a preferential directional tendency, with the result that an electric current arises in the substance, and, in a given electric field, increases rapidly with rising temperature according to the exponential law characteristic of semiconductors, with an activation energy of $\sim$0.1 ev

Fig. 17. Unit inverse-spinel cell of Fe_3O_4.

in the case of Fe_3O_4. The above exposition should facilitate understanding of the difference between the mechanism of conductivity in the ordinary semiconductors of the germanium, silicon, lead sulfide and other types, which were discussed previously, and the oxide semiconductors with their metallic ions of variable valency. In the former, an increase in temperature causes an increase in the number of current carriers in both the intrinsic and extrinsic regions of conductivity; in the latter, temperature affects not the number of current carriers but their effective mobility outside the conduction band. Thus impurities also affect the electric properties of these two semiconductor groups in completely different ways. We already know that the introduction of infinitesimal amounts of impurities into pure specimens of germanium and lead sulfide produces a sharp increase in their conductivity. In semiconductors with variable valency, the injection of impurities disturbs the system of electron transitions which we have described and causes a decline in conductivity. It is known, for example, that on the addition of nickel or zinc impurities to the good conductor magnetite (Fe_3O_4), the conductivity of the stoichiometric composition declines. This may be explained as a result of inhibition of electron transitions by Ni^{++} or Zn^{++}

ions distributed statistically among the octahedral vacant spaces in the crystal cell. The transition of an electron from Ni^{++} to Fe^{+++}, which is possible in principle, involves a considerable change in the energy of the crystal, and therefore its occurence is of small probability.

STAT

CHAPTER IV

DEPENDENCE OF SEMICONDUCTOR CONDUCTIVITY ON TEMPERATURE

In order to analyze the effect of temperature on the electric conductivity of semiconductors, one should proceed from the expression (5), the more general form of which is

$$\sigma = en_1u_1 + en_2u_2 + en_3u_3 + \ldots \qquad (12)$$

This sum is arithmetical for the reason that if the signs of the charges e of the current carriers differ, the sign of the mobility of a current carrier will change simultaneously with that of its charge, so that the products of these values, which enter the expression for specific conductivity (12), have the same sign for all current carriers. The physical significance of this statement is obvious: in an electric field, current carriers with different charge signs move in opposite directions, and the total current equals the arithmetic sum of the individual currents. Although we generally deal with only one type of current carrier in the region of extrinsic conductivity, and in the region of intrinsic conductivity with two such types which differ with respect to both sign of charge and degree of mobility, the participation of an even larger number of current-carrier species in the electric current is still possible. Thus, for example, it is now known that two species of hole-type current carriers of differing mobility exist in germanium; moreover, the participation in the electric current of impurity-band current carriers whose mobility differs from that of the valence-band holes becomes apparent in the extrinsic-conductivity region of germanium at low temperatures. Two types of hole have been observed in the valence band and two types of electron in the conduction band of tellurium.

The theoretical determination of the specific conductivity of a semiconductor is such a complex problem that its solution has never been fully achieved for any specific semiconductor. True, there exist theoretical papers in which the width of the forbidden band is determined for diamond, silicon, germanium and certain other semiconductors, but the results of these calculations are still not sufficiently in keeping with experimental data. Our knowledge of the other values by which conductivity is determined -- such as the mobility and the effective mass of current carriers -- has thus far been obtained from experimental data. The experimental method for determining concentration and mobility by means of the Hall effect was mentioned at the beginning, on page 16. Eight independent methods for the determination of the effective mass m exist at the present time; however, the results obtained from investigations of one and the same substance by different methods do not, for the most part, coincide -- sometimes not even in order of magnitude. This can be explained both by the imperfection of the theory according to which the relationship of the effective mass to the experimental data is determined and by the complex nature of this value. The values for the effective masses of current carriers in semiconductors investigated thus far range from several units to hundredths of the true mass of the electron. The most complete and authentic information on the effective mass which, in a crystal, constitutes an anisotropic value -- i.e., may, as a consequence, have different values for different directions in the crystal -- is furnished by the diamagnetic-resonance method. We cannot dwell here on the explanation of this method since it is of no direct concern to the present problem of the dependency of conductivity upon temperature.

The experimental determination of the width of the forbidden band of a crystal and the activation energy of impurities is carried out on the basis of the temperature curve of either the Hall coefficient R_H or the electric conductivity σ, and directly concerns the subject of this paragraph. Modern theory satisfactorily explains the effect of such physical factors as temperature, impurities, pressure, dielectric constant, elasticity coefficient, the crystal-lattice parameters, the nature of binding forces acting in the lattice, lattice defects and the like on the concentration and mobility of current carriers. Of greatest interest in practice are the effects of impurities and temperature. The effect of impurities has been discussed above. Let us turn now to the effect of temperature.

The effect of temperature on the conductivity of a semiconductor is a consequence of changes in the concentration and mobility of current carriers with variation of the temperature of the substance. In the first place, therefore, let us clarify the influence which temperature exerts on each of these factors individually. We shall do this with reference to the simple, yet practically most important cases of the existence of a single type of current carrier in the extrinsic-conductivity region of the semiconductor, and of two types in its region of intrinsic conductivity -- for which last case, as we know, $n_n = n_p$. The dependency of concentration on temperature has already been discussed for these two cases, and is summarized by Formulas (4) and (7). The general result obtained for the change in current-carrier concentration with temperature is conveniently represented in the form of a graph whose axes are calibrated for values of $\log n$ and $\frac{1}{2kT}$. Fig. 18 shows the typical form of this graph. The choice of semilogarithmic coordinates is convenient because the entire curve of the variation in concentration may be presented in the form of a broken line consisting of three rectilinear segments. The slopes of the segments ab and cd determine, according to Formulas (7) and (4), the activation energies ΔE_I and ΔE_0. Such graphs are plotted in experimental determinations of these energies. More accurate current-carrier activation-energy values are obtained when the experimental results are

STAT

Fig. 18. Graph showing typical dependency of current-carrier concentration on temperature in a semiconductor.

plotted on a graph with log (RT*) laid off on the axis of ordinates if ΔE_I is to be determined or log (RT*) if ΔE_0 is to be determined and $\frac{1}{2kT}$ plotted against the axis of abscissas; see Formulas (3), (4), (6) and (7). In the low-temperature region, on segment ab, we deal with extrinsic conductivity only, since the number of intrinsic current carriers is still very small due to the low temperature, and the slope of the line is determined in this segment by the impurity-activation energy ΔE_I. With rising temperature, (segment ab) the number of impurity carriers will increase until the electron reserves of the impurity atoms are depleted. In the segment bc, impurity reserves are already exhausted, while the intrinsic conductivity has not yet become apparent; hence this temperature range, in which the concentration does not vary, is called the region of depletion. Finally, the temperature in segment cd has become so high that the number of intrinsic current carriers begins to increase rapidly, so that we enter the region of intrinsic conductivity of the semiconductor, which is characterized by the activation

---

* Trans. note: the exponents are illegible in original.

energy ΔE_0. Experiment indicates that the slope of the segment ab depends on the impurity concentration N_I. The physical cause of this phenomenon has already been discussed above in connection with Formula (10), which determines the dependency of ΔE_I on N_I in the particular case of silicon. Fig. 18 illustrates the rule set forth on page 61 according to which the temperature of the transition from extrinsic to intrinsic conductivity shifts toward higher temperatures with increasing impurity concentration.

The slope of the line does not depend on impurity concentration in the region of intrinsic conductivity and is a constant characteristic of the electric properties of a given semiconductor. At some higher impurity concentration ($\sim 1\%$), when $\Delta E_I \approx 0$, the entire curve may be represented in the form of the two segments a"c" and c"d". In the extrinsic-conductivity region a"c", the current-carrier concentration of the semiconductor is independent of temperature from the lowest temperatures (a'') to the temperature of onset of the intrinsic conductivity (c"). This behavior of semiconductors -- i.e., their assumption of properties characteristic of metals (constancy of concentration) in the region of extrinsic conductivity -- induces us to segregate them into a special group of semimetals. It should be clear from the above that any semiconductor may become a semimetal at a sufficiently high concentration of dissolved impurities, and conversely that a semimetal should become a typical semiconductor after thorough purification, with the concentration of its current carriers strongly dependent on temperature. The current-carrier concentration in semimetals is usually so high that the electron gas in these materials enters the region of degeneracy (see page 58).

Before discussing the mechanism by which temperature affects the mobility of current carriers in semiconductors, let us first explain the factors which determine its absolute value. After all, it may seem strange at first glance that the mobility of current carriers in certain semiconductors whose conductivities are, as a rule, insignificant by comparison to those of metals proves to be hundreds of times greater than those of typical metals (see page 18). The fact, also mentioned in the same place, that certain insulators show mobilities higher than those of metals may seem even more surprising. What is the reason for this? Why can electrons sometimes move more "freely" in poorly conducting crystals than in the highly conductive metals? First of all, let us remember that according to one of the most basic conclusions of the electronic theory of crystals, as discussed on page 41, it is not the question of why the free paths of an electron within a crystal may exceed the interatomic distances by tens and even hundreds of times which we must subject to scrutiny, but rather the question of what, in a real crystal, limits the infinitely large mobility which an electron should theoretically have in an ideal crystal (at absolute zero).

According to Formula (2), which determines the mobility of the current carriers

$$u = \frac{e}{m}\tau = \frac{e}{m}\,\frac{l}{v}, \tag{2'}$$

a high mobility may be due to a small effective mass (m) of the current carrier and a large value of the free time or, more precisely, of the relaxation time τ.

The numerical value of the effective mass is, as we know, determined by the conditions of motion of the electron over the length of its free path in the internal field of the crystal. In the theory of conductivity, as in classical mechanics, this value essentially determines the directional acceleration which the electron acquires in an external electric field in the intervals between collisions; it is this directed acceleration which creates a current in a solid (see page 11). The general conditions of the motion of electrons, which determine their effective masses, are identical in metallic and nonmetallic crystals, so that we cannot draw distinctions between these two classes of solid substances in discussing the problem of the effective mass of current carriers. In both metals and semiconductors the effective mass of the current carriers may be either larger

STAT

or smaller than the mass of a free electron.

The relaxation time τ in Formula (2) determines the interval of time in the course of which the decay of the originally established current occurs following cutoff of the external field. The numerical value of this quantity is determined by collision processes or, as one would say, the scattering processes to which electrons are subject when moving within a crystal. It is obvious that the more intensive and more frequent these processes are, the smaller will be the time of relaxation and, consequently, the mobility as well.

Let us dwell in somewhat greater detail on the physical pattern of the electron-scattering process in a crystal and clarify the significant differences between these processes in metals and semiconductors which affect relaxation time and, consequently, mobility.

The conditions of scattering of electrons in a crystal are determined by their wave properties. It will be recalled that these wave properties of electrons have been observed directly in certain electron-diffraction experiments (see Shpol'ski, Atomic Physics). Essentially, the wave properties of electrons form the foundation on which the entire modern theory of conductivity in solid substances is based. Consideration of these properties has made it possible to demonstrate the physical significance of such important concepts as the length of free path of a current carrier -- a concept which found no satisfactory interpretation in the old prequantum theories.

In modern quantum theory, it is not the classical Newtonian equation which describes the motion of an electron within a crystal, but Schroedinger's wave equation (see Shpol'ski, Atomic Physics). In the simplest case, in which the interaction between an electron and the atoms (or ions) of a crystal lattice is accounted for by introduction of the effective mass into the Schroedinger equation (see page 45), the wave which describes the behavior of such a quasi-free particle is a plane wave. As is known from the elementary physics course, a plane wave is characterized by two parameters: the vibrational frequency ν and wavelength λ. On comparison of the wave presentation of the motion of the electron with the corpuscular presentation, in which the basic parameters of the motion are the values of mass, velocity, and energy, we should assume that the frequency ν is determined by the energy: $\nu = \frac{E}{h}$, and the wavelength λ by the momentum of the electron: $\lambda = \frac{h}{m\,v}$

At room temperature, the average random velocity of the free electrons in a nondegenerate semiconductor (and, if they are present, in an insulator) is 10^7 cm/sec. According to the formula given above, the wavelength will be $7 \cdot 10^{-7}$cm. In a metal, where the discrete, quantized character of the energy spectrum is essentially in evidence, the velocity of the free electrons is determined not by the temperature of the substance, but by their concentration (this circumstance is also referred to as **degeneracy**). At the high values assumed by the free-electron concentration in typical metals ($\sim 10^{22}$ cm^{-3}), the velocity of their motion attains values which exceed the average velocity of electrons in semiconductors by several thousand percent. Thus the wavelength of the electrons participating in the conductivity of a metal proves to be one order less ($\sim 5 \cdot 10^{-8}$cm) than in a semiconductor.

Let us note here that the very fact of the higher electron velocities in metals must result, according to Formula (2), in decreased mobility (other conditions being equal) by comparison with semiconductors.

Let us now ascertain the manner in which the established difference in wavelengths affects the process of electron scattering. Proceeding from the wave presentation of the motion of the electron, we should assume that the process of electron-wave scattering is basically determined, just as with any other waves, by those nonuniformities in the disposition of atoms which extend to distances of the order of one-quarter of an electron-wavelength. Nonuniformities of smaller dimensions do not produce any appreciable wave scattering.

As an example which explains this conclusion of wave theory, let us recall

[redacted]

STAT

the explanation for the experimental fact of the transparency of glasses and liquids. The characteristic nonuniformities of such substances, which are caused by the random distribution of atoms in amorphous media, are of atomic dimensions ($\sim 10^{-8}$cm), while the wavelength of light is a thousand times longer ($\sim 10^{-5}$cm). This dimensional ratio being the case, the propagation of light waves in a disordered amorphous medium occurs in the same manner as in an ordered transparent crystal -- i.e., with no appreciable scattering. In turbid media, in which the dimensions of the spatial nonuniformities attain values comparable to the wavelength of light, a strong scattering effect is to be observed. These nonuniformities of turbid media may be caused either by pollution of the medium by foreign particles of appropriate dimensions, or by fluctuations in molecular density, such as occur in gases as a result of chaotic thermal motion of molecules. This motion, as we know, provides the explanation for the blue color of the sky (the short-wavelength rays of the solar spectrum experience the most intensive scattering in the atmosphere). Spatial nonuniformities of atomic dimensions may cause appreciable scattering in the instance when the wavelength of the electromagnetic oscillation is of the same order of magnitude. As we know, such a condition occurs in the propagation of X-rays in glasses or liquids.

The scattering of electron waves in a crystal, just as that of light in a turbid medium, is effected by nonuniformities of two different types: disturbances of the homogeneous structure of the crystal due to impurities, as well as by lattice defects and by fluctuating nonuniformities which arise from the thermal vibration of the atoms in the lattice. The above-established difference in the electron wavelengths enables us to understand the nature of the difference between the conditions for electron scattering in metals and in semiconductors (as in insulators), and why, in certain cases, the free path of an electron may be longer in the latter than in the former. As we see, this consists in the fact that due to the long wavelength of the electrons in semiconductor and insulator crystals, nonuniformities of an atomic order do not give rise to appreciable electron scattering, while in a metallic crystal, in which the wavelength of the electron is shorter by one order, such nonuniformities cause appreciable scattering, due to which the mobility of the electrons is reduced.

It does not follow from the above, of course, that semiconductors in general are distinguished from metals by the higher mobility of their current carriers. From the table presented at the end of Chapter IX it may be seen that high mobilities, exceeding 10,000 cm^2/v.sec, are observed in only a few semiconductors (which are prepared, moreover, in a very pure form). In many instances, the mobility of the current carriers in semiconductors proves so small that the free paths computed from these data do not amount to more than fractions of the interatomic distances in the lattice. Serious difficulties arise in theoretical interpretation of the results of measurement in these instances, since by its very physical significance in the theory of electric conductivity, the free path may not be smaller than the distance between neighboring atoms. This means, essentially, that the conductivity of semiconductors with low current-carrier mobilities cannot be investigated within the scope of the existing theory and that one of the problems to be met in further development of this theory is that of the clarification of this question.

The above differences in the physical conditions in which electron scattering occurs in metals and in semiconductors are manifested not only in the numerical values for mobility, but also in the dependency of mobility on temperature, with which we are concerned in this chapter. Also essential in this dependency is the fact that in semiconductors, as opposed to metals, electron velocity is a function of temperature (as in the classical ideal gas), which has a direct effect on conductivity, as seen from Formula (2).

The manner in which temperature affects the length of free path is variable, depending on the mechanism of scattering.

STAT

We should consider the following fundamental mechanisms of scattering for semiconductors:

(1) by thermal vibrations of the atoms or ions of which the crystal lattice is composed,

(2) by the impurities which may be present in a semiconductor, in either the ionized or neutral state,

(3) by all types of lattice defects, such as vacancies, distortions caused by atom implantations, shears, cracks, crystal-grain boundaries, etc.; modern semiconductor theory neglects reciprocal scattering by the electrons themselves (due to the small concentration).

In atomic lattices (Ge, Si, Te and others) we deal primarily with two mechanisms of current-carrier scattering which produce two essentially different temperature dependencies of mobility -- one with scattering due to the thermal vibrations of the lattice, and another with scattering due to ionized impurities. Theory indicates that for scattering due to thermal vibrations of the lattice, the free path l has the same value for all current carriers irrespective of velocity, and is inversely proportional to the **absolute temperature** of the substance -- i.e., $l \sim \frac{1}{T}$. This last result can be understood from the following simple reasoning. It follows from the general classic considerations that the scattering of current carriers should be directly proportional to the cross section of the volume occupied by the vibrating atom; on the basis of simple geometrical considerations, this cross section may in turn be considered directly proportional to the squared amplitude of vibration of the atom, while the latter, which determines the energy of the lattice, is known to increase as a linear function of temperature. Since, according to Formula (2)

$$u = \frac{e\bar{\tau}}{m} = \frac{e\bar{l}}{mv}, \text{ and } v \sim \sqrt{T},$$

$$u_T \sim \frac{1}{T^{3/2}} . \qquad (13)$$

This law is not always confirmed experimentally. Along with instances in which the dependency takes the form (13), instances of still more marked dependency of mobility upon temperature (up to $u \sim T^{-3}$) are also encountered. At low temperatures the effectiveness of thermal scattering according to (13) diminishes, and another mechanism of scattering becomes predominant in atomic lattices -- the mechanism of scattering on ionized impurities. In essence, this mechanism of scattering is identical to the well-known mechanism of alpha-particle scattering, which was first investigated by Rutherford in his classical experiments on the structure of atoms. Characteristic of this mechanism is the decrease in the effectiveness of scattering of a moving charge as its velocity v increases, with the result that the length of free path of the current carrier increases with velocity according to the law $l \sim v^4$. In this case, for the mobility of a current carrier u_c, the theory gives the expression:

$$u_c \sim \frac{T^{3/2}}{N_I}, \qquad (14)$$

If both the thermal and ionic mechanisms participate in the scattering of current carriers, the resulting mobility may be presented approximately in the form

$$\frac{1}{u} = \frac{1}{u_c} + \frac{1}{u_T} ,$$

and the temperature curve of mobility by

$$\frac{1}{u} = aT^{-3/2} + bT^{3/2} . \qquad (15)$$

At low temperatures the first term predominates, while the second term comes to the fore at higher temperatures.

The approximate graphic form of the relationship (15) is presented in Fig. 19. The position of the maximum of this curve depends on the concentration of impurities; as the latter increases, the maximum moves in the direction of higher temperatures. The scattering due to neutral impurities does not depend on temperature and its importance is usually secondary.

Turning to the question of the mobility of current carriers in ionic crystals (oxides, sulfides, etc.), the following general statement should be made at the outset. The interaction between the current carriers and the vibrating charged particles in the ionic lattice (i.e., the ions) is, on the whole, much stronger than the interaction between the current carriers and the neutral atoms of the atomic lattice. Therefore, the scattering of current carriers due to thermal vibrations in ionic crystals occurs with far greater intensity, and the mobility proves to be lower in most cases in this semiconductor group. The qualitative curve of the temperature-dependence of mobility is the same in this case as in atomic crystals: mobility diminishes with rising temperature, but the specific form of the theoretical formula differs for high and low temperatures, the boundary between these two regions being determined by the maximal frequency ν_m of the so-called longitudinal optical vibrations of the ions, which is a characteristic parameter for any given ionic crystal. At low temperatures, $u \sim e^{\frac{h\nu_m}{kT}}$, and at high temperatures $u \sim \frac{1}{\sqrt{T}}$. Experimental data pertaining to this problem are still meagre and, on the whole, do not agree satisfactorily with the theoreti-

Fig. 19. Typical curve of temperature-dependency of current-carrier mobility in a semiconductor.

[redacted]

STAT

The approximate graphic form of the relationship (15) is presented in Fig. 19. The position of the maximum of this curve depends on the concentration of impurities; as the latter increases, the maximum moves in the direction of higher temperatures. The scattering due to neutral impurities does not depend on temperature and its importance is usually secondary.

Turning to the question of the mobility of current carriers in ionic crystals (oxides, sulfides, etc.), the following general statement should be made at the outset. The interaction between the current carriers and the vibrating charged particles in the ionic lattice (i.e., the ions) is, on the whole, much stronger than the interaction between the current carriers and the neutral atoms of the atomic lattice. Therefore, the scattering of current carriers due to thermal vibrations in ionic crystals occurs with far greater intensity, and the mobility proves to be lower in most cases in this semiconductor group. The qualitative curve of the temperature-dependence of mobility is the same in this case as in atomic crystals: mobility diminishes with rising temperature, but the specific form of the theoretical formula differs for high and low temperatures, the boundary between these two regions being determined by the maximal frequency ν_m of the so-called longitudinal optical vibrations of the ions, which is a characteristic parameter for any given ionic crystal. At low temperatures, $u \sim e^{\frac{h\nu_m}{kT}}$, and at high temperatures $u \sim \frac{1}{\sqrt{T}}$. Experimental data pertaining to this problem are still meagre and, on the whole, do not agree satisfactorily with the theoreti-

Fig. 19. Typical curve of temperature-dependency of current-carrier mobility in a semiconductor.

cal formulas. This may possibly be related to the general observation on ionic semiconductors put forth on page 70.

Now, having explained the separate effects of temperature on the concentration and mobility of current carriers, let us describe the general course of the variation of conductivity with temperature. Since in atomic lattices (and, at higher temperatures, in ionic lattices) mobility may be a comparatively low power function of temperature, while concentration may vary in accordance with a very high exponential law, the overall temperature curve of conductivity will have approximately the same appearance as the curve for concentration shown in Fig. 18, with the exception that in the depletion region, in which concentration is constant, the temperature curve of conductivity is determined by the variation of temperature with mobility. If the chief factor in the depletion region is the thermal mechanism of scattering, in which mobility diminishes with rising temperature, the overall conductivity curve will have the form abcd indicated in Fig. 20; if mobility is determined in the region of depletion by scattering due to ionized impurities, in which case mobility increases with rising temperature, the general form of the curve will approximate a"c"d.

At high temperatures, when the current-carrier concentration becomes so large that the electron gas enters the region of degeneracy (see page 61), σ increases with temperature at a slower rate than indicated by the exponential law presented in Fig. 20 for the intrinsic-conductivity segment cd in the non-degenerate case. Fig. 20 illustrates the frequently-applied experimental method by means of which the width ΔE_0 of the forbidden band of a crystal and the activation energy of the impurities ΔE_I are determined. Figs. 18, 19 and 20 do not apply to any specific semiconductor; they are plotted on the basis of the general theoretical laws set forth above for the influence of temperature on the concentration, mobility and conductivity of an idealized semiconductor.

[redacted]

STAT

In real semiconductors, the shape of these curves may deviate widely from the above -- both for the reason that the physical phenomena are themselves described approximately in the theory, and because not one, but several types of impurity centers, the activation energies of which may differ, exist in the materials with which we deal in practice.

Fig. 20. Typical graph of temperature-dependence of conductivity in a semiconductor.

Figs. 21, 22 and 23 show experimental curves for the temperature-dependence of conductivity for the atomic semiconductor gray tin, the ionic semiconductor cuprous oxide, and the intermetallic compound Mg_2Si, with different quantities of impurities. Figs. 24, 25 and 26 show the temperature dependencies of concentration mobility and conductivity for silicon containing different quantities of phosphorus as an impurity.

Fig. 21. Experimental curves of dependence of conductivity of gray tin atomic crystal on temperature for specimens of differing purity.

Fig. 22. Experimental curves for dependency of conductivity of cuprous oxide ionic crystal on temperature for specimens with different excess oxygen contents: 1: cuprous oxide of vitrometric composition; quantity of oxygen increases in the order 2 - 3 . . . . 7.

In Fig. 26, as is often done, the specific resistance ρ, rather than the conductivity σ, is plotted on the axis of ordinates. Fig. 27 shows experimental results of the study of the temperature-dependency of the resistance of germanium with different injections of antimony impurity.

Fig. 23. Experimental curves of temperature-dependency of conductivity for the intermetallic compound Mg_2Si with different impurity contents.

In conclusion let us note that the strong dependence of the conductivity of semiconductors upon temperature is utilized in the design of high-quality technical equipment for high-precision temperature measurements, regulation of current increase rate (time relay), stabilization of circuit voltage, vacuum and radiant-energy measurement, control and regulation of upper-limit current, voltage, sound level, etc. If the temperature-dependence of the conductivity of metals were applied in the design of these devices, which are called thermistors, their sensitivities would have been tens of times lower than those obtained by the use of semiconductors. Another important advantage of the latter consists in small inertia, since the thermosensitive element of a semiconductor thermistor can be made very small.

[redacted]

STAT

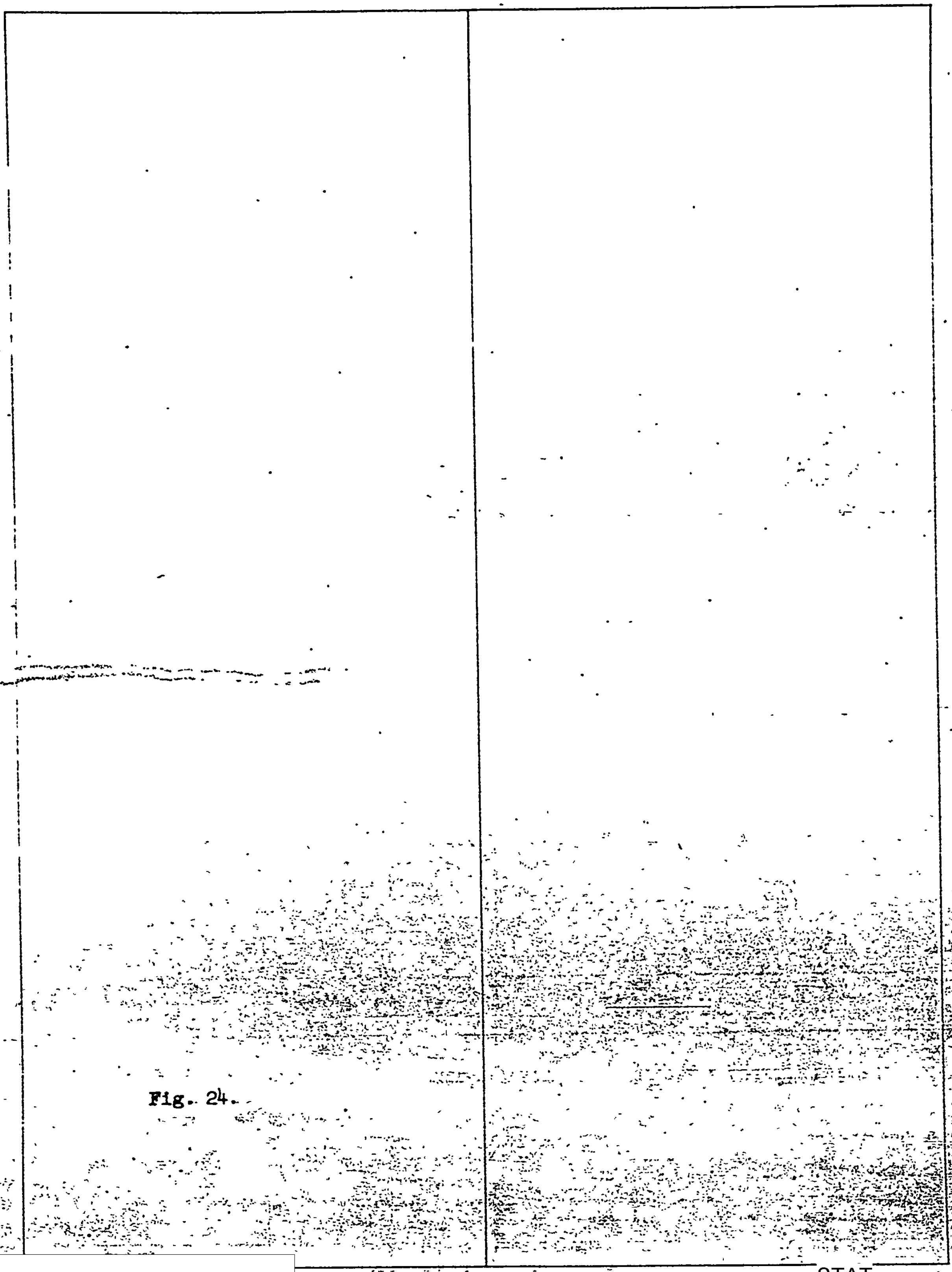

Fig. 24.

[redacted]

STAT

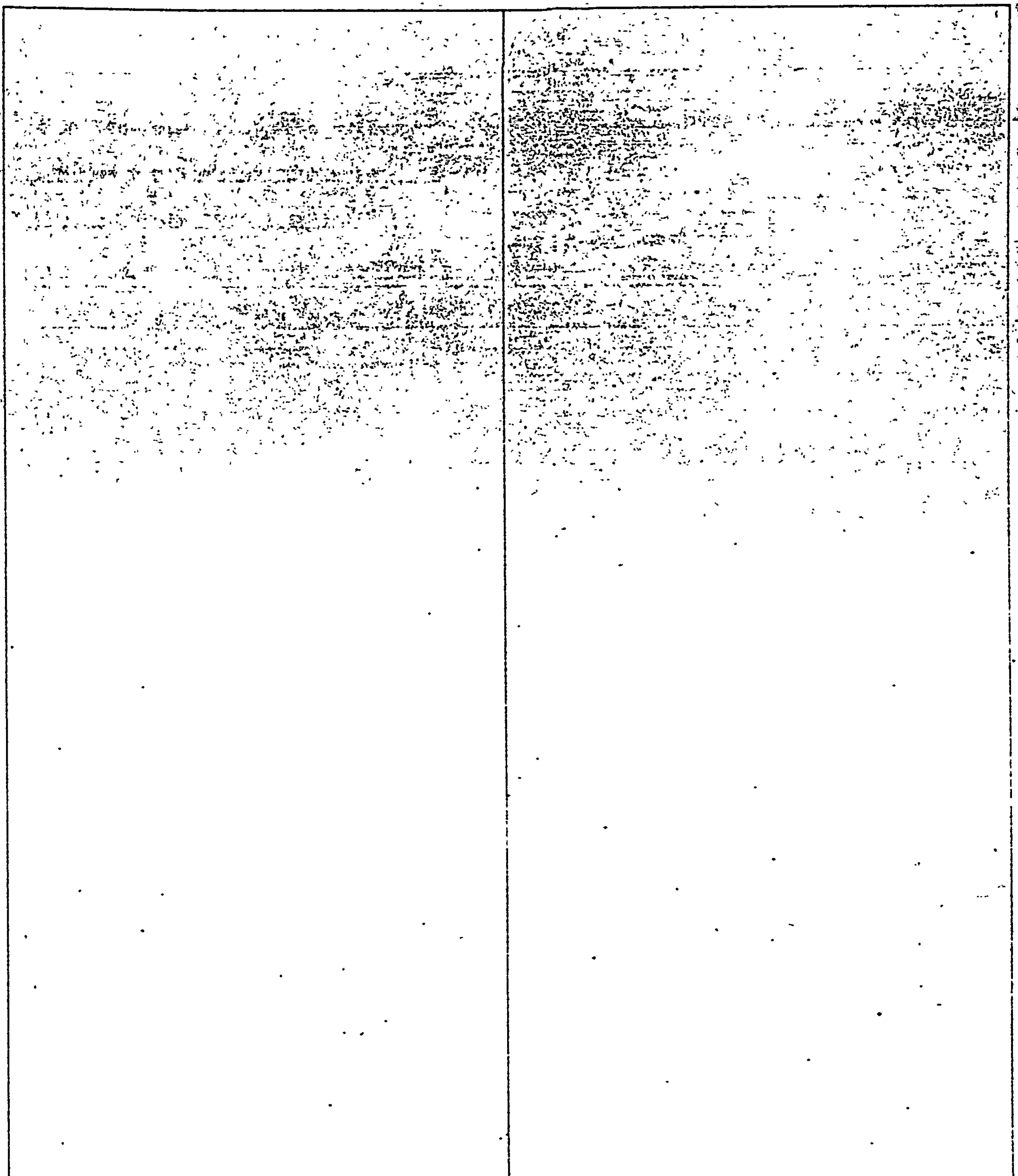

Fig. 25. Experimental temperature curves of resistance of silicon specimens with different injected contents of extrinsic phosphorus atoms: specimen A...$4.7 \cdot 10^{17}$ cm^{-3}; B ...$2.7 \cdot 10^{18} cm^{-3}$; C ... $4.7 \cdot 10^{19} cm^{-3}$; D ...$4.7 \cdot 10^{20} cm^{-3}$.

Fig. 26. Experimental temperature curves of current-carrier mobility in silicon specimens with different injected contents of extrinsic phosphorus atoms: specimen A ... $4.7 \cdot 10^{17}cm^{-3}$; B ... $2.7 \cdot 10^{18}cm^{-3}$; C ... $4.7 \cdot 10^{19}cm^{-3}$; D ... $4.7 \cdot 10^{20}cm^{-3}$.

Fig. 27. Experimental temperature curves of conductivity for germanium specimens with different injected contents of extrinsic antimony atoms.

[redacted]

STAT

CHAPTER V

THE INFLUENCE OF ELECTROMAGNETIC RADIATION (LIGHT) ON THE CONDUCTIVITY OF SEMICONDUCTORS

From the basic conclusion that the conductive state of a semiconductor is a state of excitation, it follows that any energy effect exerted upon its atoms which is capable of ejecting electrons from the valence band of the crystal or the impurity levels into the conduction band must influence the conductivity of the semiconductor. Along with the thermal mechanism of excitation of electrons, the mechanism of excitation by photons--i.e., by quanta of electromagnetic radiation--is of great practical and theoretical importance. This physical process of internal liberation of electrons is called the photoconductive effect and the additional conductivity which it produces is referred to as photoconductivity. It will be recalled that the energy of the photon is equal to $h\nu$, where h is Planck's constant = $6.62 \cdot 10^{-27}$ erg-sec, and ν is the frequency of the electromagnetic vibrations. In the process of electromagnetic excitation, one electron receives the total energy of one photon. Therefore photoexcitation is only possible if $h\nu \geqslant \Delta E_I$ for the excitation of impurity centers, or if $h\nu \geqslant \Delta E_0$ for the excitation of intrinsic conductivity. This fact distinguishes photoexcitation from the thermal excitation which may, in principle, occur at any temperature other than zero. The excess energy $h\nu - \Delta E$ is transferred to the excited electron in the form of kinetic energy which, as a result of frequent collisions of this electron with atoms, is rapidly reduced to a level corresponding to the average kinetic energy of the entire electron ensemble, which is in a state of

STAT

thermal equilibrium with the lattice. Since the activation energy for many semiconductors is expressed in fractions of an electron volt, the threshold frequency -- i.e., the minimum frequency ν_0 of the electromagnetic quantum which produces the photoconductive effect in such semiconductors -- is usually found in the infrared region of the spectrum. Even when the activation energy of a semiconductor reaches 1 ev, the threshold frequency, as determined by the relationship $\nu_0 = 1 \cdot 1.6 \cdot 10^{-12} / h$, is equal to $2.5 \cdot 10^{14} \text{sec}^{-1}$, which is still within the infrared band of the spectrum. It should be recalled that the frequencies of visible light range from $4 \cdot 10^{14} \text{sec}^{-1}$ to $8 \cdot 10^{14} \text{sec}^{-1}$, which corresponds to light-quantum energies from 1.5 to 3 ev.

The threshold frequency ν_0 of the photoconductive effect is called the red photoconductivity threshold. Electromagnetic radiations with frequencies $\nu < \nu_0$ are ineffective since the energy of a single photon of this radiation is insufficient to excite an electron. On the other hand even frequencies ν which exceed ν_0 become less and less effective with increasing departure from ν_0. This may be due to the rapid increase of the absorption coefficient $\nu > \nu_0$, with the result that the total electromagnetic energy incident on the body is absorbed in a thin layer close to the surface of the body, and the increase in the number of current carriers occurs only in the thin surface layer. Such an increase in number of current carriers may exert a slight influence on the total conductivity of a massive body both because the rate of recombination of electrons and holes is higher at the surface than in the volume, and because of the intensified recombination in the volume due to diffusion of secondary current carriers into an impurity semiconductor.

A negative photoeffect which is accompanied by a decrease, not an increase, in the conductivity of a body exposed to radiation is sometimes observed and is explained as a consequence of this intensified recombination. Capture of current carriers of the opposite sign by the impurity centers is believed to be the cause

STAT

of the diffusion of current carriers with only one sign deep into the body. These phenomena of capturing ("sticking") of current carriers have not been experimentally demonstrated in all cases, but are still often referred to because they support concepts of the complex mechanisms of the photoconductive effect which aid in explaining the complex relationships observed in experiments in the study of photoconductivity. The future development of this branch of semiconductor physics will show to what extent these concepts are true. Thus far, the experimental fact is that the frequency range of photoconductivity is comparatively very narrow. The situation which prevails in this respect is shown schematically in Fig. 28. In this figure the wavelength $\lambda = \underline{c} / \nu$, where c is the velocity of light, is plotted against the x-axis rather than the frequency ν. The absorption coefficient 1 and the photoelectric current 2 are plotted against the y-axis. It is seen from the curves in Fig. 28 that the photoconductivity of nonmetallic crystals is observed to be within a comparatively narrow spectral region at the long-wave edge of the continuous-absorption band of the crystal. The curve which corresponds to the photoeffect in the impurity centers is not shown on the schematic graph in Fig. 28; it would have formed one or several smaller peaks in the long-wave band of the spectrum.

Since the experimental curves have extended and unclear extremities and are not as distinct as the schematic curves in Fig. 28, a wavelength situated on the right-hand descending arm of the photoconductivity curve and corresponding to the half-value of photosensitivity ($\lambda_{\frac{1}{2}}$ in Fig. 28) is taken as the threshold wavelength which determines the activation energy. The activation energy $\Delta \underline{E}_0^{ph}$ determined from optical data may be compared with that determined from the temperature curve of conductivity, $\Delta \underline{E}_0^{t}$ (Fig. 20).

There are theoretical considerations according to which the above two methods of measuring the forbidden-band width may, in some cases, produce numerically different results. We shall discuss these considerations at somewhat greater length,

STAT

for they involve important problems of semiconductor physics.

The first consideration proceeds from the specific properties of the ionic lattice. In this case theory leads to the conclusion that $\Delta \underline{E}_o^{ph} > \Delta \underline{E}_o^{t}$ and explains this discrepancy on the basis of polarization of the ionic lattice. Let us take into account the fact that an electron is displaced into space upon excitation. A displacement of the surrounding electrons and nuclei into new equilibrium positions occurs as a response, with the result that the energy of the crystal diminishes. In the atomic lattice, this polarization process is effected solely by displacement of electrons, which is very rapid. The polarization process in the ionic lattice is due to both the rapid displacement of light electrons and the comparatively slow displacement of the heavy ions. The energy of optional excitation in an ionic crystal is determined by the difference in energy between the two states of the crystal: the nonexcited initial state and that after polarization -- i.e., after displacement of the ions into new equilibrium positions. Therefore the energy $\Delta \underline{E}_o^{ph}$ of optical excitation is greater than the energy $\Delta \underline{E}_o^{t}$ of the thermal excitation by the quantity of energy which the lattice receives as a result of the displacement of ions. $\Delta \underline{E}_o^{ph} = \Delta \underline{E}_o^{t}$ in atomic crystals, since the polarization of the medium, which is effected by the displacement of electrons only, keeps up with the process of excitation.

Fig. 28. Illustrative spectral distribution curves for optical absorption 1 and photoconductivity 2 in crystals.

The experimental fact which we are now discussing may become more comprehensible if explained in the following manner. The phenomenon of the electron excitation in an atomic crystal is distinguished from that in an ionized crystal

[redacted]

STAT

by the absence of the energy of polarization displacement of ions in the energy balance of the atomic crystal, which is present in the energy balance of the ionic crystal; this energy is converted into thermal energy of the crystal in both optical and thermal excitation mechanisms. Therefore if the excitation in an ionic crystal is caused by a source of thermal energy, the final excitation energy expended by the source is reduced by the amount of energy returned to the source by ionic polarization. In the case of the optical excitation mechanism, the electromagnetic quantum must furnish the electron undergoing excitation with the required quantity of energy in full.

The difference in the numerical values of $\Delta \underline{E}_o^{ph}$ and $\Delta \underline{E}_o^{t}$ for the substance may be linked to yet another circumstance. The fact is that optical electron transitions in a solid (as in individual atoms) obey specific laws of selection which, in some cases, do not permit transition from the valence band to the conduction band by the shortest energy route; in such cases, the only possible transitions are those occurring between levels at some distance from the energy-band boundaries. This circumstance, which applies equally to both ionic and atomic crystals, may also give rise to an inequality of the type $\Delta \underline{E}_o^{ph} > \Delta \underline{E}_o^{t}$. No such restrictive rules apply for the thermal mechanism of excitation, and the thermal transfer of electrons can proceed along the shortest energy route between bands.

Finally, a third and last cause of possible discrepancies between the experimental values for $\Delta \underline{E}_o^{ph}$ and $\Delta \underline{E}_o^{t}$ may arise from the excitonic mechanism of excitation in the photoconductive effect, which will be discussed on page 108. We indicate here only that the discrepancy in experimental data on the width of the forbidden band which results from this is expressed by the converse inequality

$$\Delta \underline{E}_o^{ph} < \Delta \underline{E}_o^{t}.$$

STAT

It should be noted, however, that existing experimental data from determinations of the forbidden-band width in atomic and ionic crystals by optical and electrical methods reveal noticeable divergency between ΔE_o^{ph} and ΔE_o^{t} only in rare cases.

The next problem to be dealt with is the establishment of the dependency of the equilibrium concentration of the current carriers excited by electromagnetic radiation on the intensity of the latter. Theory and experiment indicate that the result depends on whether the photoconductivity constitutes a small addition to the conductivity due to thermal excitation (i.e., to the dark conductivity) or is the primary generator of current carriers in the crystal. The equilibrium concentration of current carriers results from two competitive processes -- excitation, which depends on the radiation intensity L and the reverse process of recombination, which depends on the total equilibrium concentration. When the dark conductivity of the semiconductor is small and can be disregarded, the equilibrium concentration n_{ph} is determined by a formula of the type:

$$n_{ph} = \alpha\sqrt{L}\ \mathrm{cm}^{-3},$$

where α is a coefficient of proportionality which depends on the photoelectric activity of the light and the rate of recombination of current carriers.

In the other case, in which the number of current carriers created by the photoconductive effect is small as compared to the number of dark current carriers, the additional concentration at equilibrium is determined by the expression

$$n_{ph} = \beta L\ \mathrm{cm}^{-3}.$$

Experiments indicate that the mobilities of both the photocarriers and the dark carriers are equal. From this it follows that the photoconductivity as determined by the product $\sigma_{ph} = e n_{ph} u_{ph}$ is proportional either to the square root of the radiation intensity when the photoconductive effect is the principal factor

[redacted]

STAT

in the excitation of current carriers, or to the radiation intensity itself when photoconductivity constitutes a small addition to the dark conductivity. It should be noted, however, that these laws are not universal; for some semiconductors the dependence of photoconductivity on radiation intensity is expressed either by a different exponent or is of more complex character in general. As was noted earlier, some semiconductors even show a decrease rather than an increase in conductivity when exposed to radiation in a certain temperature range (negative photoeffect).

Experimental studies indicate that not only the mobility of the photocurrent carriers but also the sign of their charge is the same as that of the dark carriers. This is natural in the case of electromagnetic excitation of current carriers from impurity levels. However, it is also found that when electrons are excited from the valence band into the conductivity band -- in which case, as we know, two current carriers are formed (an electron and a hole) -- the charge carrier which participates in dark conductivity usually possesses mobility in the crystal lattice. This experimental fact is not yet accounted for; it is unclear why the carrier with the opposite sign becomes trapped, as it were, in the lattice very shortly after formation and does not participate in the current. The fact that such carriers actually arise in the photoconductive effect is confirmed by our experiments with secondary radiation: additional carriers with the opposite sign appear when a semiconductor previously excited by photons which transfer electrons from the valence band into the conduction band is subsequently exposed to radiation by photons whose wavelength is such that they are unable to effect the above transfer ($h\nu < \Delta E_0$). These additional current carriers do not appear when such a crystal is exposed to radiation by the latter photon type only. Hence we conclude that the secondary exposure to radiation does not create new current carriers but merely liberates those charges which have been trapped in the various secondary "capture" levels formed by impurities and crystal defects in the first stage of excitation.

[redacted]

STAT

The dependence of the intensity of the photoelectric current on the strength of the external electric field acting on the photoconductive substance (the so-called current-voltage characteristic) has its own characteristic properties. The primary photocurrent created by those charges which are directly activated by light conform to Ohm's law only in a low-field region with an upper boundary numerically different for different photoconductive substances. As the field increases beyond this boundary, the primary photoelectric current increases at a rate slower than that prescribed by a linear relationship, with a tendency to saturation in strong fields which is illustrated by the theoretical curve in Fig. 29.

Physically, this phenomenon stems from the mechanism of generation of the primary photoelectric current itself. Charge carriers activated by light migrate at random in the crystal in the absence of a field during the period τ_o of their free existence, and then recombine or are captured by impurities or defects of the crystal. If a weak external field E is applied to the body exposed to radiation, the photocurrent carriers drift in the field during the period from the moment of excitation to the moment of their recombination, undergoing displacement toward the electrodes through a distance d, which is determined by the product $d = uE\tau_o$, where u is the mobility. The quantity d is called the drift path of the current carriers and τ_o their lifetime (see page 38). The strength of the photocurrent, which is proportional to the number n_o^{ph} of primary photocarriers liberated by light in unit time and to their mean drift d ($I_{ph} \sim n_{ph}^{o} d$) in the field, will increase linearly with the field only until the drift path attains the magnitude of the interelectrode distance D. The increase of the primary photocurrent with increasing field strength slackens as the drift of the individual current carriers begins to experience restriction by the distance to the electrode, and is limited by a certain value I_{max} equal to the total charge of all current carriers liberated by the light per unit time over the entire crystal and

STAT

transported to the electrodes by the field: $I_{max} = en_{ph}^{o}$.

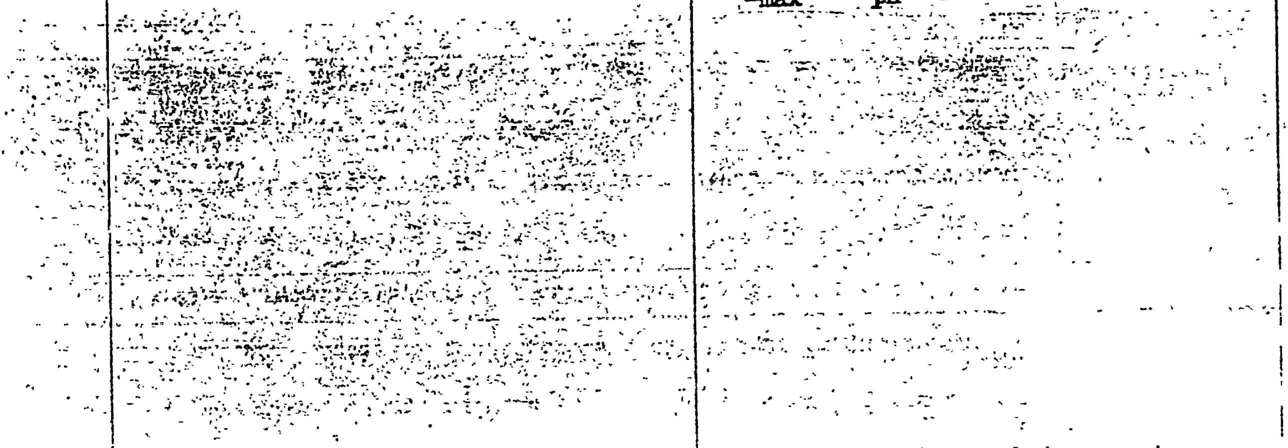

Fig. 29. Current-voltage characteristic of primary photocurrent.

This picture is observed in pure form only in nonconductor crystals. Thus the saturation point of photocurrent was attained in a 15,000-v/cm field in the diamond specimens investigated and in a 2,000-v/cm field in AgBr. The current-voltage characteristics of many typical semiconductors exposed to light differ considerably from the theoretical curve shown in Fig. 29. Two circumstances are responsible for this departure: first, the fact that the above limiting condition ($I_{ph} < I_{max}$) concerns only the primary photoelectric component of the total current and not its dark component (which considerably exceeds the former in many semiconductors) and, second, the fact that various secondary phenomena which distort the above mechanism of formation of the photocurrent appreciably are superimposed on the primary photocurrent in semiconductors.

The fact that photoconductivity is observed both in semiconductors and insulators confirms the correctness of the consistent concept of electroconductivity for all solids which has been established by the band theory. The experimentally observed differences in the manifestations of photoconductivity in semiconductors and insulators arises basically from the fact that, in the former, the photocurrent often constitutes only an inconsiderable addition to the dark current. It is an empirically established fact that photoconductivity does not occur in all pure semiconductors and insulators, but only in those with dielectric constants ϵ greater than 4. The photosensitivity may be imparted to substances which fail to meet this condition by adding impurities to them. In addition to impurities, the condition of a solid's surface (e.g., its roughness) also influences its photosensitivity, especially in cases in which the absorption of light occurs in a thin surface layer. This is explained by the fact that the condition of the surface of a body strongly affects the rate of surface recombination -- i.e., the rate of the process which returns the current carriers to their nonconductive state.

Also of interest is the question of the number of current carriers which correspond to 1 absorbed quantum (photon); this value is called the quantum yield of the photoconductive effect. In the majority of cases the quantum yield within the photoelectrically active band of the electromagnetic spectrum is equal to unity -- i.e., each photon engenders one pair of current carriers in the excitation of the basic lattice. This is shown clearly, for example, in the case of germanium, where each photon with a wavelength from 1.0 to 1.8 μ forms a single pair consisting of one electron and one hole. Sometimes, however, comparatively small amounts of radiation produce intensive photoconductivity in experiments with certain substances, and the quantum yield is found to exceed unity. According to the general concept of the mechanism of the photoconductive effect which was developed earlier, each primary excitation event in the intrinsic lattice can create only one pair of current carriers. To account for experimental facts involving a quantum yield which exceeds unity, therefore, various secondary processes are assumed to arise from the imperfection of the crystal after excitation of primary current carriers. As an example, the very strong photoconductivity effects observed in fine-grained crystalline specimens of some semiconductors are explained by the presence of insulating, oxidized intercrystalline layers whose conductivity changes sharply on the appearance of the primary photoelectrons. Experiments indicate that the influence of light on such a semiconductor manifests itself in a sharp decline in

STAT

the resistance of the specimen, although the conductivity of the material of the semiconductor itself changes inconsiderably inside the crystals. Experiment indicates that secondary processes do actually occur in the photoconductive effect. The secondary processes are distinguished from the primary by the noticeable but sometimes quite small time lag in their development. Therefore an ultra-short-pulse irradiation technique which prevents the development of the secondary processes is used in experimental separation of these processes.

The quantum yield is found in some cases to exceed 1 due to the fact that the photoelectric activity of the light extends into the frequency range in which the energy of a single photon $\underline{h}\nu$ is several times greater than the excitation energy of the current carriers $\Delta \underline{E}_0$. The primary current carrier in this case proves capable of exciting additional secondary current carriers by virtue of the excess energy $(\underline{h}\nu - \Delta \underline{E}_0)$ which it acquires on absorption of a photon.

It was indicated in the discussion of Fig. 28 that in the majority of cases only a comparatively narrow range of the electromagnetic spectrum at the long-wave edge of the continuous-absorption band of a substance is photoelectrically active. This means that photons which do not belong to this narrow range do not excite current carriers even when absorbed by a solid and possessed of an energy sufficient to accomplish an excitation event $(\underline{h}\nu > \Delta \underline{E}_0)$. This photoelectrically inactive absorption of light has suggested to physicists the idea of the existence of a certain mechanism of excitation which produces no current carriers. However, it was recently proposed that this mechanism of excitation also participates in photoelectrically active excitation -- i.e., that the formation of current carriers under the influence of light proceeds in two stages. During the first stage a photon transfers an electron not directly into the conduction band but into an excited level located just below the bottom of the conduction band. An electron in this state does not completely escape the attraction of its atom, but remains related to the hole which it has left behind in the latter. Such an excited state of an atom is

STAT

called an exciton and may be transmitted from one atom to another -- i.e., migrate at random within the atom. It must be understood here that in transmission from one atom to another, the excited state returns the first atom to its normal unexcited state. The application of an external electrical field has no influence on the random movement of excitons in the crystal lattice and therefore does not create a current, since the exciton as a whole constitutes a neutral combination of an electron and a hole and the bond cannot be broken by a low external field. According to the above concept, an exciton can "break" into two current carriers (an electron and a hole) or recombine -- i.e., return the atom to its normal state -- only upon collision with impurity centers. The former (i.e., the second stage of the photoconductive effect) occurs at the expense of the thermal energy which it is necessary to impart to the exciton to effect the break -- or, in other words, the energy which is necessary to transfer the electron from the exciton level to the conduction band. The latter (i.e., the recombination of the exciton) is accompanied either by emission of an electromagnetic quantum or -- which is more frequently the case -- by transfer of the exciton energy to the lattice in the form of heat. An electron and a hole are formed after the second stage of the photoeffect, but one of these current carriers may prove ineffective due to its low mobility or capture by impurity centers. This, briefly, is the physical essence of the modern theory of excitons.

The above enables us to understand one of the causes of the phenomenon which was discussed on page 101. It may be seen that the explanation reduces to the fact that when the photoconductive effect is due to the excitonic mechanism of excitation, the threshold excitation frequency ν_0 may be below $\Delta E_0/h$ and, consequently, the forbidden-band width (ΔE_0^t) determined by the optical method as ΔE_0^{ph}, may be found smaller than ΔE_0^t. It should be noted that the excitonic concept has received experimental confirmation in the absorption lines recently observed in the long-wavelength edge of the fundamental absorption bands of some

[redacted]

STAT

semiconductors (i.e., at the right end of the absorption curve, Fig. 28).

How does variation in the temperature of a body influence photoconductivity?

The relative importance of photoconductivity must, naturally, increase with decreasing temperature, since the dark conductivity which serves as the background against which the photoconductivity manifests itself declines with decreasing temperature. Theory and experience indicate that the absolute magnitude of photoconductivity increases with decreasing temperature. This is explained by the fact that as the concentration of dark current carriers -- the electrons and holes -- decreases, the probability of recombination of current carriers is also reduced.

Variation in temperature also influences the red photoconductivity threshold ($\lambda_{\frac{1}{2}}$ in Fig. 28); here, the algebraic sign of the shift is found to be different for different semiconductors -- i.e., some substances show a displacement of the threshold to the right with decreasing temperature, while in others it shifts to the left. This experimental result is explained as due to constriction of the forbidden band in some semiconductors and its expansion in others with increasing temperature (see page 37).

Characteristic of the photoelectric effect of electromagnetic radiation is the fact that the change in the electrical properties of semiconductors exposed to radiation is of temporary nature (with the exception of the case where γ-quanta induce artificial radioactivity). This means that the pre-irradiation value of the dark conductivity is recovered more or less rapidly in the majority of cases after irradiation ceases; some semiconductors require microseconds for this process, while in others it requires minutes and even hours. This circumstance is of extreme importance for a whole series of technical applications of the phenomenon of photoconductivity, since high specifications are sometimes imposed on devices of this class (photoresistors) with respect to their response inertia.

The most suitable materials for photoresistors are selenium, germanium, and silicon, as well as sulfides and selenides of cadmium, lead and bismuth. Thin layers of PbS, PbSe and PbTe are photosensitive in the far infrared band of the spectrum extending to 4 μ . This photosensitivity boundary moves in the direction of longer wavelengths with decreasing temperature.

CHAPTER VI

THE INFLUENCE OF STRONG ELECTRIC FIELDS ON THE CONDUCTIVITY OF SEMICONDUCTORS

It follows from the physical considerations by which we substantiated Ohm's law see Formulas [(1) and (2)] that this law must apply only if the current carrier mobility and concentration are independent of the electric-field strength. Actually, the specific conductivity σ is not a function of the field strength E in this case, and the current is proportional to the field in agreement with Ohm's law.

Let us discuss the physical significance of these two premises. Let us begin with mobility. According to (2), $u = e\tau/m$, so that the dependence of the mobility u on the field E may manifest itself only through the free time τ. The value $\tau = l/v$ itself begins to depend on the field when the increase in the velocity v over the free path l becomes comparable to the velocity -- i.e., when $\overline{\Delta v}$, as determined by Formula (1), is not negligibly small as compared to v. The quantity τ is both an explicit (denominator) and implicit function of v, since the numerator, the length of free path l, also depends on v in some important scattering mechanisms. Consequently, the physical premise of the field-independence of mobility is reduced to one of negligibly small influence of the field on the velocity. On the other hand, however, it follows from the considerations on which Ohm's law is based that the influence of the electric field consists precisely in a change in velocity, and it might appear that no explanation of the occurrence of the current could be found if this change were completely disregarded. How to account for this? The fact is that the velocity of a material particle is a vector quantity, and the el STAT

[redacted]

field acts on this quantity in two ways: it changes both the absolute magnitude of the velocity and its direction in space. In the low-field range in which Ohm's law is valid, the change in the magnitude of electron velocity over the free path is small by comparison with the absolute magnitude of the initial velocity of the electron in its free path, and may be neglected in this case. The other effect of the field -- the change of the direction of electron velocity -- suffices to explain the occurrence of the current and substantiate Ohm's law.

In the absence of an external field, the directed average velocity found by geometrical summation of the velocities of the individual electrons is equal to zero (see Fig. 2). Application of an external field produces a certain directed resultant velocity $\overrightarrow{\Delta \underline{v}}$ which, though small by comparison with the absolute average velocity of the electrons, appears distinctly enough against the background of the zero directional velocity which prevails in the absence of the field.

It can be imagined that with increasing field strength, the change in the absolute value of the velocity over the free path of an electron becomes, at sufficiently high fields, comparable to the magnitude of the initial free-path velocity.

An increase in the average random energy of electrons represents, essentially, an increase in their average thermal energy measured in degrees of the temperature scale. This means that in a strong enough field the steady-state temperature of the electron gas may noticeably exceed that of the lattice. As a result, the transfer of energy by current carriers to the lattice increases, and a stationary condition is established in which the total energy received from the external field is transmitted by the current carriers to the lattice. Since the thermal velocity of current carriers increases in this case in proportion to the square root of the field strength, the mobility of semiconductors with the length of free path independent of velocity must vary according to the law $\underline{u \sim 1/\sqrt{E}}$ [see Formula (1)]. For these fields, which we refer to as strong fields, Ohm's law no longer holds,

[redacted]

since a noticeable increase in the velocity over the length of the free path gives rise to an appreciable dependence of both the free time τ and the mobility u on the field strength E; this results in dependence of the conductivity σ on E -- i.e., a departure from Ohm's law. It follows from the relationship $\tau = l(v)/v$ that τ and, consequently, the mobility u, may either decrease or increase with increasing field strength E. Thus, τ and, consequently, u must decrease with increasing field strength (causing a marked increase in v) when the free path l does not depend on the velocity v, as is the case in the thermal mechanism of scattering in atomic crystals. But τ and, consequently, u must increase with increasing field strength when the free path $l \sim v^4$, as is the case in the scattering of current carriers by ionized impurities.

Thus a departure from Ohm's law due to dependence of mobility on the field must occur when the external field causes a perceptible increase in the velocity of electrons. The minimum field strength which causes the departure from Ohm's law to become perceptible is called the critical field E_c. What are the substances in which this departure should occur earliest? Since $\overline{\Delta v} = uE$, a noticeable increase in velocity with increasing field will manifest itself earlier in substances with greater mobility than in those with lower mobility. The mobility of electrons in germanium at room temperature, for example, is $u = 3900$ cm^2/v.sec, so that even a field with an intensity of 10^3 v/cm should produce in it a noticeable increase in velocity over the free path. The relative change in velocity in this case amounts to

$$\frac{\Delta v}{v} = \frac{uEl}{v} = \frac{3900 \cdot 10^3}{3 \cdot 10^7} = 13\%.$$

[redacted]

Fig. 30. Experimental curves illustrating departure from Ohm's law due to change in mobility of current carriers in germanium in strong fields.

Experiment has actually shown that a perceptible departure from Ohm's law may be observed in electronic germanium at as low a field as 900 v/cm at room temperature (see Fig. 30). The critical field of hole-type germanium at room temperature is somewhat stronger -- 1400 v/cm; this is in accord with the fact that the mobility of holes in germanium ($\underline{u}$ = 1900 cm^2/v.sec) is lower than that of electrons. The experimental results for silicon may be considered a confirmation of this rule for the dependency of critical field upon mobility. Thus for electronic silicon E_c = 2500 v/cm ($\underline{u}$ = 1200 cm^2/v.sec); for hole-type silicon E_c = 7500 v/cm ($\underline{u}$ = 500 cm^2/v.sec). It follows further from this rule that the critical field should decrease if mobility increases with decreasing temperature. Indeed, Fig. 30 shows that the critical field in electron-type germanium falls to 120 v/cm at T = 77° K. In all of the above cases, the departures from Ohm's law manifest themselves in transitions from the direct proportionality $\underline{I} \sim E$ with $E < E_c$ to the theoretically supported law $I \sim \sqrt{E}$ with $E > E_c$. The latter corresponds to the change in the mobility which occurs, according to the law $\underline{u} \sim 1/\sqrt{E}$, upon a change in field. It is seen from Fig. 30 that the experimental current-voltage characteristic is saturated when $E > 3500$ v/cm, indicating the transition to the dependency $u \sim 1/E$, in which the current no longer depends on field strength. At a field strength $E = 40{,}000$ v/cm, the mobility of electrons in germanium at room temperature is one-tenth of the corresponding value in weak fields.

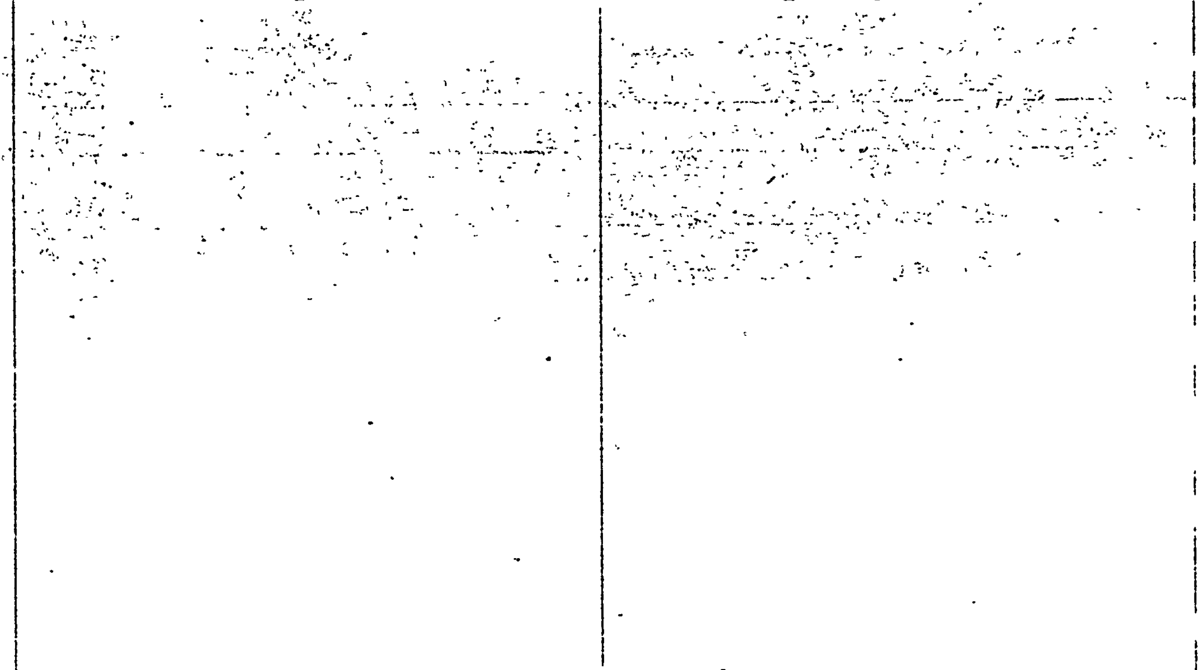

Fig. 31. Experimental curves showing current-carrier mobility as a function of electric-field strength at low temperatures.

We have already explained why the mobility u of atomic semiconductors such as silicon and germanium should decrease with increasing field in the strong-field region in the case of the thermal scattering mechanism. The mobility must increase with field strength at temperatures sufficiently low so that scattering in atomic semiconductors occurs primarily at the ionized impurities, and the free path is $l \sim v^4$. Experimental results for electronic germanium indicate that at $T = 20^\circ$ K, E_c amounts to a mere 8 to 10 v/cm and the mobility increases instead of decreasing with increasing field strength (see Fig. 31). As the velocity and the mobility of electrons noticeably increase, the thermal mechanism of dispersion becomes effective, in which case $u \sim 1/\sqrt{E}$; this accounts for the further decreases in mobility with $E > 20$ v/cm. However, there exists another body of opinion as

[redacted]

STAT

to the treatment of these last experimental facts (Fig. 31) which maintains that the observed departure from Ohm's law is due to violation of its second premise -- i.e., to nondependence of the current carrier concentration on the field. Let us now turn to this problem.

Fig. 32. Curve illustrating the influence of strong fields on conductivity with mechanism of secondary increase in current carrier concentration due to the electric field.

Theory predicts that when the intensity of the field in a pure semiconductor exceeds 10^4 to 10^5 v/cm additional current carriers produced by the field should appear in it. Experiments have actually shown that in the range of fields with $E = 10^4$ to 10^5 v/cm and sometimes even with $E \cong 10^3$ v/cm, the current-carrier concentrations of many semiconductors show a noticeable increase with increasing field, so that the current rises at a much faster rate than that required by Ohm's law.

The experimental data may be divided into two groups. The dependence of the conductivity on the field for those in the first group is characterized by Pool's* empirical law $\sigma = \sigma_0 e^{\alpha (E - E'_c)}$ (see Fig. 32), where $\sigma_0 = c e^{-\Delta E_0/2kT}$, α is a coefficient which depends upon temperature and usually decreases as the latter increases, and E'_c is the critical field, which depends on the nature of the semiconductor, temperature and the impurity concentration; the dependency of current on field for the second group may be represented in the form of the Frenkel's theoretically-derived law $\sigma = \sigma_0 e^{\beta\sqrt{E}}$.

What then, is the physical essence of this phenomenon -- i.e., the mechanism of the influence of the electrical field which is responsible for the increase in

---

* Translator's note: Name transliterated (Pul'). Spelling not verified.

current carrier concentration? It develops that this phenomenon involves not one, but as many as four different mechanisms, which characteristically become operative not simultaneously but at different critical fields. We shall concern ourselves only with the two mechanisms which produce stationary states without causing decomposition of the substance, since the other two (electrostatic ionization and Stark's effect), according to theoretical estimates, first become effective in the region of fields $\sim 10^7$ v/cm, when breakdown due to the first two mechanisms has already occurred.

With increasing field strength, the mechanism of Frenkel's thermoelectronic ionization is the first to become effective. The action of this mechanism will become clear when we consider that the external electric field acts on the electron with a force eE and changes its energy state in the atom. As a result, the energy expenditure necessary to excite an electron into the conductive state is reduced by a quantity $2e\sqrt{eE/\epsilon}$, while the probability of thermal excitation increases correspondingly by a factor $\exp(2e\sqrt{eE/\epsilon}$ which results directly in an increase in current-carrier concentration in accordance with Frenkel's law, above. A more exact theory of this phenomenon gives a dependency of the type of Pool's law for conductivity in comparatively weak fields and of the type of Frenkel's law for stronger fields. Thus the action of the thermoelectronic ionization mechanism reduces to facilitation of the escape of the electron from the atom; the separation itself is effected by thermal excitation.

To understand the physical essence of the second mechanism, which, following the mechanism of thermoelectronic ionization, becomes effective in stronger fields, it should be taken into consideration that a free electron undergoing acceleration on its free path, may, in a strong enough field, acquire an energy sufficient to excite either a bound electron, an impurity atom, or an atom of the basic lattice. In accordance with the physical essence of the phenomenon, this mechanism of excitation of secondary current carriers is called impact ionization. The mechanism

STAT

of impact ionization may bring about an increase in current-carrier concentration only under the condition that the ionizing electron, after exciting a bound electron, itself remains excited -- i.e., in the conductive state. Clearly, this can only occur when the kinetic energy of the ionizing electron is so great that it retains (after effecting the excitation) an energy sufficient to keep it in the conduction band -- i.e., when the ionizing electron is displaced only from a higher to a lower level within the limits of the conduction band in the process of impact ionization. This, in turn, requires that the conductivity-zone width exceed the forbidden-band width when electrons in the valence band are excited. Similar conditions may also be formulated for the impact ionization which may be produced in the valence zone by a hole. Since the activation energy of impurities is usually less than the forbidden-band width, the excitation of current carriers in an increasing electric field occurs first from impurity levels, following which the impact ionization of basic atoms of the lattice becomes effective.

The above concept of the mechanism of impact ionization may give rise to the misgiving that it cannot achieve a stationary state and must always lead to breakdown. Indeed, does not an avalanche-like process of formation of steadily increasing numbers of current carriers -- a process which should lead to breakdown -- occur at a given high field strength when conditions are such that the ionizing electrons, after exciting secondary electrons, themselves remain in the conduction band? This would be the case if the process of excitation of current carriers were not concurrent with another opposed and stabilizing process -- that of recombination of current carriers. The formation of secondary current carriers by impact ionization is accompanied by either additional free impurity levels or additional holes in the valence band; the excited electrons recombine with these with a probability which increases with their number. As a result of the competition between the excitation and recombination processes, some increased concentration of current carriers is established as a result of impact ionization, and increases

STAT

with the field but remains unchanged in a given strong field.

When the field is strong enough (usually at E $> 10^5$ to 10^6 v/cm) the process of impact ionization assumes an avalanche-like character and leads, of course, to a breakdown.

Since great quantities of energy which may considerably or even dangerously heat it are developed in the test specimen in strong fields, experimental studies of the effects of strong electrical fields on the conductivity of semiconductors are often carried out using the transient-pulse method.

Let us note in conclusion that the conductivity increase in semiconductors in strong electric fields is a phenomenon utilized in technical devices for the protection of electric transmission lines from overvoltage.

STAT

CHAPTER VII

THE INFLUENCE OF VARIOUS TYPES OF CORPUSCULAR RADIATION ON THE CONDUCTIVITY OF SEMICONDUCTORS

There are four reasons for the change in conductivity of semiconductors under bombardment by streams of high-speed electrons, neutrons, deuterons and α-particles.

1) Corpuscular radiation excites (ionizes) atoms, with the result that the concentration of current carriers increases. This excitation is, in general, of the same ephemeral character as in the case of the photoconductive effect, and vanishes rapidly after withdrawal of the radiation.

2) Bombarding particles cause displacements of ordered atoms (or ions), with the result that stable and unstable defects of two types appear in the lattice: vacancies and interstitially implanted ions. Like chemical impurities, these defects alter the energy spectrum of electrons, causing a change in the equilibrium concentration of current carriers which arises from thermal excitation. The work required to remove an atom from a germanium lattice point amounts to about 25 to 30 ev, so that the displacements of atoms from the lattice points during bombardment of this semiconductor by heavy high-speed particles are effected not only by the bombarding particles themselves but also by those primarily displaced atoms which have received energy sufficient for secondary displacement of other atoms in the lattice. However, the number of stable secondary current carriers dislodged by one bombarding particle does not exceed 4 or 5, although the total number of atoms displaced may amount to several hundred. This is explained by the fact that defects of the lattice are for the most part quickly corrected after radiation

STAT

ceases. In many cases, even complete restoration of the original state of the specimen may be effected by means of appropriate annealing.

3) Under corpuscular radiation, as well as when exposed to the γ-rays of the electromagnetic spectrum, radioactive transformations in the body exposed to radiation may occur, resulting in the appearance of stable atoms or ions of elements which were absent in the original substance. This development is similar in its results to the process of injecting chemical impurities, which is often used to impart definite, specified electrical properties to a semiconductor. Since the injection of chemical impurities is a very important procedure in semiconductor technology, it is likely that techniques by which impurity centers are created by various types of corpuscular radiation, which have their own peculiar advantages, will compete successfully with the chemical method and find important technical application.

4) Defects of the lattice and impurities in the form of foreign elements created by the action of corpuscular radiation reduce the current-carrier mobility (and, consequently, the conductivity) insofar as mobility is determined by scattering at impurity centers.

After thus generally characterizing the effects of all kinds of corpuscular radiations, let us discuss in somewhat greater detail those properties which are characteristic of certain particular forms of corpuscular radiation. Further discussion of this problem will be chiefly and necessarily concerned with germanium and silicon, since up to the present time these semiconductors had been studied in the majority of experiments made on the action of corpuscular radiation.

<u>Electrons</u>. It has already been indicated that a work of 25 ev per atom is required to create stable effects resulting from displacement of atoms from the lattice points of germanium. Calculation indicates that an electron may transmit such an energy to an atom in an elastic collision only if its initial energy exceeds 0.55 Mev. If the energy of the bombarding electrons is less than this

[redacted]

quantity, their effects on the semiconductor are reduced, as in the case of the photoconductive effect, to a temporary change in the current-carrier concentration which vanishes when radiation ceases. Experiments with the exposure of germanium to radiation by medium-energy electrons (< 0.55 Mev) have confirmed this conclusion and indicated that conductivity increases during the exposure and is restored to its initial value in $\sim 10^{-5}$ sec upon withdrawal of the radiation. This time interval is hundreds of times smaller for other semiconductors. The working principle of electron- and α-particle counters of the crystal type (diamond, CdS, ZnS), which offer convenience in operation and provide high amplifications (to 1,000) but still possess a number of serious shortcomings, is based on this phenomenon of temporary change in conductivity upon exposure to radiation.

When specimens of electron-type germanium are exposed to radiation by high-speed electrons with an energy of 1 Mev, the change in their electrical properties takes a course different from that in specimens of hole-type germanium. The stable conductivity of n-type germanium, i.e., the conductivity established after suspension of radiation, decreases as the radiation begins and then passes through a minimum and increases, as radiation continues, to a value which may exceed its initial value. However, n-type germanium is converted into the p-type modification in the process of these changes in conductivity. This effect of high-speed electrons on n-type germanium stems from the acceptor character of the lattice defects created by the bombardment. The small number of acceptors produced in n-type germanium during a short exposure to radiation capture electrons from the conduction band and the donor levels, thus lowering the conductivity of the irradiated specimen; with a further irradiation, the number of acceptors increases, with the result that the number of conducting electrons in the conduction band decreases and becomes practically negligible. At the same time, holes appear in growing numbers in the valence band as the exposure to radiation continues. Finally the electron-type germanium, having been converted into the hole type after a long

STAT

enough exposure, acquires current carriers (holes) in concentrations which increase with the duration and intensity of its bombardment by the high-speed electrons. The condition of thermal equilibrium $n \cdot p = n_i^2$ between the concentrations of electrons and holes [see Formula (9)], where n_i is a constant of equilibrium which depends on temperature for a given substance, applies for all stages of the process in which electron-type germanium is converted into the hole type. For germanium, $n_i^2 = 3.6 \cdot 10^{26} \mathrm{cm}^{-6}$ at T = 300° K.

When hole-type germanium is exposed to radiation by high-speed electrons, its stable conductivity increases from the very beginning of the exposure. This experimental fact is in full accord with the conclusion that the impurity centers created in germanium during exposure to radiation by fast electrons are of acceptor character.

Experiment also confirms that current-carrier mobility continues to decrease with longer and more intensive bombardment in both hole- and electron-type germanium. However, this has only a slight effect on the character of the change in conductivity, since the relative change in concentration which occurs when pure germanium is exposed to radiation far exceeds the relative change in mobility.

Neutrons. Neutron bombardment of a semiconductor has its peculiar characteristics. First, the ionization of atoms by neutrons is a secondary process brought about not by the action of neutrons themselves but by the action of the atoms to which neutrons have imparted their energy in collisions and by radioactivity induced by neutrons in the atoms of the basic substance. Second, upon irradiation by neutrons, radioactive transformations take place in a semiconductor, with the result that foreign-element chemical impurities are formed in the original pure substance. Thus, for example, gallium, arsenic and selenium impurities appear in germanium upon neutron bombardment. This variety of products of radioactive transformation of germanium atoms is a result of the complex isotopic structure of the latter element. The atoms of gallium, an element of group III of the periodic table, act as acceptors; those of arsenic, an element of group V, behave as donors (see page 65), while the selenium atoms, as indicated by special investigations, exert no influence on current-carrier concentration in germanium. Since the formation of gallium proceeds three times as rapidly as that of arsenic, the effect of the radioactivity factor during bombardment by neutrons is to cause preferential accumulation of acceptor impurities. Slow neutrons produce a predominant radioactivity effect, although it should be noted that the cross section of this reaction is comparatively small in germanium, and in addition, fast neutrons cause crystal defects which, as was shown before, are of acceptor character in germanium. As a result, the bombardment of germanium by fast and slow neutrons is found to lead in general to the same result as does irradiation by fast electrons: n-type germanium is gradually transformed into the p-type as irradiation continues, its conductivity changing in accordance with the changes in the concentration and mobility of current carriers. The conductivity of p-type germanium increases monotonically throughout the process of neutron bombardment.

Fig. 33. Schematic diagram illustrating the process of change in the conductivity of germanium under neutron bombardment.

Fig. 33 presents a band diagram which illustrates the above process of transformation of electron-type germanium into the hole type: with the progressive

[redacted]

accumulation of acceptor impurities, the compensation of the donor chemical impurities which were present in the n-type germanium before irradiation are compensated first; the total current-carrier concentration decreases in the course of this compensation process; then, with further irradiation, which leads to the formation of an excess of acceptor impurities, hole-type current carriers appear in the valence band, their number increasing with the duration of bombardment.

It might be assumed on the basis of general considerations that the irradiation of silicon would lead to the same qualitative results as irradiation of germanium. It is true that the crystalline structure of silicon is the same as that of germanium, and the influence of chemical impurities on the conductivity of silicon is of the same character as that in the case of germanium. Experiments, however, have shown that the above assumption is not the case. The conductivity of silicon irradiated by high-speed neutrons decreases throughout the process of irradiation from the outset irrespective of the sign of the initial conductivity. The sign of conductivity affects only the rate of the increase in electrical resistance. Silicon with a specific resistance of 10^4 ohm·cm has been obtained in this way; this figure, however, is still 30 times smaller than the specific resistance which perfectly pure silicon should exhibit at room temperature (see Fig. 13). This behavior of silicon may be understood if it is assumed that the donor and acceptor levels created by bombardment with high-speed neutrons are located somewhere in the middle of the forbidden band -- which is 1 1/2 times as wide in silicon as in germanium (ΔE_0 for silicon = 1.1 ev for germanium 0.75 ev) -- at a removal from the boundaries of both the valence and conduction bands. Such an assumption appears quite plausible since it implies essentially that the activation energy of the donor levels created by atoms which have entered the interstices and of the acceptor levels created by lattice vacancies considerably exceed the activation energy of the chemically injected impurities which usually occupy points of the lattice. Thermal excitation of donors and acceptors situated in the middle of the

STAT

forbidden band in silicon is ineffective at room temperature, but the capture of electrons from the conduction band by these acceptors and the recombination of holes in the valence band by these donors in n-type silicon lead alike to both a decrease in current carrier concentration and an increase in specific resistance. The increase in the resistance of silicon is further promoted by the decrease in current-carrier mobility due to secondary scattering of current carriers by the compensated impurities present in such semiconductors in the ionized state at room temperature ($\Delta \underline{E}_{\underline{I}} \ll \underline{kT}$).

Neutron bombardment of the intermetallic semiconductor InSb results in radioactive conversion of In into Sn, the cross section of capture of slow neutrons by the In nuclei being especially large as distinguished from the case of the Ge nuclei. The Sn atoms which take positions at points of the InSb lattice act as donors.

The effect of other forms of corpuscular radiation has thus far been studied only in germanium. Irradiation by deuterons and α-particles was found to lead to results qualitatively similar to those obtained by irradiation with high-speed electrons and neutrons.

As distinct from neutrons, which cause uniform changes in electrical properties over the entire depth of even comparatively thick specimens, charged particles penetrate solid substances only insignificantly. Thus, for example, α-particles with an energy of 5 Mev penetrate germanium only to depth of $2 \cdot 10^{-3}$ cm.

Clarification of all aspects of the physical processes which occur during bombardment of atoms of a semiconductor by various types of corpuscular radiation is of great theoretical and practical importance. As an example of the practical value of such irradiation, it might be pointed out that by irradiation of a piece of electron-type germanium with a directed narrow beam of charged particles, electron-hole transitions may be produced at desired places in the piece. It is even possible to construct a new type of crystal radiation counter with the aid of the

STAT

p-n junctions created in this way (see page 122). It has already been indicated that the electron-hole transitions represent the basis for the technical utilization of semiconductors in the fabrication of rectifiers, amplifiers, and photocells.

STAT

CHAPTER VIII

THE INFLUENCE OF DEFORMATION OF SEMICONDUCTORS ON THEIR CONDUCTIVITY

Deformation of a crystalline solid affects its conductivity in two ways.

1) Upon a change in the interatomic distance, the energy depths of the valence, forbidden, and conduction bands, as well as the location of the impurity levels, also change (see Page 37). These changes result in changes in the activation energy of intrinsic and extrinsic current carriers as well as in the effective masses of the electrons and holes, which depend on the width of the valence and conduction bands; either change has a direct influence on current-carrier concentration in the solid [see Formulas (4) and (7)].

2) As the atoms approach one another, the forces of interatomic interaction increase, with the result that the amplitudes of thermal vibrations of these atoms in the lattice decrease. The decrease in the vibrational amplitudes of the atoms asserts itself in decreased thermal scattering of current carriers (see Page 87), which leads to an increase in mobility in a crystal under omnilateral compression. The above factor of variation of the effective masses must also affect the current-carrier mobility.

The change in mobility is the basic factor in metals. Therefore most metals, when compressed omnilaterally so that the interatomic distances decrease, show an increase in conductivity. The few anomalous cases of decreasing conductivity are accounted for by the action of the second factor -- the increase in the effective mass of current carriers as the atoms approach each other.

STAT

The change in the activation energy, which produces a change in the current carrier concentration, is the factor which plays the principal part in semiconductors. On page37, general theoretical considerations were given according to which the forbidden-band width of a crystal may either increase or decrease as the atoms draw together. Therefore the same deformation may cause both an increase and a decrease in conductivity in different semiconductors.

Experiments with elongation and omnilateral (hydrostatic) compression of semiconductors confirm this general conclusion.

Experimental study of the influence of elastic elongation on the conductivity of germanium and silicon monocrystals at room temperature has shown that the effect of tension depends on the orientation of the crystal and the sign of the conductivity of the semiconductor. Ge-n, Ge-p, and Si-n show an increase, and Si-p shows a decrease in conductivity when the tension and the current are directed along the [100] axis. Ge-n and Si-n show an increase, and Ge-p and Si-p a decrease in conductivity when the tension and the current are directed along the [110] axis. In the latter case, the change in conductivity attains a value of 0.01 per cent per 1 kg/cm^2 of stress. It may be concluded from the results of these tension tests with monocrystals that as the atoms draw together, the current-carrier mobility increases in a normal manner in Ge-p, Si-n, and Si-p, while it decreases in Ge-n. The latter is probably a result of the considerable increase in the effective mass of electrons in the conduction band.

The influence of hydrostatic pressure on the conductivity of germanium is also different for electron-type and hole-type specimens. While the resistance of Ge-n at room temperature increases by a factor of 4.5 under a pressure of 30,000 kg/cm^2, that of Ge-p decreases by only a few per cent under the same conditions. Hydrostatic pressure has the reverse effect on silicon: the resistance of Si-n is reduced by half at a pressure of 30,000 kg/cm^2, and that of Si-p under the same pressure increases by several tenths. In the latter case, a strongly marked

[redacted]

STAT

hysteresis effect is observed: the change in the resistance of Si-p with a decrease in pressure from 30,000 kg/cm^2 to zero is represented by a curve completely different from that for increasing pressure and having a sharply defined resistance maximum.

Wide conductivity variation is observed in tellurium. At a hydrostatic pressure of 30,000 kg/cm^2, the resistance of tellurium decreases, at room temperature, by a factor of 600. A calculation indicates that a reduction in forbidden-band width to 1/25 of normal (from 0.36 to 0.015 ev) corresponds to such a decrease in conductivity. At a hydrostatic pressure of 45,000 kg/cm^2, tellurium undergoes a phase transformation into a metallic state. It is of interest to note that hydrostatic pressure produces anomalous deformation in tellurium: compression occurs only in the direction perpendicular to the main axis of the crystal; a tellurium monocrystal is elongated along this axis.

The resistance of selenium decreases by 30 per cent at a hydrostatic pressure of 1000 kg/cm^2, and increases by a factor of 10^4 with a further increase in pressure to 100,000 kg/cm^2.

At a hydrostatic pressure of 13,000 kg/cm^2 and a temperature of 200° C, yellow phosphorus, which is an insulator with a forbidden-band width of 2.1 ev, undergoes an irreversible phase-transformation into the semiconductor black phosphorus, whose forbidden-band width is only 0.33 ev.

At room temperature, the resistances of the oxide semiconductors ZnO, NiO, TiO_2, V_2O_5, and U_3O_8, which have been studied under pressures up to 50,000 kg/cm^2, decrease severalfold (by a factor of up to ten) with declining pressure; the resistance of Mn_3O_4 increases with pressure at approximately the same rate.

Experiments made in studies of the effect of deformation of semiconductors on their electrical properties enable us to draw important theoretical conclusions concerning the structure of the electronic energy spectrum.

The dependence of the resistance of semiconductors on elastic deformation finds

STAT

practical application in the tensometers used for the determination of stresses which arise in structural elements and machine parts.

STAT

CHAPTER IX

ON THE CONDUCTIVITY OF LIQUID, AMORPHOUS, AND POLYCRYSTALLINE BODIES

The band aspect of the modern quantum theory of solids enables us to explain satisfactorily all of the various experimentally-observed electrical properties of semiconductors, metals and insulators from a single point of view. A basic position of this theory is the concept of the solid as an aggregate of atoms and ions arranged in an orderly manner in a lattice. According to this theory, the electronic conductivity of solids is determined by the band structure of the energy spectrum of the electrons. This band structure of the energy spectrum proceeds from very general considerations related to the periodic structure of the crystal. Therefore in discussing the various electrical properties of semiconductors we frequently emphasized the crystalline structure of the substances which were of interest for us, as the only case to which the conclusions of the band theory are actually applicable. Experiments confirm that the crystalline structure of a solid is, in many cases, a highly important factor in determining its electrical properties. Indeed, the width of the forbidden band, which exerts an essential influence on the electrical properties of a semiconductor, depends in a very definite manner on the spatial arrangement of atoms in the lattice. Carbon, which occurs in two modifications with sharply differing forbidden band widths depending on the nature of the crystal lattice (diamond and graphite), may serve as an example of this dependence. A similar situation, as we have seen, is encountered in tin and phosphorus. Since the great majority of solids have crystalline structure, we may study their electronic conductivities with the aid of the band theory, in which this unifying property is a point of departure.

On the other hand, it is known that noncrystalline substances also show electronic conductivity of both metallic and semiconductor character. It is known, for example, that the conductivity of the majority of typical metals declines only by a factor of 1.5 to 2 upon fusion, so that the metallic character is preserved. The conductivities of bismuth, gallium, and antimony are even doubled upon fusion. It is also found that the electronic conductivity of many crystalline semiconductors is not essentially changed by melting, when the crystalline structure disappears. Both increases and decreases in conductivity after fusion are observed in this case, for which we may discern in the experimental data a distinct relationship between the sign of the change in conductivity and that of the change in density during melting. Thus, for example, germanium shows an increase in density and an increase in conductivity at the melting point. Selenium shows a considerable decrease in density (13 percent) accompanied by a sharp drop in conductivity (by a factor of 3,000). It is also known that the electronic conductivity of liquid solutions of alkali metals in ammonia is of a semiconductor nature. Finally, it is known that electronic conductivity of a semiconductor nature is observed in amorphous selenium, amorphous antimony, and certain glasses.

Thus there exists a whole series of experimental facts which indicate that the crystalline structure of a body radically influences its conductivity, and another series of experimental facts which indicate that even the disappearance of the crystalline structure affects conductivity only slightly. From these facts, which are mutually contradictory in terms of modern band theory, we may draw the conclusion that the crystalline structure of a body does not exert decisive influence on electronic conductivity -- the latter being determined not by long-range order, or the orderly disposition of the atoms over great distances (as compared with the interatomic distances), but by short-range order, which is established by the arrangement of the closest neighbors of a given atom, by their number, and by their distances

STAT

from one another.

It is known that when the long-range order is disturbed upon the fusion of crystalline bodies, the short-range order is preserved; only with further heating of the liquid toward the critical temperature does the latter gradually disappear. This development explains both the facts observed upon fusion of electronic semiconductors and the correlation between the changes in conductivity and density which was noted above.

The modern band theory of the conductivity of solids, which overestimates the part played by long-range order in crystalline bodies, is, in essence, an approximate theory which disregards many of the important physical factors active in dense systems. It is surprising, therefore, not that this theory fails to explain the electrical properties of liquid and amorphous bodies, but that the varied electrical properties of crystalline bodies find their explanation within the framework of this comparatively simple theory.

The viewpoint that the electrical properties of solids are basically determined by the nature of the chemical inter-particle bonds has recently won increasing recognition. Bonds of the homopolar type in particular give rise to semiconductor properties. It is maintained that the fundamental difference between metals and nonmetals is related not to the width of the forbidden band of the crystal but to the degree of saturation of chemical bonds: those of the metals are not saturated, while those of nonmetals are saturated to a considerable degree. The "chemical" theory is empirically confirmed by the established relationship between the nature of chemical bonds in semiconductor materials and such factors as activation energy and current-carrier mobility. Rules formulated on this basis have made it possible to predict the electrical properties of certain materials when these are prepared in a specific manner.

Yet another general observation should be made in connection with the fact that the majority of real solids which we encounter in practice are not monocrystals but polycrystalline aggregates consisting of monocrystalline grains with differently oriented crystallographic axes. Since the dimensions of these grains are usually large as compared to the interatomic distances in the lattice, the conclusions of the band theory may be applied to them.

In what way can the polycrystalline structure of a solid affect its conductivity? A transition atomic layer of the order of several interatomic distances must exist as a result of the difference between the regular arrangement of atoms inside the grain and the distorted arrangement at the grain boundaries. If the current-carrier free path were comparable to the dimensions of crystalline grains, secondary scattering at their boundaries would noticeably reduce the current-carrier mobility and thereby the conductivity of the polycrystal as well. Under ordinary conditions, however, the current-carrier free path in metals and semiconductors ($< 10^{-5}$ cm) is smaller by far than the dimensions of the crystalline grains, so that the presence of transition layers between the grains of a polycrystalline aggregate will not, in itself, influence conduction in it; the conductivity of a polycrystal would be determined by simply averaging over all crystallographic directions.

The possibility of accumulation of the bulk of impurities at the grain boundaries in a polycrystalline material is found to exert an essential influence on conductivity. Such contaminated layers, when sufficiently thick, increase the electrical resistance of metallic specimens. The presence of such transition layers with elevated impurity concentrations in semiconductors with small volume conductivity may in some cases cause a significant part of the current to flow over the surface of grains having increased extrinsic conductivity, rather than through monocrystalline grains of the basic substance. In other cases in which the transition layers have elevated resistance, the conductivity of the polycrystal is lower than that of the substance itself. In all cases, however, the contaminated layers reduce the effective current-carrier mobility noticeably.

The presence of adsorbed impurities in the grain surfaces may exert a strong influence on the photoconductivity of semiconductors, since the current carriers

STAT

liberated by the light may be captured by the intercrystalline layers, with the result that the photocurrent-carrier mobility is sharply limited. On the other hand, when the intercrystalline layers have higher resistances (oxides, for example), the photoconductivity of the substance may prove to be anomalously high (see page 106).

The polycrystalline structure of a solid may have an effect on its conductivity in still another way. The fact is that, as noted previously, the energy state of all atoms at grain boundaries is different from that of atoms of the basic substance at a distance from the boundaries even though impurities are not adsorbed at the surface of the crystalline grains. Due to this circumstance, separate "surface levels" (Tamm levels) are formed in the electron-energy spectrum which, like the impurity levels, may also occur in the forbidden band of the crystal involved. If electrons are present in these levels in sufficient numbers, surface conductivity due to the surface states of the electrons may occur even in substances with virtually no volume conductivity (due to the great width of the forbidden band), just as in the case of impurities. The experimental study of these problems is still in its initial stage. The presence of surface levels in the electronic energy spectrum of a solid may have noticeable effect not only on its dark conductivity, but on its photoelectric, contact, and thermoemission properties as well.

In any case, the conclusion can be drawn from the experimental data that the mobility and particularly the lifetime (see page 38) of the current carriers are, as a rule, smaller in polycrystals than in monocrystals of the same substance. This may be explained by the fact that monocrystals, for the most part, are purer than polycrystals due to the displacement of impurities to the boundaries of the monocrystalline grains. It is therefore necessary to use monocrystalline semiconductors in those applications in which the mobility and lifetime of the current carriers are important factors.

Nonuniform distribution of impurities may exert a strong influence on the electrical properties of a semiconductor. Of particular interest is the case of

STAT

the so called p-n transition which occurs between two adjacent zones of a specimen when one of these has an excess of donors and the other an excess of acceptors.

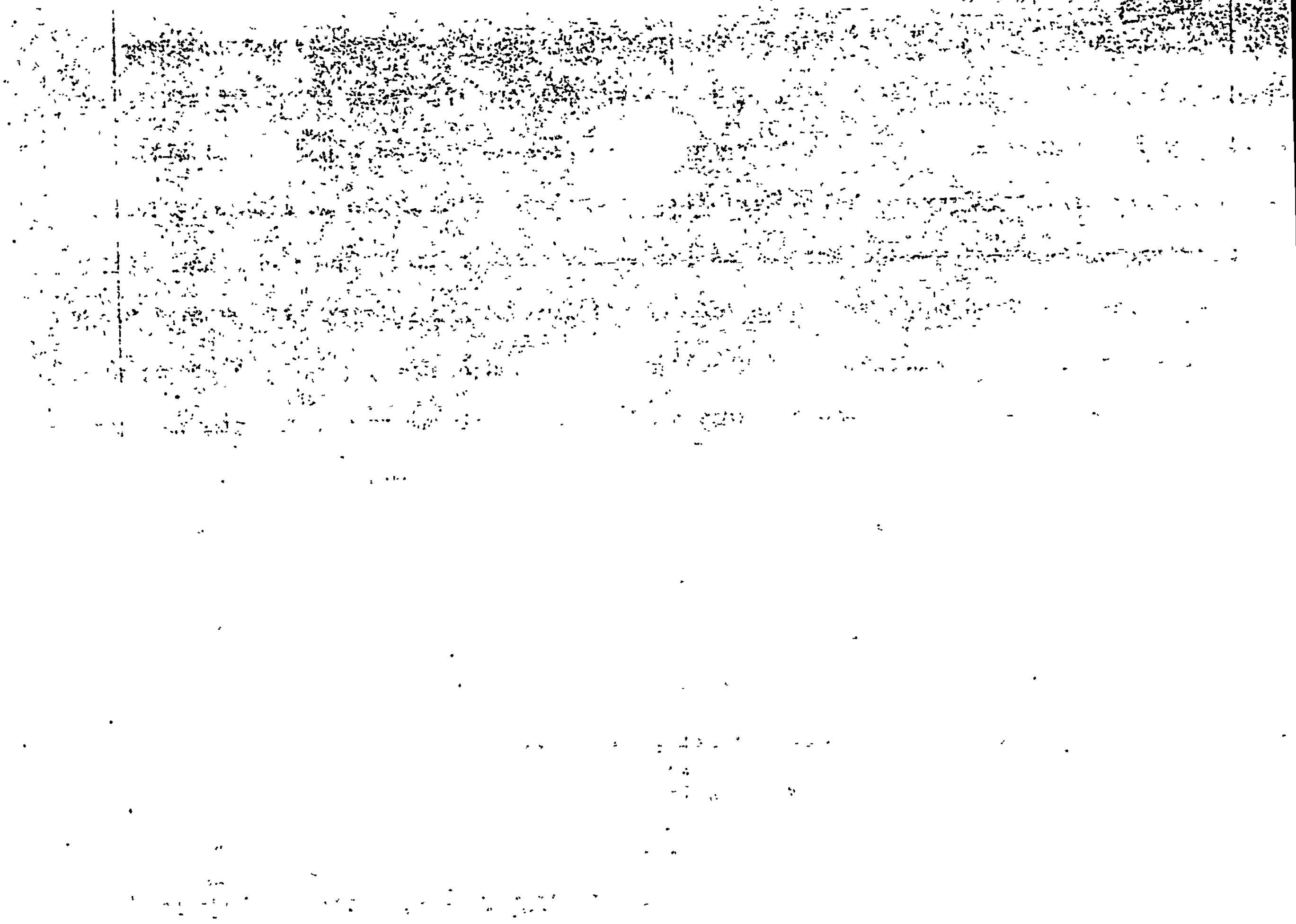

Fig. 34. Positions of impurity levels of various elements in the forbidden bands of silicon and germanium.
D -- donor level; A-- acceptor level; D. F. -- Frenkel defects.
(The width of the forbidden bands corresponds to $T = 0$.)

The important physical processes which occur in such a boundary layer in a semiconductor will be discussed in the next chapter.

In the table below are given numerical values of the basic quantities determining the intrinsic conductivity of the most important nonmetallic solids--semiconductors and insulators--at room temperature. The numerical values for the mobilities of the electrons and holes depend on the degree of perfection of the crystal lattice of the solid: the more pefect the crystal the higher the mobility. Therefore, values obtained by different researchers for these quantities frequently disagree. The table gives the most reliable data available in 1956.

This table should be supplemented by the value of the activation energy of the impurity centers as a characteristic of the extrinsic conductivity of the listed substances. However, data on these values are as yet fragmentary for most substances. Some data on the distribution of impurity levels in the forbidden bands of semiconductors have been obtained only for germanium and, to a lesser extent, for silicon (see Fig. 34). The ability of a single chemical element to create several local energy levels is explained by the ability of atoms of this element to assume different degrees of ionization. Thus, for example, gold atoms, which create four types of impurity levels in the forbidden band of germanium, may occur in the ionization states Au^{+}, Au^{-}, Au^{--}, and Au^{---}. A donor level corresponds to the first state, and acceptor levels to the others. It is supposed that the donor level is formed by implantation of an atom of gold in the germanium lattice. However, when this atom enters the germanium lattice as a substitution atom, it may possess three degrees of negative ionization in conformity with the tendency of a substitution atom to form, in the germanium lattice, the tetrahedral bond system which is characteristic of this lattice (see page 65), since a four-electron group which establishes a tetrahedral bond between an extrinsic atom (or, more exactly, an ion) and four adjacent germanium atoms is formed when three more electrons are added in succession to an atom of gold with one electron in its outer shell. The assumption that substitution atoms tend to ionize according to the bond system of the lattice into which they are built also explains the fact that univalent copper produces three acceptor levels in germanium. It can be understood from this standpoint why elements of group III form single levels in germanium while two acceptor levels are formed by elements of group II (according to the most recent data, which are not considered in Fig. 34).

STAT

TABLE

Numerical values for basic physical parameters determining intrinsic conductivity of nonmetallic solids at room temperature [see Formulas (2) and (4)].

T_m - melting temperature; T_c - upper limit of the temperature range in which a given semiconductor modification is stable; ΔE_0 - width of forbidden band; u_n and u_p - mobilities of electrons and holes.

Substance	T_m, °C	ΔE_0, ev	μ_n in $\frac{cm^2}{v.sec}$	μ_p in $\frac{cm^2}{v.sec}$	Substance	T_m, °C	ΔE_0, ev	μ_n in $\frac{cm^2}{v.sec}$	μ_p in $\frac{cm^2}{v.sec}$
				Elementary semiconductors					
B	2300	∼1.1	∼10	∼10	As - Sublimated Layer	T_c=300 °C	∼1.2	∼50	
C - graphite	>4000	<0.1			Sb - Sublimated Layer	T_c=0 °C	∼0.12		
- diamond		5.3	1800	1200	S	113	∼2.4		
Si	1420	1.1	1200	500	Se crystalline	220	∼1.2		∼5
Ge	936	0.65	3900	1900	Se - amorphous		∼2.3		
Sn - gray	T_c=13°C	0.08	∼2000	∼1000	Te	452	0.35	1700	1200
P - yellow	44	∼2.1	—	—	I	113.5	∼1.3	∼25	
- red	—	∼1.5	—	—					
- black	—	∼0.33	200	350					
				Complex nonmetallic crystals					
Ag_2Te	955	0.17	4000		Fe_2S	689	1.2	100	
$InSb$	523	0.18	80000	4000	Ga_2Te_3		1.2		
$HgTe$	670	0.2	17000	∼100	InP	1050	1.3	3500	700
Mg_2Sn	778	0.2	300	250	$GaAs$	1280	1.35	4000	400
Bi_2Te_3	585	0.25	600	150	$CdTe$	1045	1.5	450	100
$HgSe$	690	0.3	>10000		$AlSb$	1060	1.52	50	150
$InAs$	942	0.35	30000	200	SiC	2700	1.5	60	10
$PbSe$	1065	0.55	1500	1500	Cu_2O	1232	∼1.5		100
$PbTe$	917	0.6	2100	800	$CdSe$	1350	1.8	600	
PbS	1114	0.6	600	800	Ca_2Si	920	1.9		
Ca_2Pb	1110	∼0.45			$ZnTe$	1239	2.2		100
Cd_3As_2	721	0.5			$AlAs$	>1600	∼2.2		
$CoSb_3$	858	0.5	300		Fe_2O_3	∼1500	∼2.2		
Cs_3Sb		∼0.55	500		GaP		2.25		
Cs_3Bi		0.55		350	CdS	1750	∼2.4	200	
$CdSb$	456	0.55		350	Al_2O_3	2050	2.5		
$ZnSb$	546	0.6	60	100	AgI	552	2.8	30	
Mg_2Ge	1115	∼0.65	500	1000	$ZnSe$	1240	2.8	∼100	
$GaSb$	725	0.7	5000	70	AlP	∼1500	∼3.0		
Mg_2Si	1102	∼0.75	400	80	ZnO	1975	3.2	200	
Mg_3Sb_2	1228	0.8	20		ZnS	1850	3.7		
K_3Sb	812	0.8			$CuBr$	498	5	∼30	
KSb	605	0.9			KBr	741	6.6	12	
Ca_2Sn	1122	0.9			KCl	770	7.6		
Ag_2S	842	∼1.0	50		$NaCl$	800	7.8	25	
Zn_3As_2	1015	1.0							

CHAPTER X

THE ELECTRICAL PROPERTIES OF BOUNDARY LAYERS OF SEMICONDUCTORS

The above discussion was concerned only with phenomena occurring inside the volume of a semiconductor. Some important technical applications of semiconductors are based on phenomena which occur at the boundary of a semiconductor at which it comes into contact with a metal or another semiconductor. The study of these contact effects is also of great importance in the theory of semiconductors.

Thorough description of the contact effects and the methods of their practical utilization is a subject of the special issues of this series devoted to semiconductor rectifiers, amplifiers, and photocells. Here we shall attempt to identify only the specific properties of semiconductors responsible for the peculiar phenomena which occur at the junction between a semiconductor and a metal or another semiconductor but not at contacts between two metallic conductors. In doing this we shall proceed from the general physical conception of semiconductors as developed above.

First, let us discuss briefly the physical sense of the two basic quantities which characterize the contact phenomena - the chemical potential μ and the work function R.

In Part I it was shown that many important electrical properties of semiconductors and metals may be understood on the basis of the simple model of the free electron gas. According to this concept, the conduction electrons in a conductor are in a state of continuous chaotic motion; migrating freely in the conductor, they may occur at any point of its volume but may not under normal conditions

STAT

[redacted]

escape outside it. We conclude from this that at the conductor boundaries there exists an electric field which prevents expulsion of free electrons -- or, in other words, a potential barrier exists at the surface of a conductor.

Physically, the formation of this barrier is due to the peculiar state of the atoms at the surface of the body. Unlike the atoms in the interior, which are acted upon by forces from all sides, the atoms at the surface experience a unidirectional inward attraction. Therefore, the interatomic distances and the energy states of electrons at the surface of the body differ from those in its interior. The resulting displacement of charges (electrons and nuclei) creates at the surface of the body a double electric layer (as in a capacitor) whose field presents a potential barrier which prevents free electrons from being expelled outside the body. Moreover, an electron moving away from the surface causes polarization of the body, with the result that an electrical image force begins to act upon it, attracting it back to the surface.

Experience also indicates that electrons may escape from a conductor under certain conditions. To effect this, it is necessary either to increase the energy of electrons or to reduce the height of the potential barrier. The latter is achieved by means of a strong electric field directed in such a way as to facilitate the escape of electrons. The energy of the free electrons of a conductor may be increased, as is known, in various ways: heating the body, illuminating it, or bombarding the body with corpuscular radiation. These experimental facts indicate that the free electrons of a conductor are strongly attracted to ions present in the conductor; the latter are formed by the collectivization of free electrons which endows them with freedom of motion throughout the volume of the specimen. If it is assumed that the potential energy of an electron which is infinitely distant from the conductor is zero, a high negative potential energy should be ascribed to the electrons inside the conductor.

The simplest energy model of a conductor may be presented in the form of the

STAT

rectangular potential well shown in Fig. 35. The energy of the electrons is plotted against the vertical axis: 0 is the potential-energy value of an electron outside the conductor, which is taken as the origin of the energy scale; A is the potential energy of an electron inside the conductor; ϵ_i and ϵ_k are total electron-energy values which compose the energy spectrum. The kinetic energies of the various electrons are defined as the differences $(\epsilon_i - A)$.

It is known from the opening chapters of the present edition (Part I) that the electron-energy levels in a solid are quantized and that no more than two electrons with oppositely directed spins may be present in each level. This is represented in Fig. 35 by a system of horizontal lines with a double row of oppositely directed arrows (ϵ_i, ϵ_k). On the basis of the already familiar concepts, we may consider that the level A corresponds to the bottom of the conduction band of the solid; the energy difference $(0 - \epsilon_i)$ determines the energy which must be provided to extract an electron from the level ϵ_i.

In metals the number of free electrons does not depend on temperature, and at $T = 0$ they occupy all levels ranging from A to a certain ϵ_{max}. The level ϵ_{max} to which the potential well of the metal is filled is called the Fermi level ϵ_F, or the chemical potential μ. Theory indicates that the numerical value of this quantity is determined in a metal by the concentration n of electrons and their effective mass m:

$$\epsilon_F = \frac{h^2}{2m}\left(\frac{3n}{8\pi}\right)^{2/3}.$$

The position of the Fermi level in metals is practically independent of temperature.

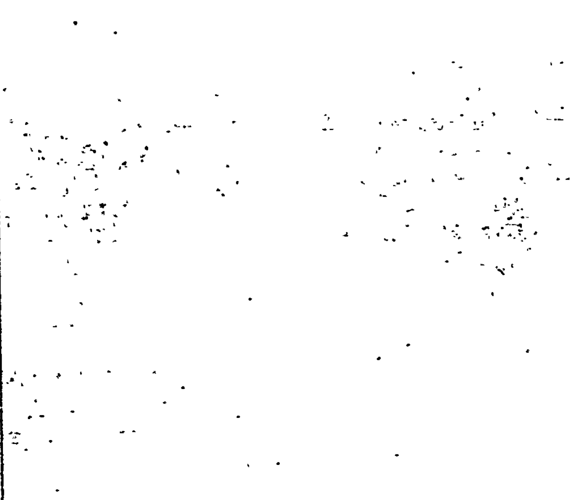

Fig. 35. Energy spectrum of electrons in a metal.

From Fig. 35 it is seen that the minimum energy which must be expended to remove one electron from a metal is determined by the distance $(0 - \epsilon_F)$. A larger energy is required for removal from some intermediate level ϵ_i located below ϵ_F. At thermodynamic equilibrium, however, when the electrons present in the body tend to occupy all lower levels of the energy spectrum, the electron from the highest level will descend, either directly or stepwise, into a vacated intermediate level. An energy $(\epsilon_F - \epsilon_i)$ will be released during such a transition. When thermionic emission takes place, in which case the energy consumed in extracting the electron is thermal energy, and if, in addition, the energy liberated in the transition of another electron downward to a vacated intermediate level is also thermal energy, then the same energy $0 - \epsilon_F = \underline{R}$ will be always consumed by removal of an electron from any energy level. This energy, which is determined by the position of the chemical potential level μ in the energy spectrum of the solid, is called the thermodynamic work function.

Thus it is found that an energy equal to the work function $\underline{R}$ is required to extricate an electron from a solid; the same energy is released when an electron is injected into this solid from without. Since the position of the Fermi level in metals is practically independent of both concentration and temperature, the work function $\underline{R}$ is a physical constant for a given metal and determines both the intensity of its thermionic emission and the red threshold of the photoemissive effect.

The facts are otherwise in the case of semiconductors. As we know, the conduction band of this class of electronic conductors becomes vacant at $\underline{T} = 0$: the free electrons which were present in it with $T > 0$ descend as $T \rightarrow 0$ either to the valence band in the case of an intrinsic semiconductor or to the local impurity levels in the case of an extrinsic semiconductor. The concept of the chemical potential μ (which is also called the Fermi level) is applied in theoretical descriptions of electronic properties of semiconductors just as in the case of metals. Actually, this factor in all cases determines the probability of occupation of the electronic energy levels of a crystal and, consequently, the current-carrier concentration as well. All energy levels below μ are occupied with a probability near unity; the probability of occupation of energy levels above μ declines with increasing distance of the energy level from μ. The positions of the chemical potential on the energy diagrams for intrinsic, donor and acceptor semiconductors at $T = 0$ are indicated in Fig. 36.

The position of μ on these diagrams at $T > 0$ varies according to the following approximate formulas:

for an intrinsic semiconductor:

$$\mu = -\frac{\Delta E_o}{2} + kT \ln \left(\frac{m_p}{m_n}\right)^{3/4}$$

for an impurity semiconductor:

$$\mu = -\frac{\Delta E_n}{2} + kT \ln \sqrt{\frac{N_n}{2}} \frac{h^{3/2}}{(2\pi mkT)^{3/4}},$$

The notation used in these formulas is the same as in Formulas (4) and (7), and the energy scale begins at the bottom of the conduction band.

The thermodynamic work function of electrons is determined in any semiconductor, as in the case of metals, by "the depth of occurrence" of the chemical potential. However, since, as seen from the formulas presented above, the position of the chemical-potential level on the energy diagrams of semiconductors (as distinguished from metals) depends appreciably on the temperature of the body and the impurity concentration, the thermodynamic work function of a semiconductor depends noticeably

[redacted]

upon these values.

The thermodynamic work function of semiconductors, as distinguished from the case of the metals, determines only the intensity of its thermionic emission and not the energy of the photoemissive effect. Therefore in discussing certain problems of semiconductor electronics it is useful to consider, along with the thermodynamic work function, the so-called external work function R_e, which is defined as the distance between the level of an electron which has been removed from the body and the bottom of the conduction band.

It is known from the opening chapters of the present volume that the position of the bottom of the conduction band on the energy diagram of a semiconductor depends neither on temperature nor on impurity concentration and is determined solely by the crystalline structure of the material of the semiconductor. Therefore the external work function is distinguished from the thermodynamic work function by its nature as a characteristic constant of the material of a semiconductor. When considering contact effects we have recourse to the thermodynamic work function, which is often called simply the work function.

Fig. 36. Position of chemical-potential level on the energy diagrams of a - intrinsic, b - donor, c - acceptor semiconductors.

Now, after briefly clarifying the physical significance of the concepts of chemical potential and work function, we can turn to our immediate concern -- the processes which occur in the boundary layer at the contact between any two conductors.

It will be appropriate here to refer to certain simple physical examples. Let us recall what occurs when communication is established between two vessels filled with gas under different pressures or filled with liquid to different levels. In the first case, as we know, the gas will flow from the vessel with the higher pressure into the vessel with lower pressure until such time as the pressures in both vessels are equal; with equal temperatures in both vessels, the condition of equality of pressures reduces to that of equality of concentration. In the second case, as is known, the transfer of liquid ceases when the levels are equally high.

Something similar (to a certain degree) occurs when conductors, which are filled with an electron gas, come into contact. When two conductors are drawn close to one another they begin to interchange electrons. What will be the direction of the resultant electron flow and how is the equilibrium condition formulated under which the flow ceases? Such are the basic questions which we must ask here.

The electron theory answers these questions in the following way. When electron conductors are brought into contact, they begin to interchange electrons. At first, the electronic counterflows are not equal and their difference originates a resultant flow. The direction of the resultant flow is determined by the relation between the work functions of the contacting conductors: the flow is always directed from the body with lower work function to the body with higher work function. The resultant electron flow imparts a negative charge -- relative to the body with lower work function -- to the body with the higher work function. A double electric charge forms in the boundary layer and a contact potential difference is set up between the two bodies. The electric field which is created in this double layer at the conductor boundary by the resultant-charge flow retards the electrons which participate in this flow, gradually reducing the latter to zero with increasing field strength. An equilibrium state is established in which the resultant electron flow is zero and a contact potential difference determined by the difference in the work functions of the two bodies is created at the contact layer.

As in the case of the gas equilibrium, the fact that the resultant electron flow is equal to zero does not result in cessation of the particle interchange between the contacting bodies: the established equilibrium is a dynamic one.

The creation of the contact potential difference gives rise to mutual displacement of the electron-energy spectra of the contacting bodies: both the band boundaries and the chemical-potential levels are displaced. All levels are displaced upward in a negatively charged body and downward in a positively charged body. This is easy to understand when we remember that the potential energy of an electron increases in the direction from the plus sign to the minus sign as it moves in the contact field of the electric double layer. Thus the presence of a contact potential difference leads to a relative displacement of the Fermi levels in the contacting bodies. Since the electrical potential of the body has an influence on the position of the Fermi level, the latter is in this case referred to as the electrochemical potential. The theory states that equality of Fermi levels is a condition for the equilibrium at which the resultant electron current reduces to zero, and plays the same part in the contact between electronic semiconductors as the equality of pressures in the vessels filled with an ordinary gas or equality of levels in the communicating vessels filled with a liquid. It should be emphasized that these general theoretical conclusions relate both to metals and to semiconductors (which we designate collectively in this case by the term electronic conductors).

How, then, is the difference between metals and semiconductors manifested, and what important effects does it produce in the contact phenomena?

This can be best explained by discussing three cases in succession: 1) metal-to-metal contact, 2) metal-to-semiconductor contact, and 3) semiconductor-to-semiconductor contact.

1. Metal-to-Metal Contact

Fig. 37 gives a schematic representation of displacement of energy levels and creation of a contact potential difference when two metals come into contact. The

Fig. 37. Scheme of formation of contact potential difference between two metals. R - work function; d - thickness of contact layer in which contact electric field is concentrated.

contact potential difference V_c in the boundary layer d is

$$V_c = (1/e)(R_1 - R_2).$$

The thickness d of the electric double layer in which the contact field is concentrated is readily estimated.

The contact potential difference of both metals and semiconductors constitutes a quantity of the order of 1 volt. The double charge layer may be regarded essentially as a plane capacitor. Using the customary formula for the capacitance per unit area of a plane capacitor,

$$C = 1/4\pi d$$

and the formula relating the surface charge density q to the potential difference V_c and the capacitance C,

$$q = CV_c,$$

we can obtain an estimation formula for the double-layer thickness:

$$d = V_c/4\pi q.$$

The surface free-electron density in metals may be determined from the number of atoms present per 1 cm^2 of surface. Since the latter constitutes approximately 10^{15} cm^{-2}, the number of free electrons will be the same -- i.e., $10^{15} cm^{-2}$ -- when one free electron is present in each atom. Analysis of the latter formula shows that the thickness of the contact layer in metals cannot exceed the interatomic distances -- i.e., a value of the order of $3 \cdot 10^{-8}$ cm. Indeed, at $d = 3 \cdot 10^{-8}$ cm,

$$q = \frac{V_c}{4\pi d} = \frac{1/300}{12.5 \cdot 3.10^{-8}} \approx 10^4 \text{ statcoulombs/cm}^2 = 2 \cdot 10^{13} \text{ electrons/cm}^2.$$

Thus the migration of only a few per cent of the number of free electrons present in the surface layer from one metal into the other is sufficient to create a contact potential difference of 1 v. Such a change in the concentration of the monatomic suface layer of a metal exerts practically no influence on the conductivity of the contact layer, so that a current flows through the contacting metals as freely as through the volumes of the metals themselves. Therefore the metal-to-metal contact shows no rectifying effect (as distinguished from semiconductors).

2. Metal-to-Semiconductor Contact

Conditions quite different from those discussed above may arise at metal-to-semiconductor contacts. The fact is that if the current-carrier concentration in the volume of the semiconductor is no greater than 10^{15} cm^{-3}, the concentration in the monatomic surface layer amounts to only $10^{15} \cdot 5 \cdot 10^{-8} = 5 \cdot 10^{7}$ cm^{-2}. Such a charge concentration can create a potential difference of only 10^{-6} v when the thickness of the double layer is of atomic dimensions. Therefore the leveling of electrochemical potentials at a metal-semiconductor contact occurs at the expense of formation of a space charge which extends to a considerable depth into the semiconductor. The field of the contact potential difference penetrates into the semiconductor to a depth which may be determined by the formula

$$d = \sqrt{\frac{(R_m - R_{sc})\,\epsilon}{2\pi e^2 n_o}}, \qquad (16)$$

where ϵ is the dielectric constant of the semiconductor; n_o is the volume concentration of free charges, which is equal to the impurity concentration. (This formula was derived on the assumption that all impurity centers are ionized and that the contact field removes all free charges from a layer of depth d.) Assuming $\epsilon = 10$, $n_o = 10^{15}$ cm^{-3}, and $R_m - R_{sc} = 1$ ev ($1.6 \cdot 10^{-12}$ erg) we obtain

$$d = \sqrt{\frac{1.6 \cdot 10^{-12} \cdot 10}{6.3 \cdot 25 \cdot 10^{-20} \cdot 10^{15}}} = 10^{-4} \text{ cm.}$$

At $n_o = 10^{13}$ cm^{-3}, d amounts to 10^{-3} cm; at $n_o = 10^{17}$ cm^{-3}, $d = 10^{-5}$ cm.

The fewer the ionized impurities (and, consequently, current carriers) in the semiconductor, the greater will be the depth of penetration of the contact field. The contact fields of insulators with very small current-carrier concentrations may penetrate the specimen to a depth of several centimeters.

The dependence of the thickness of the contact layer into which the field penetrates on the contact potential difference V_c and the impurity concentration determined by Formula (16) may be obtained from the following simple and conceptual considerations. It is known that the relationship between the charge q, the capacitance C and the potential difference V_c for a plane capacitor is expressed as $q = CV_c$.

The charge which contributes to the formation of a contact field is equal to the product of the space-charge density and the layer thickness: $q = en_o d$. The capacitance per unit area of the capacitor is $C = \epsilon/4\pi d$. Therefore $V_c = \frac{q}{C} = \frac{en_o d}{\epsilon/4\pi d} = \frac{4\pi e n_o d^2}{\epsilon}$, from which $d = \sqrt{\frac{V_c\,\epsilon}{4\pi e n_o}}$, and since $V_c = \frac{1}{e}(R_m - R_{sc})$, $d = \sqrt{\frac{(R_m - R_{sc})\,\epsilon}{4\pi e^2 n_o}}$.

This expression differs from Formula (16) only by a factor of $1/\sqrt{2}$.

What is the influence of a contact field penetrating into a semiconductor on the electronic energy spectrum? The structure of this spectrum is essentially determined, after all, by the electric field of the atoms of the crystal. Therefore, if the strength of the electric field were a quantity comparable with the strength of the internal electric field of the crystal, it would be necessary to determine first the changes in the structure of the electron-energy spectrum due to the contact potential difference. But the strength of this field never exceeds 10^6 v/cm, while the intensity of atomic fields is at least 100 times greater. Hence it may be considered that the contact field does not change the structure of the energy spectrum itself; it does not change, for example, either the forbidden-band width of the crystal or the impurity activation energy, and it causes only parallel curvature of energy levels. A pattern obtained for the energy spectrum of a donor semiconductor is presented in Fig. 38. This pattern applies to the case where the work function R_{sc} for the semiconductor is smaller than the work function R_m for the metal, and the contact potential difference arises from the transfer of some of the electrons of the conduction band of the semiconductor onto the metal, on which they form a very thin ($\sim 10^{-8}$ cm) negatively-charged layer. The thin negative-charge layer, as we saw above, does not appreciably influence the electron concentration even in a monatomic layer of the metal. A numerically equal positive volume charge of fixed ionized donors is concentrated in a contact layer of thickness d in the semiconductor. In the absence of a contact field, the electron energy at any energy level is represented on the diagram by a horizontal line. By this we denote the fact that the energy of an electron situated at any energy level (for example, at the bottom of the conduction band) is equal at all points of the crystal. If an external field is applied to the crystal, the potential energy of an electron at the bottom of the conduction band (as at any other level) will depend on the coordinate and direction of the field. When charges are arranged as in Fig. 38, the potential energy of an electron in the semiconductor increases from right to left

STAT

in the contact field. Therefore all energy levels in the semiconductor (both in the conduction band and in the valence band) are curved and, as calculation indicates, form parabolas. In the deeper regions of the semiconductor, outside the contact field, the energy levels are horizontal as before but, due to the change in electrical potential, they are displaced in such a way that they retain their original position with respect to the displaced level of the electrochemical potential.

A sharp change in the conductivity of the contact layer of a semiconductor is one of the principal consequences of the creation of the contact potential difference. If this layer were of atomic dimensions the creation of the contact potential difference would not give rise to a sharp change in conductivity, since electrons, due to their wave nature, easily pass through potential barriers of such dimensions. If, however, the contact layer extends over distances hundreds of times greater, the motion of the electron is determined by the conductivity of this layer. In the case to which Fig. 38 applies, the contact layer of the semiconductor represents a layer almost devoid of conduction electrons for which reason its conductivity proves to be greatly reduced as compared to that of the volume of the semiconductor. This accounts for the term "barrier layer" customarily applied to such a layer of a semiconductor. The positive charge of such a barrier layer is

Fig. 38. Creation of contact potential difference between metal and donor semiconductor: ++ fixed ionized donors in contact layer of semiconductor, -- electrons accumulated at boundary of metal, <u>d</u> - thickness of the barrier layer.

created by fixed ionized donors. Any external electric field which we may apply to the metal-semiconductor contact system will sharply influence the conductivity of the barrier layer. With one orientation of the external field (so that the minus applies to the semiconductor and the plus to the metal), mobile charges (in this case electrons) will be attracted from the volume of the semiconductor into the barrier layer, reducing its effective thickness and increasing its conductivity; with another direction of the external field, electrons will be displaced from the contact layer into the depth of the semiconductor, the thickness of the barrier layer will increase, and the conductivity of the contact will drop sharply. This physical presentation explains the rectifying action of semiconductor junctions.

It will be profitable for the sake of general familiarization with the contact phenomena to discuss yet another case in which the work function of the donor semiconductor is greater than that of the metal. The pattern of the energy spectrum for this case is presented in Fig. 39. In this case electrons migrate from the metal into the contact layer of the semiconductor, charging the latter negatively. Therefore the curvature of the energy levels is turned in the other direction. Here the contact layer of the semiconductor is found to be much thinner than

Fig. 39. Schematic diagram of formation of antibarrier layer at contact between a metal and a donor-semiconductor whose work function is greater than that of the metal.

for $R_m > R_{sc}$, and -- which is most important -- the conductivity of the semiconductor contact layer, in which the electron concentration is greater in this case, is higher than that in its volume. The layer thus formed constitutes an antibarrier layer and cannot serve to rectify alternating currents. Similarly, it can be shown that formation of a barrier layer in an impurity acceptor semiconductor is possible only with $R_m < R_{sc}$. When this inequality is reversed an antibarrier layer is formed in the contact layer of a hole-type semiconductor.

3. Semiconductor-to-Semiconductor Contact

Junctions between two semiconductors may be realized in three variants: a) junctions of two impurity semiconductors with current carriers of the same sign; b) junctions of two impurity semiconductors with current carriers of different signs; and c) a contact between two intrinsic semiconductors with two current-carrier species -- which will not be discussed here since it is the most complicated in theory and of the least interest in practice.

a) The processes at the junction between two semiconductors with like signs of conductivity are approximately the same as those at the junction of a semiconductor with a metal. A barrier layer with a sharply reduced current-carrier concentration is formed in one semiconductor and an antibarrier layer with a current-carrier concentration higher than that in the volume is formed in the other. The result depends both on the relation between the work functions of the contacting bodies and on the sign of their conductivities. Thus, for example, if electronic semiconductors form the contact, the contact layer of the semiconductor with the lower work function loses free charges and becomes a barrier layer; the contact layer of the other semiconductor, which has the higher work function, becomes an antibarrier layer. The picture is reversed at junctions between hole-type semiconductors: the barrier layer is formed in the conductor with the higher work function and the antibarrier layer in that with the lower.

b) The junction between two impurity semiconductors with unlike signs of conductivity is of the greatest practical interest. Its energy diagram is presented in Fig. 40. The contact field at the boundary of the electron-type (n) and the hole-type (p) semiconductors is due to the loss of mobile charges by both boundary levels. This also effects leveling of electrochemical potentials in the two semiconductors. An excess positive charge appears at the boundary of the n-region in the form of fixed ionized donors and an equal negative charge is created in the boundary of the p-region by acceptors fixed in the lattice; the entire contact boundary layer becomes a layer of high electrical resistance and, for this reason, is called a barrier layer. The thickness of the electron-hole, or, as it is often called, p-n transition, is of the order of 10^{-4} to 10^{-5} cm in the most important cases encountered in practice. The height (V_c) of the potential barrier concentrated in this layer attains values of several tenths of a volt. Since the current carriers in a nondegenerate semiconductor acquire such energies only at a temperature of several thousand degrees, neither electrons from the n-type semiconductor nor holes from the p-type semiconductor can diffuse into the contact layer in appreciable numbers at normal temperatures, so that the latter becomes a barrier layer for mobile charges. One can readily imagine that holes which diffuse from left to right in the direction of the boundary of separation are acted upon in the contact layer by the repulsive forces of the positively charged donors, which are present in excess at the boundary of the n-type semiconductor. Similarly, electrons which diffuse from right to left in their random thermal motion are also acted upon in the contact layer by a force which repels them from the uncompensated negatively-charged acceptors of the boundary layer of the p-type semiconductor into the body of the n-type semiconductor. The bend of curves limiting the forbidden band in the p-n transition shown in Fig. 40 indicates the increase in energy required for penetration of a mobile charge into one or another point of the contact layer.

Fig. 40. Schematic diagram of formation of contact potential difference V_c and barrier layer *d* at electron-hole transition; ++ ionized donors, -- ionized acceptors.

If the concentrations of donors and acceptors are equal in the contacting semiconductors, the contact field is symmetrical with respect to the boundary of separation; if the concentrations of impurities are unequal, the contact field penetrates deeper into the semiconductor with the lower carrier concentration. Here we are considering the simple case in which the impurities are fully ionized at the expense of the thermal energy of the solid.

The technical application of the phenomenon described here is based in one case (photocells) on the very fact of the formation of a contact field in which the separation of the electron-hole pairs created by light occurs; in another important case (rectifiers), the technical application of the contact field is based on the strong dependence of the thickness of the contact layer in which this field is concentrated, the energy level of the potential barrier (V_c) which forms, and, consequently, the resistance of this layer on the direction of the external voltage applied.

Of great importance for the technical utilization of the contact phenomena is the possibility of forming a junction between two semiconductors with unlike conductivity signs not by means of mechanical or welded joints but by creating an electron-hole transition within the unmutilated body of a monocrystalline semiconductor. As experience shows, only by this method of forming p-n junctions can high performance of semiconductor equipment utilizing the contact phenomena be ensured. Special numbers of this series are devoted to more detailed consideration of these problems.

Literature

1. A. V. Joffe, Poluprovodniki i ikh primeneniye (Semiconductors and Their Application), AN SSSR (Acad. Sci. USSR), Moscow-Leningrad, 1956.

2. A. V. Joffe, Fizika poluprovodnikov (Physics of Semiconductors), AN SSSR, 1957.

3. G. Bush, Elektronnaya provodimost' nemetallov (Electronic Conductivity of Nonmetals), UFN (Prog. Phys. Sci.), 47, 2, 258, 1952.

4. F. Seitz, Fizika metallov (Physics of Metals), Moscow-Leningrad, 1947.

5. J. N. Frenkel, Vvedeniye v teoriyu metallov (Introduction to the Theory of Metals), Moscow-Leningrad, 1950).

6. F. Seitz, Sovremennaya teoriya tverdogo tela (Modern Theory of Solids), Moscow-Leningrad, 1949.

7. Poluprovodnikovyye materialy (Semiconductive materials). Collection of Translations, Moscow-Leningrad, 1954.

8. W. Shockley, Teoriya electronuykh poluprovodnikov (Theory of Electronic Semiconductors), IL (Foreign Literature Publishing House), 1953.

9. Nauchnaya literatura po poluprovodnikam (Scientific Literature Pertaining to Semiconductors), bibliography assembled under the direction of V. P. Zhuze. AN SSSR, 1955.

10. A. V. Shpol'skiy, Atomnaya fizika (Atomic Physics), Moscow-Leningrad, 1949.

11. A. S. Kompaneyets, Teoreticheskaya fizika (Theoretical Physics), Moscow, 1955. Part IV, Elements of Statistical Physics.

STAT

[redacted]