Holes in Hall Effect

Lianxi Ma,[1,*] Qingli Zhao[2], Chi Chen[1]

[1]Department of Physics, Texas A&M University, College Station, TX 77843, USA.

[*]Present address: Department of Physics, Blinn College, Bryan, TX 77805, USA.

[2]Department of Automation, Tangshan University, Hebei, 063000, P.R. China.

E-mail: lianxi.ma@blinn.edu

(Received 12 December; accepted 5 January 2009)

Abstract

Hall Effect can be used to determine the signs of current carriers in metals and semiconductors. It is well known that when electrons are current carriers, the Hall coefficient is negative, i.e. $R_H<0$; when holes are current carriers, the Hall coefficient is positive, i.e. $R_H>0$. However, puzzling arises regarding that in both scenarios; the essential moving particles are electrons. Therefore, there should not have any different effect in theory. We discuss the details about two situations and point out that both quantum and classical mechanics give same current direction under external electric field. However, under the influence of external magnetic field, because the mass of electrons is negative at valence band, electrons move to the opposite direction of its Lorentz force, which behave like a positive charge and give positive R_H.

Keywords: Hall effect, Hall coefficient, Holes, Effective mass.

Resumen

El Efecto Hall se puede utilizar para determinar los signos de los portadores de corriente en metales y semiconductores. Es bien sabido que cuando los electrones son portadores de corriente, el coeficiente de Hall es negativo, es decir, R_H <0, cuando los agujeros son portadores de corriente, el coeficiente de Hall es positivo, es decir, $R_H > 0$. Sin embargo, el desconcierto surge al considerar que en ambos escenarios; las partículas que esencialmente se mueven son electrones. Por lo tanto, no debería tener ningún efecto diferente en la teoría. Se discuten los detalles de las dos situaciones y se señala que tanto la mecánica clásica y la cuántica dan la misma dirección de la corriente bajo el campo eléctrico externo. Sin embargo, bajo la influencia del campo magnético externo, ya que la masa de los electrones es negativa en la banda de valencia, los electrones se mueven hacia la dirección opuesta de su fuerza de Lorentz, que se comporta como una carga positiva y da R_H positivo.

Palabras clave: Efecto Hall, coeficiente de Hall, Agujeros, masa efectiva.

PACS: 07.79.Pk, 07.79.Lh, 03.50.De **ISSN 1870-9095**

I. INTRODUCTION

Hall Effect can be used to determine the signs of current carriers in metals and semiconductors. Hall coefficient is defined as:

$$R_H=E_y/(j_xH), \tag{1}$$

where E_y, j_x, *and* H are electric field, electric current density, and magnetic field strength (shown in Fig. 1). If $R_H<0$, then it indicates that the E_y is along $-y$ direction, and electric current carriers are negative particles (essentially they are electrons); if $R_H>0$, however, it indicates that E_y is along y direction, and electric current carriers are positive particles, which are called holes

By applying Drude model [1], we can write R_H as

$$R_H=-\frac{1}{nec}, \tag{2}$$

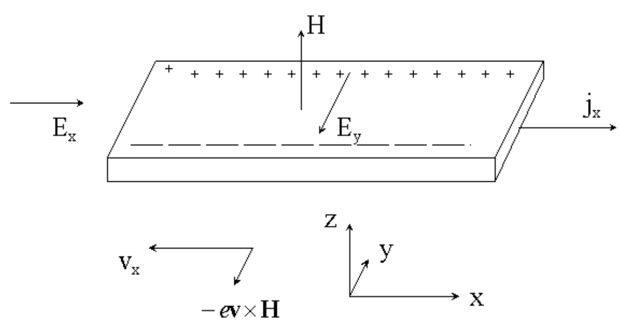

FIGURE 1. Schematic of Hall effect experiment. Usually the electric carriers are believed to be electrons. The electrons move against the external electric field and therefore, Lorentz force make them move toward downside. The resultant electric field is along –y direction resulting in negative Hall coefficient.

where n, e, c are electric charge density, electron charge (positive here, as negative sign has been taken care of), and speed of light (the cgs system is used for convenience). The other version of Eq. (2) is $R_Hnec=-1$.

Suppose that those values are measured in the experiment, their multiplication can be compared with -1. In Table I we show some experimental values.

TABLE I. Experimental values of -1/($R_H$nec) of some elements.

Metal	Number of valence electrons	-1/($R_H nec$)
Li	1	0.8
Na	1	1.2
Be	2	-0.2
Mg	2	-0.4
In	3	-0.3
Al	3	-0.3

Table I shows that $R_H nec$ is not -1. This is not a big surprise since in Drude Model we regard electrons in the crystal as free electrons. This is not true due to the interactions between electrons and atoms. However, the most surprising fact is that some metals such as Be, Mg, In, and Al have positive $R_H nec$ values. Positive $R_H nec$ values indicate that electric field E_y is along the y direction, and electric carriers are positive. General explanation for such a phenomenon is that the electric carriers are holes rather than electrons, and the holes can be regarded as positive particles. The second thought, however, will raise a question: hole itself doesn't move; under the influence of the external electromagnetic field, it is electrons that move instead of holes. Therefore, Hall coefficient should always be negative. How do we explain this paradox?

It is *impossible* to reconcile this problem in the frame of classical physics because of the wave-particle dual properties of the electrons. This is just like the solar system model of an atom proposed by Ernest Rutherford couldn't explain why the electrons don't radiate energy while moving around a nucleus. For the electrons in a crystal, we need to apply quantum mechanics to the electrons and consider the influence of the ions. Due to the huge number of ions, it is hard to get an analytical solution about electrons. Therefore, several approximation methods are proposed. The approach that "Nearly free electrons" is one of them. In this approach, mean field is used to represent the real field and the period electric field induced by ions is regarded as perturbation. Under these assumptions, Schrödinger equation yields approximations of electron energies and positions, and thereafter, the concept of energy band is proposed [2].

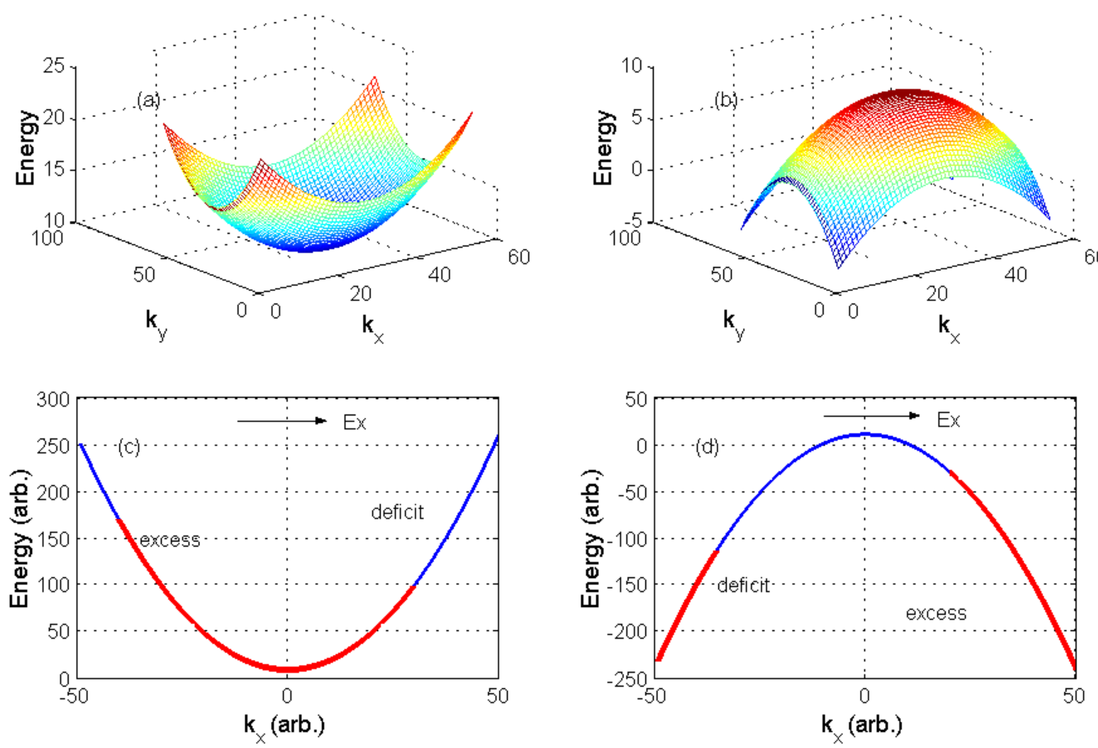

FIGURE 2. Considering the electrons are moving in 2D **k** – space. Fig. 2(a), (b) represent energy of conduction band and valence band. In the (c) and (d), the condition that k_y=0 is considered. If there is no magnetic field and the electrons are moving only under the effect of electric field, Eq. (5) leads to the conclusion that wave vector **k** decreases, i. e. the average of **k** moves to the left. Fig. 2(c) shows that under the influence of electric field, there are more electrons in the region of k_x<0, which results in negative average value of **k.** Red line stands for the positions occupied by electrons. Fig. 2(d)) shows that under the influence of electric field, there are more electrons in the region of k_x>0, which results in positive average value of **k**. Red line stands for the positions occupied by electrons.

One important conclusion of "nearly free electron approximation" is that the electron's energy can be expressed as

$$\varepsilon = \varepsilon_0 \pm A(\mathbf{k} - \mathbf{k_0})^2, \qquad (3)$$

where ε_0 is potential energy which is a constant; $A(\mathbf{k}\text{-}\mathbf{k}_0)$ is kinetic energy, and we have $A>0$; **k** is wave vector which can be an arbitrary value and it determines electron's kinetic energy. Therefore the electron's energy can still be regarded as being composed of potential and kinetic energies. We usually defines a quantity with mass dimension, m^*, with the relation $A = \frac{\hbar^2}{2|m^*|}$. Here $\hbar = \frac{h}{2\pi} = 1.05\times10^{-27}\,\text{erg}\cdot\text{s}$, is Planck constant. (The reason that we use m^*, instead of m, to stand for mass will be mentioned later). Hence, electron's energy has two possibilities: one is parabolic surface convex to the bottom (take "+" in Eq. (3), shown in Fig. 2(a)), and the other is parabolic surface concave to the bottom (take "-" in Eq. (3), shown in Fig. 2(b)).

Another conclusion of energy band theory is that electron's velocity, v, can be expressed as

$$v = \frac{1}{h}\frac{d\varepsilon}{dk}. \qquad (4)$$

And in the electro-magnetic field, the electron's equation is

$$\hbar\frac{d\mathbf{k}}{dt} = (-e)(\mathbf{E} + \frac{\mathbf{v}\times\mathbf{H}}{c}) = (-e)\mathbf{\Pi}. \qquad (5)$$

Here $\mathbf{\Pi} = \mathbf{E} + \frac{\mathbf{v}\times\mathbf{H}}{c}$, is the sum of electric and magnetic field. The left hand side of the Eq. (5) is momentum change rate, and the right hand side is the sum of the electric and magnetic forces. Therefore, it is still in the form of Newton's second law of motion: **F** = d**p**/dt. However, readers may have noticed that the momentum is $\hbar\mathbf{k}$ instead of classical form of $m\mathbf{v}$. Equation (5) shows that the sum of electric and magnetic forces, $(-e)\mathbf{\Pi}$, induces the change of wave vector **k**, rather than the change of velocity **v**. This is the essential difference from the Newton's second law $\mathbf{F} = m\mathbf{a}$. The fact that Eqs (3), (4), and (5) all include Planck constant indicates that under the frame of classical physics, it is impossible to get those equations derived.

Armed with these 3 equations, we are able to consider Hall effect now. Equation (3) shows that energy ε is symmetric to wave vector **k**, and therefore Eq. (4) shows that electron's velocity can be either positive or negative. Because the number distribution of electrons with **k** is symmetric, the numbers of electrons with positive and negative velocity are equal – no electric current is produced. Interestingly, in the case that there is no external electric field, electron's velocity is not zero and actually, oscillation is formed. After the electric field **E** is introduced, Eq. (5) shows that **k** decreases (d**k**/dt is negative). In the case that energy band is like Fig. 2(a), which means that it is not filled full and called conduction band, then there are more electrons in the –**k** region. In another word, there are more electrons with negative velocity（velocity direction is opposite to the electric field **E**), and hence electric current with the same direction of external electric field is produced. This can be said that electric carriers are electrons, which is displayed in Fig. 2(c). In the case that energy band is like Fig. 2(b), which means that it is filled full and called valence band, then **k** cannot be changed anymore. Or strictly speaking, electrons going out of -**k** border come back from +**k** border [3], and so electron distribution does not change. Resultant current is zero.

In the case that there are vacant places in the valence band, we call those vacant places holes. Since the electron's velocity is still symmetric to the k, the overall electric current is zero. When electric field **E** is introduced, Eq. (5) shows that **k** decreases (d**k**/dt is negative). But electrons going out of -**k** border come back from +**k** border [3], and so there are more electrons with +**k,** which means that there are more electrons with negative velocity (opposite to the electric field **E**). Electric current in the direction of electric field is formed as shown in Fig. 2(d). This can be said that the electric carriers are holes. Yes, under the influence of the external electric field, it is the electrons, rather than holes, that move to form current. Both produce the electric current with the same direction as the external electric field. This conclusion agrees with that of classical physics.

Readers may wonder why we used quantum mechanics rather than classical mechanics to get the essentially the same result. This is because under the frame of the classical mechanics, the electric current should appear as long as the electrons in the solid have collective motion under the influence of external electric field. But this is in contradiction to the experiment. Experimentally we have found some solids that are insulators but there are a lot of electrons in them. Quantum theory gives the result that electrons may go opposite direction and the current may cancel out. (In fact this is because of the interaction between electrons and crystal lattice.) Explaining this phenomenon is one of the great successes of energy band theory.

In the following we discuss the result derived from quantum theory, that is different from the one derived from classical mechanics.

Now we consider a magnetic field **H** that is perpendicular to the electric field and along the z direction is introduced. We still assume that electrons' motion can be described by Newton's law $\mathbf{F} = m\mathbf{a}$. For the electrons in conduction band – or non-strictly speaking, free electrons – Lorentz force and acceleration directions are straightforward: along the –y direction. Electrons move toward the –y direction and thereafter produce the electric field $-E_y$ that pointing toward $-y$, and Hall coefficient

$R_H = E_y/(j_x H)$ is negative. This can be predicted by classical mechanics too. However, what is the Lorentz force for the electrons in valence band – or non-strictly speaking, bounded electrons? Yes, the right hand rule is still right: Lorentz force is toward the –*y* direction. But electrons don't accelerate along the direction of the Lorentz force. Rather, they accelerate toward the direction opposite to the Lorentz force! This induces the electric field $-E_y$ along the *y* direction and the Hall coefficient $R_H = E_y/(j_x H)$ is positive. Why? Because the electron's mass is negative now!

Considering the motion of electrons in a crystal lattice and if we still treat the electrons as classical particles, the following equation can be used:

$$\mathbf{F} = m^* \mathbf{a}, \qquad (6)$$

which is very similar to Newton's second law of motion. Here the mass m^* is called effective mass which is not a constant. Rather it is the function of wave vector **k**. Strictly speaking it is a tensor. Therefore generally speaking the directions of external force **F** and the acceleration **a** are different. The definition of effective mass m^* is

$$\frac{1}{m^*} = \frac{1}{\hbar} \frac{\partial^2 \varepsilon}{\partial k^2}. \qquad (7)$$

For the Eq. (3), electron's mass has two possibilities: one is positive, $m^* = \frac{\hbar}{2A}$, which corresponds to convex parabolic surface (Fig. 2(a)) and the acceleration direction of the electrons in the conduction band is the same as the external force; the other is negative, $m^* = -\frac{\hbar}{2A}$, which corresponds to concave parabolic surface (Fig. 2(b)) and the acceleration direction of the electrons in the valence band is opposite to the external force. Holes are formed in the concave surface, and so the mass of the electrons is negative and the electrons accelerate opposite to the magnetic force.

The negative mass means that when electrons are moving, the momentum obtained from electric field is less than that transferred to the lattice. Their resultant momentum decreases. In fact, Eq. (6) includes the effect of interaction between electrons and crystal lattice.

In summary, we can see that quantum mechanics and classical mechanics give the same direction of the electric current, but give the different acceleration direction in the magnetic field in for the valence band.

Therefore, the Hall Effect with the holes as electric carriers is explained.

Actually, it has been rigorously proved [4] that the collective motion of the electrons in the valence band when there are holes is equivalent to the motion of positive change *e* of the holes. And the mathematical description for the latter is much easier and the physical picture is clearer. Therefore, we usually say that hole has positive charge *e*.

REFERENCES

[1] Marder, M. P., *Condensed Matter Physics*, (Corrected Pinting, pp. 413, John Wiley & Sons, Inc., USA, 2000).
[2] Kun, H., *Solid state physics*, (pp. 124, People's publishing bureau. Jan., China, 1979). In Chinese.
[3] Kun, H., *Solid state physics*, (pp. 165, People's publishing bureau. Jan., China, 1979). In Chinese.
[4] Kittel, C., Introduction to Solid State Physics, (7^{th} edition, John Willey & Sons, Inc., USA, p. 206, 1995).