

# The Hall effect – an important diagnostic tool

E.H. Rhoderick, UMIST, Manchester, UK

The quantitative measurement of carrier concentrations and mobilities is of vital importance in the assessment of semiconductor materials. One of the best-established methods of doing this relies on the Hall effect, which was discovered as long ago as 1880 by E.H. Hall in his investigations of metals. This article introduces the basic theory of the Hall effect at a fairly elementary level, with a minimum of mathematics, and discusses how it can be used to determine semiconductor parameters. It does not attempt to cover the more advanced developments of the subject, such as the quantum Hall effect.

Hall's observation was that, if a current is passed through a conductor (in his case, a metal) in a magnetic field which is perpendicular to the current, then an electric field is generated which is mutually at right angles to both the current and the magnetic field. This electric field can be detected as a potential difference between two suitably placed electrodes, and is proportional to the magnitudes of both the current and the magnetic field.

In practice one usually considers a specimen in the form of a

rectangular bar, and the basic geometry is as shown in Figure 1. A right-handed system of co-ordinates is chosen, with the direction of the current taken as the  $x$ -direction and that of the magnetic field as the  $z$ -direction. The induced electric field is in the  $y$ -direction, and is detected by the voltage induced between two contacts, A and B, opposite each other. We will assume that the current flows in the positive  $x$ -direction and that it is carried by electrons, as in an  $n$ -type semiconductor, so the electrons must move in the negative  $x$ -direction.

The Hall effect arises because if a particle carrying a charge  $q$  is moving with velocity  $\mathbf{v}$  in a magnetic field of flux density  $\mathbf{B}$ , where the bold type denotes vectors, it experiences a force (Lorentz force) in a direction mutually perpendicular to both  $\mathbf{v}$  and  $\mathbf{B}$ , of magnitude  $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$ . (A similar force is experienced by electrons in a magnetically deflected cathode-ray tube.) For the particular choice of directions shown, and bearing in mind that  $\mathbf{v}$  is in the negative  $x$ -direction,  $\mathbf{F}$  must be directed downwards in the negative  $y$ -direction, and of magnitude  $F_y = q_e v_x B_z$ , where  $q_e$  is the charge on the electron ( $q_e$  and  $v_x$  are both negative).

This downwards force tends to make the electron trajectories bend downwards, as shown by the dotted lines in Figure 1, thus concentrating the electrons at the bottom surface of the slab\*. Since the electrons are negatively charged, this generates an electric field  $E_y$  (the 'Hall field') vertically downwards (i.e.  $E_y$  is negative), which in turn gives rise to a force  $q_e E_y$  on each electron in the upward direction. Equilibrium is reached when the upwards force  $q_e E_y$  balances the downwards force of

\* In many elementary treatments of the Hall effect, the electrons are shown as being bent upwards. This is because they take the magnetic field to be into the paper, i.e. in the negative  $z$ -direction.

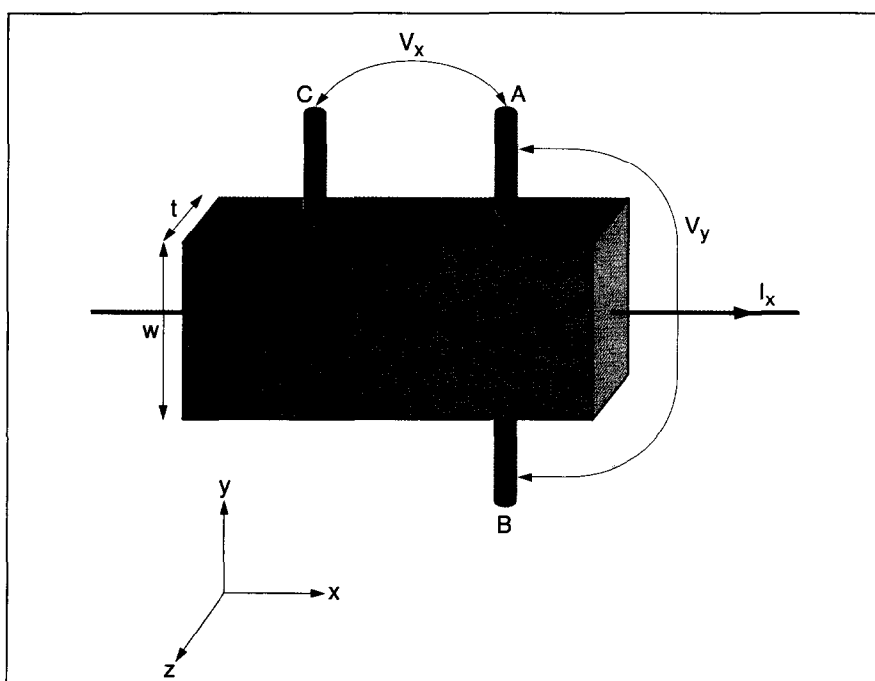


Figure 1. Schematic drawing of specimen geometry, showing dimensions and directions of axes. Note that five contacts to the sample are indicated.

magnitude  $q_e v_x B_z$ , in which case we must have

$$E_y = v_x B_z. \quad (1)$$

The Hall field  $E_y$  generates a potential difference  $V_y$ , called the Hall voltage, between the two contacts A and B. The ratio  $E_y/J_x B_z$  is called the 'Hall coefficient',  $R_H$ . Since  $E_y$  is negative,  $R_H$  is negative for *n*-type semiconductors.

An alternative expression for  $R_H$  can be obtained if we can find the value of  $v_x$  in equation (1). To find  $v_x$  we shall make the very crude assumption that we can neglect the random motion of the electrons due to their thermal agitation, and that they all have a uniform velocity  $v_x$  equal to the drift velocity  $-\mu E_x$ , where  $\mu$  is the mobility of the electrons. (The minus sign arises because the electrons move in the opposite direction to the electric field.) Hence the current density in the *x*-direction,  $J_x$ , is given by

$$J_x = q_e v_x n \quad (2)$$

where  $n$  is the electron density per unit volume. From equations (1) and (2) we have

$$\begin{aligned} R_H &= E_y/J_x B_z = v_x B_z/q_e v_x n B_z \\ &= 1/q_e n. \end{aligned} \quad (3)$$

Since  $q_e$  is negative, this is in accordance with the previous result that  $R_H$  is negative for *n*-type semiconductors.

A very useful result can be obtained if we also measure the conductivity  $\sigma = J_x/E_x$ . The conductivity can be measured by making use of the results  $J_x = I_x/wt$  and  $E_x = V_x/d$ . As we saw in equation (2),  $J_x = q_e v_x n$  where  $v_x = -\mu E_x$ , so that  $\sigma = -\mu q_e n$ . Hence, from (3), we have

$$\sigma R_H = -\mu. \quad (4)$$

This forms the basis of a very useful method of measuring  $\mu$ .  $V_x$  is measured across contacts C and A (Figure 1).

If the charge carriers are holes, as in a *p*-type semiconductor, the analysis remains valid with the substitution of  $q_h$  for  $q_e$  and the hole concentration  $p$  for  $n$ , so that

$$R_H = 1/q_h p. \quad (5)$$

Since  $q_h$  is positive,  $R_H$  is positive. The physical explanation of this is that because  $v_x$  and  $q_h$  are both positive, the Lorentz force remains downwards and the holes are deflected downwards just as in the case of electrons, as shown in Figure 1. But since the holes are positively charged, the Hall field  $E_y$  is positive (i.e. directed upwards), making the Hall coefficient positive.

## Units

Using mks units, with  $n$  in  $m^{-3}$ ,  $q$  in coulomb (C),  $B$  in Tesla,  $J$  in  $Am^{-2}$ , and  $E$  in  $Vm^{-1}$ ,  $R$  has dimensions of  $m^3C^{-1}$ . (Using cgs units, with  $n$  in  $cm^{-3}$ ,  $q$  in C,  $B$  in gauss,  $J$  in  $Acm^{-2}$ , and  $E$  in  $Vcm^{-1}$ ,  $R$  has dimensions of  $cm^3C^{-1}$ .)

One does not normally measure the Hall field  $E_y$  or the current density  $J_x$  directly, but rather the current  $I_x$ , the width  $w$  and thickness  $t$  of the specimen, and the Hall voltage  $V_y = wE_y$ . We then have, for mks units,  $R_H = V_y t/I_x B_z$ , (cgs units,  $R_H = 10^8 V_y t/I_x B_z$ ). If we recast these last two relationships to give expressions for  $V_y$  in terms of the other parameters, it is evident that  $V_y$  is inversely proportional to  $t$ , so that to obtain an easily measurable Hall voltage we need to use very thin specimens.

## Refinements of the theory

The above theory was developed making the naive assumption that charge carriers move in straight lines with the drift velocity  $v_x$ .

This is an extremely crude assumption, since in reality the carriers move in all directions in a random manner due to thermal agitation, and the drift velocity is

merely a relatively small velocity in the *x*-direction superposed on this thermal motion.

Taking the thermal motion into account is a much more complicated problem involving the higher realms of transport theory, and we shall not attempt it here, but merely state the main results. The most important effect of the thermal motion is to substitute for equation (3) the more complicated form

$$R_H = r_c / q_e n \quad (6)$$

where  $r_c$  is the 'scattering factor' for electrons, which is a complicated function of the energy distribution of the electrons (e.g. whether they are 'non-degenerate' as in a weakly-doped semiconductor, or 'degenerate' as in a metal) and also depends on the scattering mechanism (e.g. whether the electrons are scattered by collisions with impurities or by vibrations of the crystal lattice). The result is that  $r_c = \langle \tau^2 \rangle / \langle \tau \rangle^2$ , where  $\tau$  is the average time between collisions. In words,  $r_c$  is the ratio of the mean value of  $\tau^2$  to the square of the average value of  $\tau$ ; it can be calculated theoretically and ranges between 1 and 2, depending on what assumption is made about the scattering mechanism. If we do not know the exact value but assume it to be 1.5, we can be reasonably confident that the value of  $n$  obtained from Hall measurements is accurate to about 30%, apart from any experimental errors that may occur.

Another consequence of the thermal motion following from the revised expression for  $R_H$  given by (6) is that the product  $\sigma R_H$  now gives  $-r_c \mu$  for electrons, rather than  $-\mu$ , so that determinations of the mobility obtained from Hall measurements always give  $r_c \mu$ . Because of this, the product  $r_c \mu$  is often termed the 'Hall mobility' and denoted by  $\mu_H$ , to distinguish it from the ordinary drift mobility  $\mu$ , which is denoted by  $\mu_d$ .

A similar result applies to holes, so that equation (5) becomes

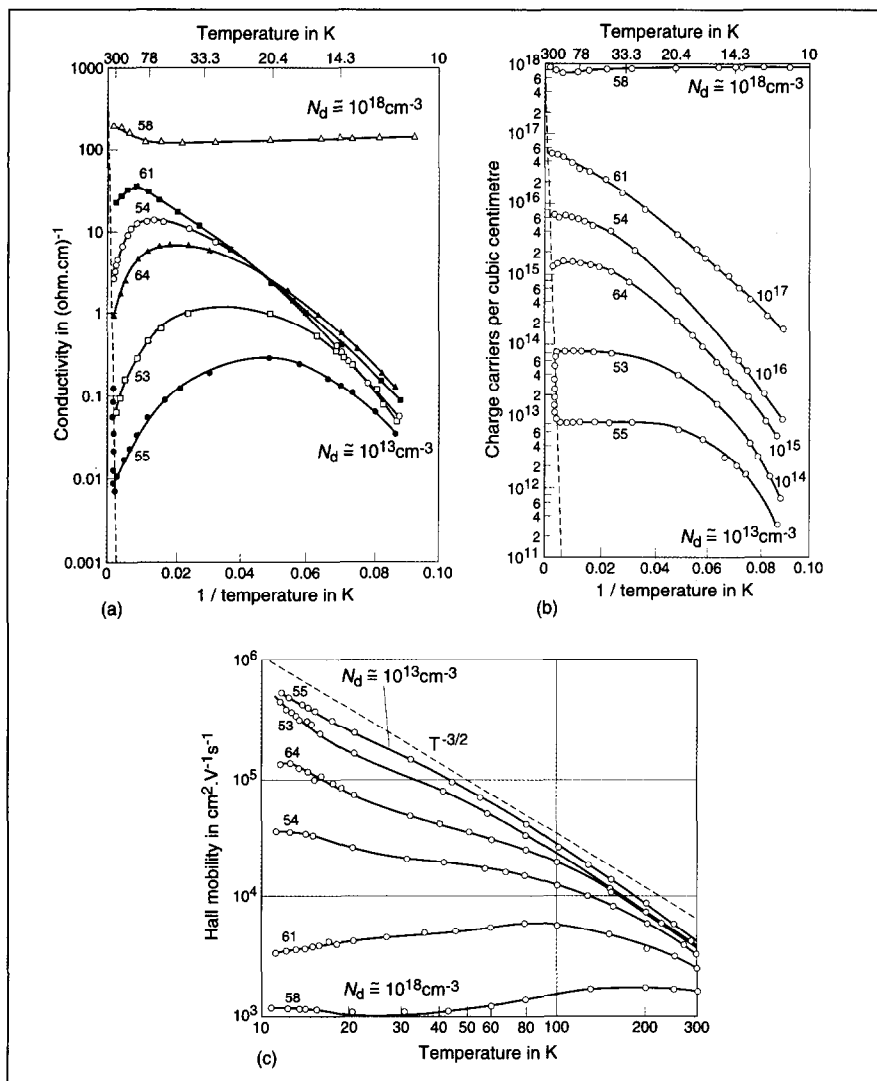


Figure 2. (a) Conductivity, (b) electron concentration, and (c) Hall mobility of *n*-type germanium. Data by P.P. Debye, after E.M. Conwell, *Proc. I.R.E.* 40, 1327 (1952).

$$R_H = r_H / q_H p. \quad (7)$$

The scattering factor for holes is not necessarily the same as that for electrons because the holes generally will not have the same energy distribution as electrons and do not respond to the scattering mechanisms in the same way.

## Some typical results

The use of Hall effect and conductivity measurements to measure carrier concentrations and mobilities is well brought out by the classic measurements of Debye and Conwell on *n*-type germanium containing arsenic donors. Figure 2(a) shows the conductivity  $\sigma$  measured from  $\sigma = I_x d / V_x w t$ ,

where  $d$ ,  $w$ , and  $t$  are as shown in Figure 1, and Figure 2(b) the electron concentration  $n$  deduced from the Hall coefficient through equation (6), both plotted logarithmically as functions of the reciprocal of the absolute temperature,  $T^{-1}$ .

In evaluating  $n$ , the scattering factor  $r_e$  is assumed to have the theoretical value, ranging between  $3\pi/8$  ( $=1.18$ ) for weakly doped specimens, where the scattering is predominantly by thermal lattice vibrations, and  $315\pi/512$  ( $=1.93$ ) for the most strongly doped specimen, where the electrons are scattered by impurity ions.

For  $N_d \gg 10^{18} \text{cm}^{-3}$ , the donor binding energy is effectively zero and  $n$  is practically independent

of  $T$ . For low donor concentrations ( $N_d < 10^{15} \text{cm}^{-3}$ ), the electron concentration is independent of  $T$  when  $T^{-1} < 0.02$ , corresponding to the 'exhaustion' region where all the donors are ionised, but for  $T^{-1} \approx 0.04$ ,  $n$  is a linearly decreasing function of  $T^{-1}$  with a slope corresponding to a donor binding energy of about  $10^{-2}$  eV, which agrees with the known donor binding energy in germanium.

Figure 2(c) shows the electron Hall mobility,  $\mu_H$ , calculated from the product  $\sigma R_H$ , as a function of  $T$ . For the highest doping concentration,  $\mu_H$  is very low and practically independent of  $T$  because at low temperatures the scattering is predominantly due to impurity ions, while for the purest specimen the mobility is much higher and approaches a dependence on  $T^{-3/2}$ , as would be expected for scattering by thermal lattice vibrations.

## The effect of conduction by both electrons and holes

So far we have considered the current to be carried by either electrons or holes, but not both. However, it often happens that both types of carrier are important. For example, in a low-bandgap semiconductor there may be a significant concentration of minority carriers; or in some semiconductors the mobility of electrons may be so much greater than that of holes (e.g. indium antimonide, where the ratio is about 100 at room temperature) that a relatively small concentration of electrons may make a greater contribution to the current than the holes. In such cases we have to take both carriers into account in calculating the Hall coefficient.

The analysis is now complicated by the fact that it is not possible for the force due to the Hall electric field to balance the Lorentz force for both carriers simultaneously

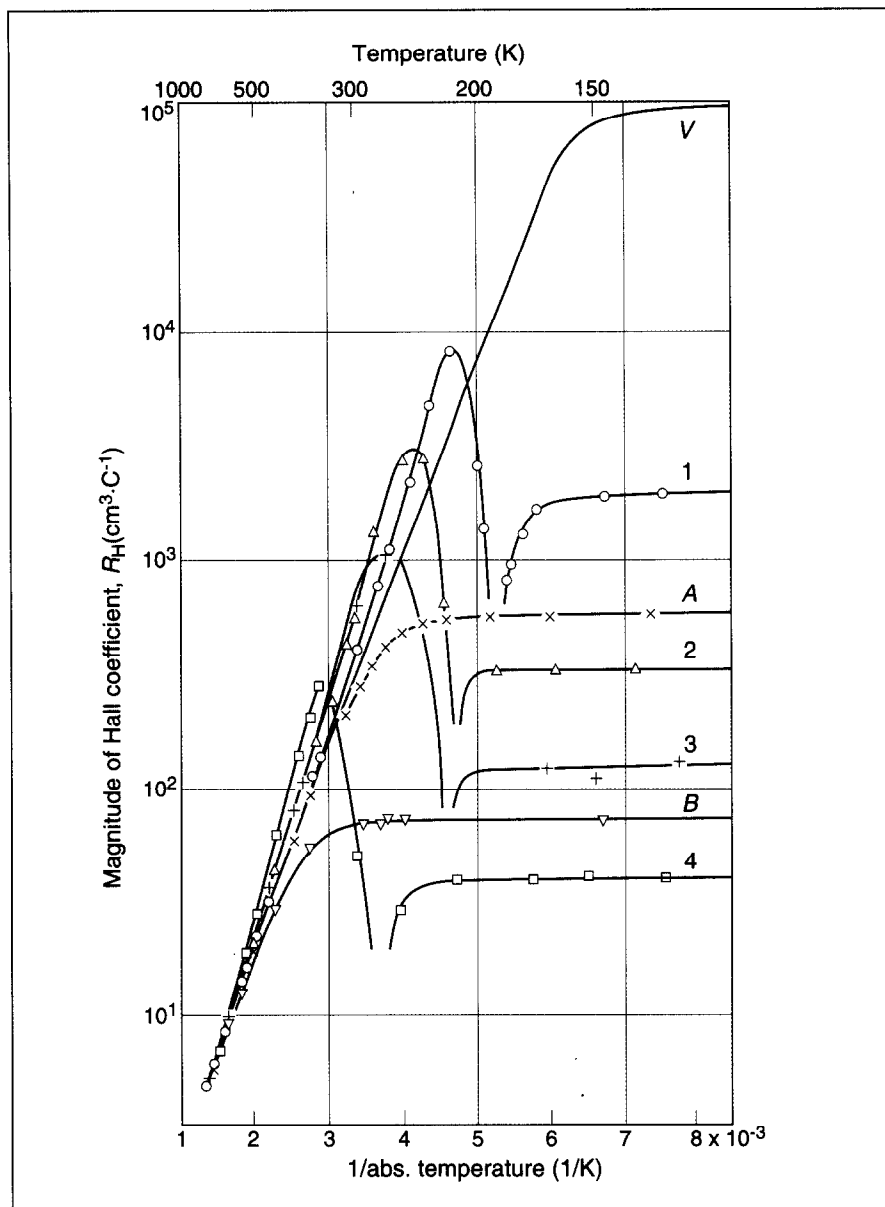


Figure 3. Magnitude of Hall coefficient of three *n*-type (A, B, V) and four *p*-type (1-4) samples of indium antimonide, after O. Madelung, *Physics of III-V Compounds*, J. Wiley & Sons (New York), 1964, p.113.

because their drift velocities are not the same, so there will be a net movement of both types in the  $y$ -direction. Hence, instead of requiring that there is no current in the  $y$ -direction, we now have to rely on the weaker condition that the sum of the hole current density  $J_{yh}$  and the electron current density  $J_{ye}$  must be zero.

This means that holes and electrons leave the top surface of the specimen in Figure 1 and arrive at the bottom surface at the same rate; in other words, hole-electron pairs must be

generated at the top surface and recombine at the bottom surface, and in order that the theoretical analysis shall be valid there must be an adequate generation-recombination rate at both surfaces. This is usually secured in practice by damaging the surfaces by sandblasting.

An analysis similar to that given previously, but making the assumption that  $J_y = 0$ , is straightforward but tedious. The result is that

$$R_H = (r_h p - r_e b^2 n) / e(p + bn)^2 \quad (8)$$

where  $r_h$  and  $r_e$  are the scattering factors for holes and electrons respectively,  $b$  is the mobility ratio  $\mu_e/\mu_h$  and  $e$  is the magnitude of the electronic charge, so that  $e = q_h = -q_e$ . We may note from (8) that if  $n = 0$  then  $R_H = r_h/ep$ , and if  $p = 0$  then  $R_H = -r_e/en$ , in accordance with equations (7) and (6) respectively.

An interesting feature of equation (8) is that since  $b$  can be quite large in III-V semiconductors, the second term in the numerator may be larger than the first even if  $n$  is less than  $p$ , so that the Hall constant may be negative in *p*-type semiconductors. This behaviour is well illustrated in Figure 3, which shows some Hall effect measurements on indium antimonide by Madelung and Weiss.

Curves 1 to 4 all refer to *p*-type specimens, and show the magnitude of the Hall constant plotted against the inverse of the absolute temperature. (In this context, *p*-type means that the specimen is doped with acceptors, not that the Hall coefficient is positive!).

At low temperatures (high values of  $1/T$ ),  $R_H$  is independent of  $T$  and is positive. This corresponds to the situation in which all the acceptors are ionised, so that  $p \approx N_a$ , and  $n$  is too small for the second term in the numerator of (8) to be greater than the first. As  $1/T$  decreases,  $T$  rises, the intrinsic concentration  $n_i$  increases, and the electron concentration increases so that the product  $pn$  is always equal to  $n_i^2$ . The contribution of  $r_e b^2 n$  increases so that the magnitude of  $R_H$  begins to fall, and eventually  $R_H$  becomes zero when  $r_h p = r_e b^2 n$  and then changes sign. The three samples A, B, & V, are *n*-type (i.e. they are doped with donors) and their Hall coefficient is always negative because the contribution of the holes to  $R_H$  is never large enough to outweigh the contribution of the electrons.

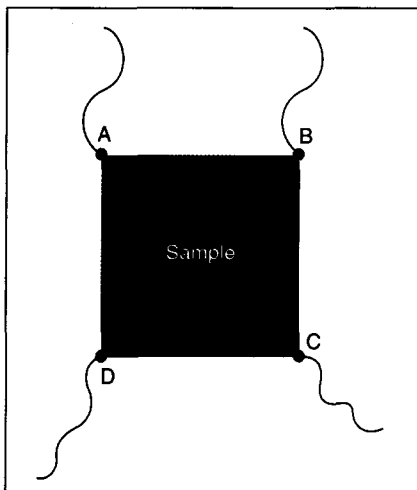


Figure 4. The Van der Pauw four-contact technique. The contacts to the periphery of the sample must be point-like and ohmic. Only if the geometry is perfectly regular will equations (9) and (10) apply (the thickness of the sample is  $t$  and it is assumed that the magnetic field passes orthogonally through the plane of the figure). Note that the distance between adjacent contacts is not a parameter in these equations.

## Experimental techniques (by M R Brozel)

### Equipment

There are three basic components to a Hall effect kit; a constant current generator to develop  $I_x$ , a high-impedance voltmeter to measure  $V_y$  and a magnet to generate  $B_z$ .

The constant current generator is a specialised power supply usually arranged to pass a current of either 1 mA or 10 mA through the sample. The current is selected to produce an easily measurable Hall voltage within the restriction that heating effects of the current must be kept to a minimum. Suitable instruments are made by Keithley.

The voltmeter must have a very high input impedance if accurate measurements are to be made. For reasons associated with the need to subtract "offset voltages" that result from non-symmetries in either the sample geometry or doping properties, the voltmeter must have an extremely high precision. A typical DVM for this application needs to have 8-digit resolution with a

precision of measurement that allows all the resolution to be useful. These DVMs have sub-microVolt resolution and do not come cheap. Many commercial instruments use DVMs supplied by Keithley.

The final component is the magnet. In general, the greater the magnetic field, the greater the Hall voltage that is produced for a given current. However, cost, size and the requirement for an electromagnet to have a large dedicated power supply often restrict options to a permanent magnet of a few tenths of a Tesla field strength.

### Technique

The contacts are often simple point contacts produced on the four corners of a square sample (Figure 4). The analysis of this simple geometry was made by Van der Pauw. If we label the four contacts in either clockwise or anticlockwise directions as A, B, C and D (the size of the sample is irrelevant), then the conductivity (without a magnetic field applied) is given by,

$$\sigma = (\ln 2 / \pi t) \cdot (I_{AB} / V_{CD}). \quad (9)$$

The Hall coefficient (with a magnetic field) is given by,

$$R_H = (t/B) \cdot (V_{BD} / I_{AC}). \quad (10)$$

It is the asymmetry of the geometry of the four contacts, or of the properties of the sample itself, that results in the offset voltages that exist when a current is passed (but before the application of a magnetic field) in the measurement of  $R_H$ .

The need to make allowances for these offset voltages (which can exceed the Hall voltage by orders of magnitude) usually requires that the currents and voltages must be switched between the two sets of opposite sample contacts. A high-precision switching unit that introduces a minimum of contact potential is needed for this.

Also, it is necessary to perform measurements either with or

without the magnetic field, or to reverse the magnetic field. Sometimes the sample and its holder are simply twisted by  $180^\circ$  in the field: in more automated kits the magnet can be removed or rotated. Of course, if an electromagnet is used, its current can simply be turned off or reversed.

## Improvements to the basic Hall effect measurement

### 1. Producing a set of Hall data over a range of sample temperatures:

This data, already shown in Figure 2, can be used to determine dopant ionisation energies, bandgap energies and the temperature dependence of the carrier mobility. The latter gives information on scattering processes and hence the purity of the semiconductor. Obviously, there is considerable extra expenditure in this type of measurement over the simple room-temperature measurement because of the need for cryostat and temperature control equipment. However, when automated, this approach can produce a large amount of data in a couple of hours or so. As a compromise, data taken at 300K and 77K only can be compared to give a rough idea of material properties.

### 2. Sequential etching / Hall effect measurements:

In this modification the differential changes in Hall parameters are measured as the surface is sequentially removed by dipping the sample in liquid etchant. This type of measurement is made at room temperature only and allows the depth profiles of carrier concentrations in layers to be determined. However, substantial errors can be incurred in the subsequent analysis. In general, this approach is suitable only for layers of one conductivity type and the presence of heterojunctions can make the method very difficult to employ. Typical applications include the assessment of the uniformity of epitaxial layers and the assessment of ion-implanted layers.

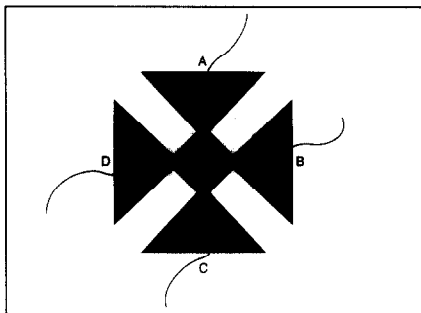


Figure 5. The use of a clover-leaf geometry to the Van der Pauw configuration makes the fabrication of the contacts much simpler. As in Figure 4, the magnetic field passes orthogonally through the sample.

3. The use of "clover leaf" or "Maltese Cross" samples (Figure 5): The geometry is optimised for this sample shape. In bulk or substrate material, the clover-leaf shape is often produced by sand blasting the rest of the sample away, the original sample being held behind a cross-shaped piece of metal during sand blasting. This approach can be applied to epitaxial layers grown on an

insulating substrate, although in this case it may be preferable, if not easier, to remove the epi-layer by a chemical etch. Only the layer need be removed, most of the substrate being retained for mechanical strength.

## Practical applications of the Hall effect

In addition to the Hall effect being a useful technique for semiconductor assessment, it is important to realise that it has been extensively used in instrumentation. This is because, once the conduction of the semiconductor has been calibrated, the Hall effect becomes a sensitive measurement of magnetic field. Thus, applications include fluxmeters for physical, geological and other uses and magnetic proximity probes. The material of choice is InSb, which can be obtained in a very pure, though

residually n-type, form and displays a large Hall coefficient

## Further reading

Those wishing to explore the subject more fully may find the following useful:

- E.H. Putley: *The Hall Effect and Related Phenomena*, Butterworths (London), 1959.
- H.H. Weider: Chap 2 of *Non Destructive Evaluation of Semiconducting Materials and Devices* (ed. J.N. Zemel) Plenum Press (New York), 1979.

### Contact:

E.H. Rhoderick,  
Emeritus Professor of Solid State Electronics,  
Department of Electrical Engineering and Electronics,  
UMIST, PO Box 88,  
Manchester M60 1QD, UK.  
E-mail: [erhod@beeb.net](mailto:erhod@beeb.net)

