14-3 The Hall effect

It is certainly a peculiar thing that in a substance where the only relatively free objects are electrons, there should be an electrical current carried by holes that behave like positive particles. We would like, therefore, to describe an experiment that shows in a rather clear way that the sign of the carrier of electric current is quite definitely positive. Suppose we have a block made of semiconductor material—it could also be a metal—and we put an electric field on it so as to draw a current in some direction, say the horizontal direction as drawn in Fig. 14-6. Now suppose we put a magnetic field on the block pointing at a right angle to the current, say *into* the plane of the figure. The moving carriers will feel a magnetic force $q(v \times B)$. And since the average drift velocity is either right or left—depending on the sign of the charge on the carrier—the average magnetic force on the carriers will be either up or down. No, that is not right! For the directions we have assumed for the current and the magnetic field the magnetic force on the moving charges will always be *up*. Positive charges moving in the direction of j (to the right) will feel an upward force. If the current is carried by negative charges, they will be moving left (for the same sign of the conduction current) and they will also feel an upward force. Under steady conditions, however, there is no upward motion of the carriers because the current can flow only from left to right. What happens is that a few of the charges initially flow upward, producing a surface charge density along the upper surface of semiconductor—leaving an equal and opposite surface charge density along the bottom surface of the crystal. The charges pile up on the top and bottom surfaces until the electric forces they produce on the moving charges just exactly cancel the magnetic force (on the average) so that the steady current flows horizontally. The charges on the top and bottom surfaces will produce a potential difference vertically across the crystal which can be measured with a high-resistance voltmeter, as shown in Fig. 14-7. The sign of the potential difference registered by the voltmeter will depend on the sign of the carrier charges responsible for the current.

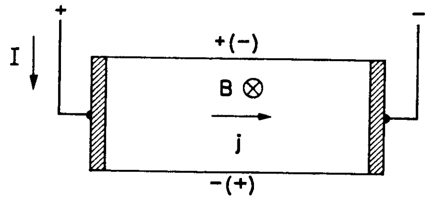

Fig. 14-6. The Hall effect comes from the magnetic forces on the carriers.

When such experiments were first done it was expected that the sign of the potential difference would be negative as one would expect for negative conduction electrons. People were, therefore, quite surprised to find that for some materials the sign of the potential difference was in the opposite direction. It appeared that the current carrier was a particle with a positive charge. From our discussion of doped semiconductors it is understandable that an n-type semiconductor should produce the sign of potential difference appropriate to negative carriers, and that a p-type semiconductor should give an opposite potential difference, since the current is carried by the positively charged holes.

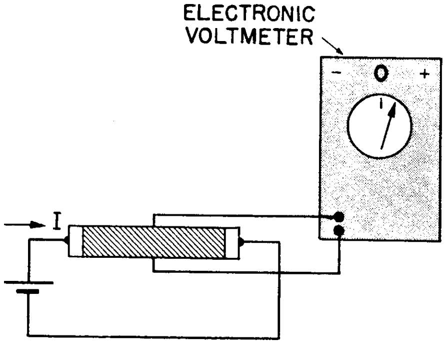

Fig. 14-7. Measuring the Hall effect.

The original discovery of the anomalous sign of the potential difference in the Hall effect was made in a metal rather than a semiconductor. It had been assumed that in metals the conduction was always by electron; however, it was found out that for berylium the potential difference had the wrong sign. It is now understood that in metals as well as in semiconductors it is possible, in certain circumstances, that the "objects" responsible for the conduction are holes. Although it is ultimately the electrons in the crystal which do the moving, nevertheless, the relationship of the momentum and the energy, and the response to external fields is exactly what one would expect for an electric current carried by positive particles.

Let's see if we can make a quantitative estimate of the magnitude of the voltage difference expected from the Hall effect. If the voltmeter in Fig. 14-7 draws a negligible current, then the charges inside the semiconductor must be moving from left to right and the vertical magnetic force must be precisely cancelled by a vertical electric field which we will call $\mathcal{E}_{tr}$ (the "tr" is for "transverse"). If this electric field is to cancel the magnetic forces, we must have

$$\mathcal{E}_{tr} = -v_{drift} \times B. \tag{14.9}$$

Using the relation between the drift velocity and the electric current density given in Eq. (14.6), we get

$$\mathcal{E}_{\text{tr}} = -\frac{1}{qN}\,jB.$$

The potential difference between the top and the bottom of the crystal is, of course, this electric field strength multiplied by the height of the crystal. The electric field strength $\mathcal{E}_{\text{tr}}$ in the crystal is proportional to the current density and to the magnetic field strength. The constant of proportionality $1/qN$ is called the Hall coefficient and is usually represented by the symbol R_H. The Hall coefficient depends just on the density of carriers—provided that carriers of one sign are in a large majority. Measurement of the Hall effect is, therefore, one convenient way of determining experimentally the density of carriers in a semiconductor.

14–4 Semiconductor junctions

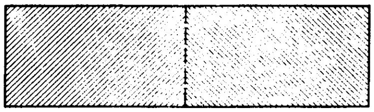

Fig. 14–8. A *p-n* junction.

We would like to discuss now what happens if we take two pieces of germanium or silicon with different internal characteristics—say different kinds or amounts of doping—and put them together to make a "junction." Let's start out with what is called a *p-n* junction in which we have *p*-type germanium on one side of the boundary and *n*-type germanium on the other side of the boundary—as sketched in Fig. 14–8. Actually, it is not practical to put together two separate pieces of crystal and have them in uniform contact on an atomic scale. Instead, junctions are made out of a single crystal which has been modified in the two separate regions. One way is to add some suitable doping impurity to the "melt" after only half of the crystal has grown. Another way is to paint a little of the impurity element on the surface and then heat the crystal causing some impurity atoms to diffuse into the body of the crystal. Junctions made in these ways do not have a sharp boundary, although the boundaries can be made as thin as 10^{-4} centimeters or so. For our discussions we will imagine an ideal situation in which these two regions of the crystal with different properties meeting at a sharp boundary.

On the *n*-type side of *p-n* junction there are free electrons which can move about, as well as the fixed donor sites which balance the overall electric charge. On the *p*-type side there are free holes moving about and an equal number of negative acceptor sites keeping the charge balanced. Actually, that describes the situation before we put the two materials in contact. Once they are connected together the situation will change near the boundary. When the electrons in the *n*-type material arrive at the boundary they will not be reflected back as they would at a free surface, but are able to go right on into the *p*-type material. Some of the electrons of the *n*-type material will, therefore, tend to diffuse over into the *p*-type material where there are fewer electrons. This cannot go on forever because as we lose electrons from the *n*-side the net positive charge there increases until finally an electric voltage is built up which retards the diffusion of electrons into the *p*-side. In a similar way, the positive carriers of the *p*-type material can diffuse across the junction into the *n*-type material. When they do this they leave behind an excess of negative charge. Under equilibrium conditions the net diffusion current must be zero. This brought about by the electric fields which are established in such a way as to draw the positive carriers back toward the *p*-type material.

The two diffusion processes we have been describing go on simultaneously and, you will notice, both act in the direction which will charge up the *n*-type material in a positive sense and the *p*-type material in a negative sense. Because of the finite conductivity of the semiconductor material, the change in potential from the *p*-side to the *n*-side will occur in a relatively narrow region near the boundary; the main body of each block of material will have a uniform potential. Let's imagine an *x*-axis in a direction perpendicular to the boundary surface. Then the electric potential will vary with *x*, as shown in Fig. 14–9(b). We have also shown in part (c) of the figure the expected variation of the density N_n of *n*-carriers and the density N_p of *p*-carriers. Far away from the junction the carrier densities N_p and N_n should be just the equilibrium density we would expect for individual blocks of materials at the same temperature. (We have drawn the figure for a

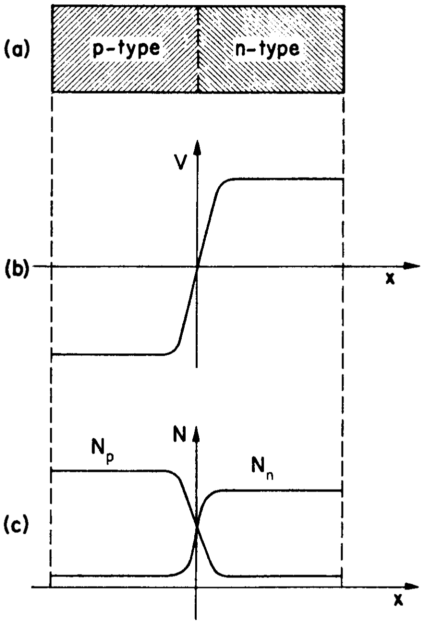

Fig. 14–9. The electric potential and the carrier densities in an unbiased semiconductor junction.