THE HALL EFFECT AND RELATED PHENOMENA

E. H. PUTLEY

SEMI-CONDUCTOR MONOGRAPHS

THE HALL EFFECT
AND RELATED PHENOMENA

E. H. PUTLEY
B.Sc., Ph.D., A.Inst.P.
Royal Radar Establishment, Malvern

LONDON
BUTTERWORTHS
1960

PREFACE

SOME years ago I was asked to write an article on the Hall effect for the *T.R.E. Journal*. That article is now seriously out of date, but judging from the number of requests I received and am still receiving for reprints, it appears to have been a useful contribution to the literature on semi-conductors.

In this book I have attempted to bring my original article up to date and at the same time to deal with closely related effects such as the magneto-resistance effect, the other galvano-magnetic effects and the thermo-electric effects. I have discussed the subject from the point of view of an experimentalist but have tried to show the contribution which the study of these effects makes to our understanding of the fundamental properties of solids and to both the intrinsic and extrinsic properties of particular semi-conductors.

It is with great pleasure that I acknowledge the debt I owe to many colleagues at R.R.E. who have helped me in many ways with the development of this study.

E. H. PUTLEY

Malvern
May, 1960

ENGLAND: BUTTERWORTH & CO. (PUBLISHERS) LTD.
LONDON: 88 Kingsway, W.C.2

AFRICA: BUTTERWORTH & CO. (AFRICA) LTD.
DURBAN: 33/35 Beach Grove

AUSTRALIA: BUTTERWORTH & CO. (AUSTRALIA) LTD.
SYDNEY: 6/8 O'Connell Street
MELBOURNE: 430 Bourke Street
BRISBANE: 240 Queen Street

CANADA: BUTTERWORTH & CO. (CANADA) LTD.
TORONTO: 1367 Danforth Avenue, 6

NEW ZEALAND: BUTTERWORTH & CO. (NEW ZEALAND) LTD.
WELLINGTON: 49/51 Ballance Street
AUCKLAND: 35 High Street

U.S.A.: BUTTERWORTH INC.
WASHINGTON, D.C.: 7235 Wisconsin Avenue, 14

Printed in Great Britain by The Whitefriars Press Ltd., London and Tonbridge

CONTENTS

CHAPTER 1

CONDUCTION PROCESSES IN SEMI-CONDUCTORS

1.1. INTRODUCTION

THIS book is an attempt to describe recent developments in the study of electronic conduction in semi-conductors. The emphasis is upon the experimental results and an attempt is made to describe their significance in terms of modern concepts of the solid state.

It will be assumed that the reader has some general familiarity with solid state physics and is seeking a more detailed account of the transport properties. It is hoped that the level of discussion will be suitable for research workers who have been concerned with other aspects of the subject but who are requiring fuller information on the steady state conduction processes than is provided in the many books of a more general nature which are now available. It should also be useful to final year undergraduates.

The first chapter reviews very briefly the development of the study of semi-conductors from the first electron theories to the very sophisticated work that has been carried out in recent years and describes the part which the study of conduction processes has had in this development.

In the following chapters the conduction properties and the experimental techniques for measuring them are described, the formal theory of conduction is considered, and the interpretation of experimental results to obtain more fundamental information is discussed. These chapters are illustrated by describing the behaviour of typical semi-conductors. The properties of the more important semi-conductors are summarized in the last chapter.

1.2. THE HALL EFFECT AND ELECTRICAL CONDUCTIVITY

In view of the important part played by the Hall effect in the study of conduction processes, a few introductory remarks will be made concerning it and its relation to the electrical conductivity. This effect was discovered by E. H. Hall in 1879 [1] during an investigation of the nature of the force acting on a conductor carrying a current in a magnetic field. Hall found that when a magnetic

field is applied at right angles to the direction of current flow an electric field is set up in a direction perpendicular to both the direction of the current and of the magnetic field (see *Figure 1.1*). If in an extended medium the current density is j_x, the magnetic induction B_z and the Hall electric field E_y, then the Hall coefficient R is defined by the equation

$$E_y = Rj_xB_z \qquad \text{.....(1.1)}$$

The conditions under which this is defined are considered more precisely in Chapter 2, where a description of the more important

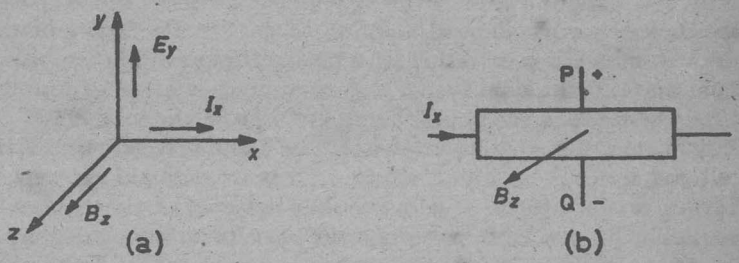

Figure 1.1. The Hall effect. (*a*) Relative direction of current, magnetic induction and Hall field for negative electrons. (*b*) Strip of conducting material: edges parallel to x and y directions. Points P and Q are at opposite ends of an equipotential in the absence of a magnetic field. When induction B_z is applied, P has a positive potential relative to Q (for negative electrons)

transport properties is given. The physical significance of the Hall effect becomes apparent when a current is considered as a stream of electrons moving through the solid. An electron travelling with a velocity $\mathbf{v}$ will experience the Lorentz force $(\mathbf{B} \times \mathbf{v})\, e$. In free space the electron will then be deflected in a direction perpendicular to the plane of $\mathbf{B}$ and $\mathbf{v}$, but the current inside a solid will be constrained within the bounds of the solid. A few electrons will at first be deflected by $\mathbf{B}$ but they will create an electric field which will counterbalance the Lorentz force acting on the bulk of the current carriers, so enabling the current to continue flowing as before. The required electric field is therefore given by

$$(\mathbf{B} \times \mathbf{v})\, e = e\,\mathbf{E} \qquad \text{.....(1.2)}$$

Hence if $\mathbf{v} = (V_x,\ 0,\ 0)$ and $\mathbf{B} = (0,\ 0,\ B_z)$, then $\mathbf{E} = (0,\ -E_y,\ 0)$. Now the relation between current density $\mathbf{j}$ and the electrons' velocity $\mathbf{v}$ (assumed in this simple discussion to be the same for all electrons) is

$$\mathbf{j} = n\,e\,\mathbf{v} \qquad \text{.....(1.3)}$$

where n is the concentration of electrons. Hence from (1.2) and (1.3) we obtain

$$E_y = -\frac{1}{ne} B_z j_x \qquad \ldots\ldots(1.4)$$

So that the Hall coefficient defined by equation (1.1) is inversely proportional to the concentration of conduction electrons:

$$R = -\frac{1}{ne} \qquad \ldots\ldots(1.5)$$

In these equations e denotes the magnitude of the electronic charge. The direction of E_y depends upon the sign of e. The convention is adopted that when e is negative R is taken as negative, as in equation (1.5).

It is because of the simple relation (1.5) between the Hall coefficient and the electron concentration that measurement of the Hall coefficient is so useful in the study of conduction processes. For this reason it undoubtedly occupies the position of *primus inter pares* amongst the transport properties of conductors. In Chapter 3 a more exact derivation of equation (1.5) will be given, when it will be found that the R.H.S. should be multiplied by a numerical factor, but this never departs very far from unity. In some conductors two or more different groups of electrons can contribute to conduction processes. In these cases equation (1.5) must be replaced by an expression combining the contributions of the various groups of carriers. Again, this will be taken up in more detail in Chapters 3 and 4.

A few words must also be said about the electrical conductivity σ. This may be defined by writing

$$j_x = \sigma E_x \qquad \ldots\ldots(1.6)$$

so that, from (1.3)

$$\sigma = ne\frac{V_x}{E_x} \qquad \ldots\ldots(1.7)$$

If Ohm's law is obeyed, as it normally is under the conditions we are concerned with, then σ must be independent of E so that (1.7) may be written

$$\sigma = ne\mu \qquad \ldots\ldots(1.8)$$

μ is the electron drift velocity per unit electric field, or the mobility. This is a characteristic of the conductor and must be, within the validity of Ohm's law, independent of E. Equation (1.8) shows that σ depends upon two characteristics of the conductor: the density n of the charge carriers and their mobility μ. The inter-

pretation of measurements of conductivity is therefore not so straightforward as that of the Hall effect, but the combination of the two yields that very informative quantity, the mobility. From (1.5) and (1.8),

$$|R|\,\sigma = \mu \qquad \text{.....(1.9)}$$

It is a convention to regard μ as a positive quantity regardless of the sign of R. When the properties of the mobility are considered in more detail it will be found that a numerical factor enters into (1.9), as in (1.5), and that it is possible to define the mobility in more than one way. Then the quantity defined by equation (1.9) is called the Hall mobility, while the mobility defined by equation (1.8) is called the conductivity mobility.

1.3. THEORIES OF ELECTRONIC CONDUCTION

The study of the electrical properties of semi-conductors goes back to the beginnings of the study of electricity itself. The discovery of the electron was a great stimulus to research, both experimental and theoretical, on the electrical properties of solids. In this early work both semi-conductors and metals were extensively studied, but although theories such as those of Drude and Lorentz [2] were able to show that the majority of the observed effects could be accounted for by the presence in conducting solids of a gas of free electrons behaving very similarly to an ordinary gas, there were a number of serious discrepancies between theory and experiment. In metals, measurements of the Hall effect indicate that from each atom present roughly one electron can join the gas of free electrons responsible for conduction. By analogy with real gases, it was expected that the free electrons would make a large contribution to the specific heat of metals, but in fact the specific heats of metals and non-conductors are very similar. A second difficulty which emerged was that in the Rutherford–Bohr atom each atomic nucleus is surrounded by a number of electrons sufficient to neutralize the nuclear charge. It therefore followed that all solids must contain many electrons per atom. Yet many of these solids are insulators, while even in the metals only about one electron per atom appears to be concerned in conduction processes. A third difficulty which the classical electron gas theory of conduction could not explain is concerned with the sign of the electronic charge. This, by established convention, is negative. Now the Hall effect and the thermo-electric effects produce electric fields whose direction depends upon the sign of the charge carriers in the conductor. For many metals, such as silver, copper and the alkali metals, these

effects behave as expected for negatively charged electrons, but for several other metals, such as zinc or cadmium, the sign of the Hall and thermo-electric effects indicates a positively charged carrier. To make matters worse, some substances were known, such as silicon and galena (lead sulphide) in which some specimens apparently had negatively charged carriers while others positively charged ones.

That was the position by the 1920s. These difficulties were finally resolved by the early 1930s by the application of quantum mechanics. The specific heat difficulty was resolved by Sommerfeld who appreciated that Fermi–Dirac statistics must be applied to the electron gas. It then turns out that in a metal only a very small fraction of the electrons in the electron gas cloud are free to move under the influence of applied fields and so able to contribute directly to any conduction processes. These electrons will contribute to the specific heat—but, because of the small number of electrons concerned, this contribution is small at normal temperatures compared with that of the atoms of the metal. The Hall effect still indicates the total concentration of electrons in the electron gas. In semi-conductors, where the electron concentration is usually less than 10^{-8} electrons per atom, Fermi–Dirac statistics become in most cases indistinguishable from the classical Boltzmann statistics, which means that all the conduction electrons can move freely.

The second and third of the difficulties of the early theories were removed by the development of the band theory of solids [3]. The electrons attached to a free atom try to arrange themselves in closed groups or shells. When the correct number of electrons is present a very stable structure results which does not take part in chemical reactions (the rare gases) but atoms of other elements do not contain the required number of electrons to completely fill the shells. It is the electrons left over after as many shells as possible have been completed which take part in chemical reactions, tending to unite with electrons from other elements to form ionic or covalent structures which simulate the stable closed-shell structures.

All this refers to isolated atoms or molecules. The band theory of solids shows that analogous behaviour takes place when the atoms or molecules combine to form a solid. It is found that the possible energy levels are arranged in groups or bands of preferred energies separated by ranges of energy containing no permitted levels. The energy levels for the innermost shells are not altered greatly from those of a free atom, but the discrete levels of the outermost electrons are broadened into groups or bands when the distance between the

atoms becomes comparable with the radii of the outer shells (see *Figure 1.2*). These bands become broader as we consider shells further out, until eventually adjacent bands overlap. Now the number of levels in each band is a multiple of the number of atoms making up the solid. Each level represents a state of motion through the crystal but these states of motion are distributed in such a way that if an arbitrary number of electrons is introduced into a band

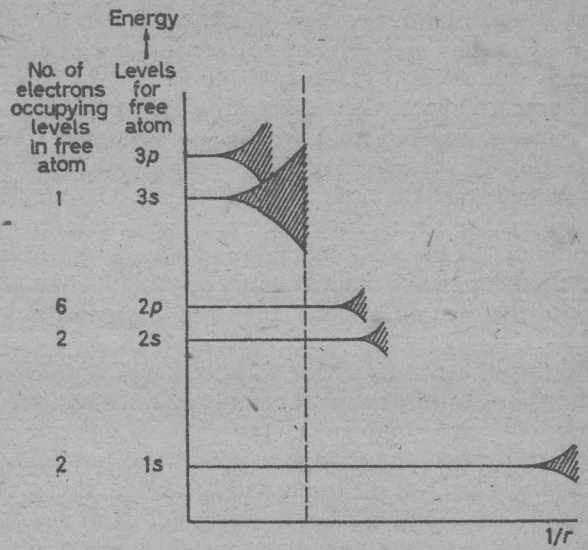

Figure 1.2. Band formation. On the left the energy levels in a free sodium atom are represented. The broadening is plotted against $1/r$, where r is the distance between atoms. The broken line indicates the lattice distance in a sodium crystal (not to scale)

the net motion of the electrons will be zero in the absence of applied forces. If the system is disturbed from this state by the application of, say, an electric field or a temperature gradient, then the electrons will re-distribute themselves amongst the levels to set up a current. In the bands the levels are so close together that for most purposes the distribution of states can be regarded as continuous, although exceptions to this may occur when high magnetic fields are applied. It follows from this description of an energy band that if a band is completely filled the electrons in it cannot re-arrange themselves when a field is applied and so these electrons cannot contribute to

conduction processes unless some of them can move into an adjacent band. This latter process is unlikely unless the bands overlap, because the energy separation between bands will usually be much greater than the mean thermal energy kT while the energy which an electron can derive from the normal values of applied fields will only be comparable with kT at the most.

The conclusion that a band filled with electrons cannot contribute to conduction processes explains the paradox that although all solids must contain large numbers of electrons, not all solids are conductors. If a solid contains the correct number of electrons to fill completely a number of the bands and then there is an energy gap, large compared with kT, between the highest filled band and the lowest empty one, the solid will be an insulator. Diamond (C), quartz (SiO_2), corundum (Al_2O_3) and periclase (MgO) are good examples of crystalline solids of this type. The highest filled band is usually called the valence band and the next empty one the conduction band. If the solid does not contain enough electrons to fill completely the highest occupied band, or if the highest occupied band overlaps the next empty one, then the solid will, in general, be a metal.

The band theory accounts for the third of the difficulties encountered by the classical theories, that of the apparent positive charge of the carriers in some conductors. As we have just said, a completely full band cannot contribute to conduction processes, but a band containing a few vacant levels will be able to do so. Bloch [4] showed that the contribution from a nearly full band was formally equivalent to that from a similar band in which the levels corresponding to the empty ones in the first band were occupied by positively charged carriers and the remaining levels were empty. Thus in a nearly full band the vacant levels are equivalent to positively charged carriers and are usually referred to as 'positive holes'. Those conductors having positive Hall coefficients or thermoelectric powers are ones in which the band containing the conduction electrons is more than half full.

It is helpful at this point to consider the application of quantum theory to a gas of free electrons enclosed in a volume V. At absolute zero the electrons would occupy a series of levels up to a maximum energy ϵ_{max}. This energy is determined by two considerations. First, to satisfy the Pauli exclusion principle only two electrons (of opposite spin) can occupy each level. Secondly, the energy spacing between adjacent levels is determined by Heisenberg's uncertainty principle to be h^3/V. From these two conditions it is possible to calculate the number of levels in an energy interval $d\epsilon$ and the

energy of the highest occupied level when the system contains N electrons. The results obtained are

$$N(\epsilon)\mathrm{d}\epsilon = 4\pi V\left(\frac{2m}{h^2}\right)^{3/2}\epsilon^{1/2}\,\mathrm{d}\epsilon \qquad \text{.....(1.10)}$$

and

$$\epsilon_{\max} = \frac{h^2}{8m}\left(\frac{3N}{\pi V}\right)^{2/3} \qquad \text{.....(1.11)}$$

The importance of these simple results is that they may be applied as a first approximation to the broader bands in a solid. The density of levels given by equation (1.10) increases indefinitely with ϵ. This of course cannot be correct for a band which only contains a finite number of levels, but equation (1.11) is often a good approximation to the shape of a band at energies near its

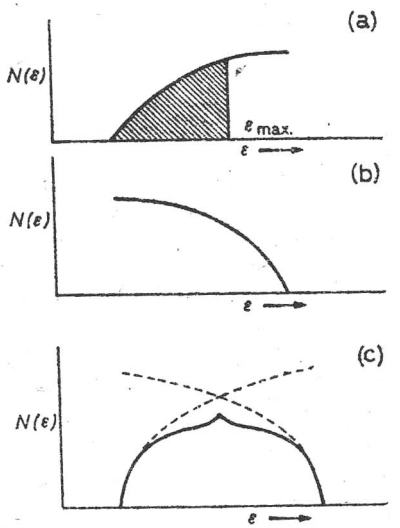

Figure 1.3. Level density in simple band. (*a*) Density of allowed levels for a quantum electron gas (equation (1.10)). (*b*) Inverted form used for approximation at energy maximum of band. (*c*) Use of (*a*) and (*b*) as approximations at extremities of actual band

minimum. When applied to a band the electron mass m must be replaced by an effective mass m^* which is usually of the order of magnitude of m. This takes account of the fact that the density of

levels in a band is determined by the extent of the interactions between the atoms of the solid. It is also found that near the energy maximum of a band the density of levels varies in a similar way to that of equation (1.10), but with $\epsilon^{1/2}$ replaced by $(\epsilon_{max} - \epsilon)^{1/2}$. *Figure 1.3* illustrates equation (1.10) and the approximate way it can be fitted to an energy band. These approximations at the extremities of bands are especially useful in considering semi-conductors because the number of electrons or holes in a band is almost always so small as to make expressions similar to (1.10) valid.

1.4. THE WILSON MODEL OF A SEMI-CONDUCTOR

So far we have mentioned conductors and insulators, and while the band theory explains the reason for the extreme differences between the properties of a good insulator and of a metal, it also shows that one should not expect an abrupt distinction between these two types of solid. In a solid in which the bottom of the conduction band is an energy ϵ_g above the top of the valence band a small number of electrons will be excited thermally into the empty band. The number will be proportional to a Boltzmann factor $\exp(-\epsilon_g/kT)$ so that it will vary very rapidly with temperature. The insulators mentioned in the last paragraph but one are all transparent in the visible. This means that the energy required to excite an electron from the valence to the conduction band is greater than that of visible light, and is in fact of the order of 5–10 eV. On the other hand, at room temperature $kT \simeq 1/40$ eV. Hence at room temperature the Boltzmann factor will permit a negligible number of electrons to be excited. However, it is easy to suppose that if in some substance ϵ_g were small enough, then an observable number of electrons could be excited at a moderate temperature (*Figure 1.4*). In such a substance the variation of the Boltzmann factor with temperature would lead to an exponential variation of conductivity. This type of behaviour had been observed by Königsberger[5] who pointed out the possibility of its being associated with the temperature variation of the carrier concentration. It was some time, however, before this simple mechanism of excitation from the valence to the conduction band was unambiguously established experimentally. When this process takes place, electrons in the conduction band and holes in the valence band both contribute to conduction. This process is known as 'intrinsic conduction'.

The difficulty was that this excitation process is clearly a characteristic of a particular substance and therefore similar behaviour

would be expected in different samples of the same substance, much in the same way that different samples of the same metal show similar characteristics. But the differences between the behaviour of different samples of semi-conductors were so great as to make this explanation seem unlikely. Thus all samples of silicon or galena do not have the same sign Hall coefficient and it was found that the conductivity of highly conducting oxides such as Cu_2O or ZnO depended upon the ambient oxygen pressure. These difficulties led Wilson [8] to propose an alternative mechanism to account for the presence of conduction electrons in those substances that were neither metals nor good insulators nor ionic conductors.

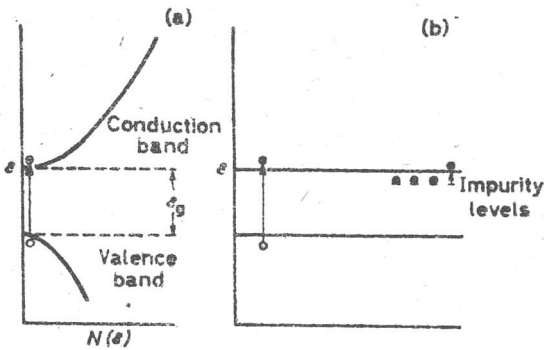

Figure 1.4. Intrinsic conduction. (*a*) Density of levels plotted for two bands separated by energy gap ϵ_g. A hole in the valence band is formed when energy not less than ϵ_g is supplied to the system. The electron from the valence band is liberated in the conduction band. The scheme is often represented more diagrammatically as in (*b*), which also shows some impurity levels which can supply electrons to the conduction band

The band scheme so far discussed supposes that the crystal is a perfectly periodic structure containing no imperfections of any kind. If now an extra atom is inserted somewhere inside the crystal it will spoil the perfect periodicity of the system. This will be reflected in the band structure. Consider, for example, the conduction band. The number of levels in the band will still be a multiple of the total number of atoms present (including the extra one), but while the levels derived from the normal atoms will be grouped closely together and represent states of motion through the crystal, the levels associated with the extra atom can be displaced

from the main group of levels and may fall into the energy gap between this band and the filled band. These levels represent a localized state centred on the extra atom. Thus the introduction of an extra atom represents a physical perturbation of the crystal and is reflected in the band scheme by a displacement of some of the levels which may lead to the appearance of levels in the gap between the full and empty bands (see *Figure 1.4*(b)).

There are, of course, many ways in which an extra atom can be introduced. For an example, consider lead sulphide. This forms an ionic crystal of rock salt structure in which Pb^{2+} ions and S^{2-} ions alternate. Suppose an extra Pb atom were inserted. It might occupy an interstitial position or it might occupy a normal Pb^{2+} site. In the latter case a vacant S^{2-} site must also be formed, so that adding a Pb atom would be equivalent to removing an S atom. In the resulting crystal, localized levels will be associated with the Pb atom and in addition the system will contain two electrons in excess of those required to ionize all the sulphur atoms. Now in a perfect crystal there are just sufficient electrons to fill completely the valence band. The two extra electrons now present will, in the configuration of lowest energy, occupy the localized levels associated with the extra lead atom. This simply means that they stay with their parent atom. If, however, these levels are located in the gap between the valence and conduction bands, then the energy required to excite them into the conduction band will be less than that required to excite electrons from the valence band. Thus this mechanism will provide mobile conduction electrons at a lower temperature than the intrinsic mechanism.

When the impurity mechanism is dominant, the number of conduction electrons will depend upon the amount of impurity in a given specimen, so that, unlike the intrinsic mechanism, this mechanism can account for the wide difference in conductivity of different specimens of the same material. If in lead sulphide an extra sulphur atom is introduced, a corresponding behaviour takes place, but in this case the sulphur centre forms levels close to the top of the valence band. These levels trap electrons from this band, leaving two holes in the band. Again conduction is now possible, but in this case the Hall effect will have a positive sign. An alternative physical picture of the behaviour when one extra sulphur atom is present is obtained by noting that all sulphur atoms tend to capture two electrons from the lead atoms. If there is one extra S atom they cannot all be ionized. At absolute zero the un-ionized atom would be located near the corresponding Pb^{2+} vacant site, but at higher temperatures the electrons