Cambridge University Press & Assessment
978-1-316-51175-6 — Quantum Hall Effect
Saurabh Basu
Excerpt
More Information

1

Quantum Hall Effect

1.1 Introduction

The date of discovery of the quantum Hall effect (QHE) is known pretty accurately. It occurred at 2:00 a.m. on 5 February 1980 at the high magnetic lab in Grenoble, France (see Fig. 1.1). There was an ongoing research on the transport properties of silicon field-effect transistors (FETs). The main motive was to improve the mobility of these FET devices. The devices that were provided by Dorda and Pepper allowed direct measurement of the resistivity tensor. The system is a highly degenerate two-dimensional (2D) electron gas contained in the inversion layer of a metal oxide semiconductor field effect transistor (MOSFET) operated at low temperatures and strong magnetic fields. The original notes appear in Fig. 1.1, where it is clearly stated that the Hall resistivity involves universal constants and hence signals towards the involvement of a very fundamental phenomenon.

In the classical version of the phenomenon discovered by E. Hall in 1879, just over a hundred years before the discovery of its quantum analogue, one may consider a sample with a planar geometry so as to restrict the carriers to move in a 2D plane. Next, turn on a bias voltage so that a current flows in one of the longitudinal directions and a strong magnetic field perpendicular to the plane of the gas (see Fig. 1.2). Because of the Lorentz force, the carriers drift towards a direction transverse to the direction of the current flowing in the sample. At equilibrium, a voltage develops in the transverse direction, which is known as the Hall voltage. The Hall resistivity, R, defined as the Hall voltage divided by the longitudinal current, is found to linearly depend on the magnetic field, B, and inversely on the carrier density, n, through $R = \frac{B}{nq}$ (q is the charge). A related

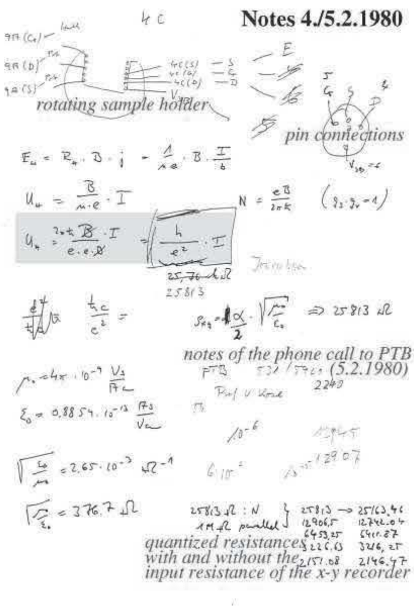

FIGURE 1.1. Copy of the original notes by K.v. Klitzing on the discovery of the quantum Hall effect. It documents that the Hall resistance ($\frac{U_H}{I}$) involves a fundamental constant h/e^2 (taken from Ref. [60]).

and possibly more familiar quantity is the Hall coefficient, denoted by $R_H = R/B$, which via its sign yields information on the type of the majority carriers, that is, whether they are electrons or holes.

At very low temperature or at very high values of the magnetic field (or at both), the resistivity of the sample assumes quantized values of the form $\rho_{xy} = \frac{h}{ne^2}$. Initially, n was found to be an integer with extraordinary precession (one part in $\sim 10^8$). This is shown in Fig. 1.3. The quantization of the Hall resistivity yields the name *quantum* (or quantized) Hall effect, which we refer to as QHE throughout the book.

Klaus von Klitzing and his co-workers [60, 61], while measuring the electrical transport properties of planar systems formed at the interface of two different semiconducting samples in the strong magnetic field facility at Grenoble, France, noted that the Hall resistivity is quantized in units of h/e^2 as a function of the external magnetic field. The flatness of the plateaus occurring at integer or fractional values of h/e^2 has an unprecedented precession and is independent of the geometry of the sample (as long as it is 2D), density of the charge carriers and its purity. The accuracy of the quantization aids in fixing the unit of resistance, namely $h/e^2 = 25.813$ kΩ, also called the Klitzing constant. Thus, among other significant properties of QHE that we shall be discussing in due course, an experiment performed at a macroscopic scale that can be used for metrology

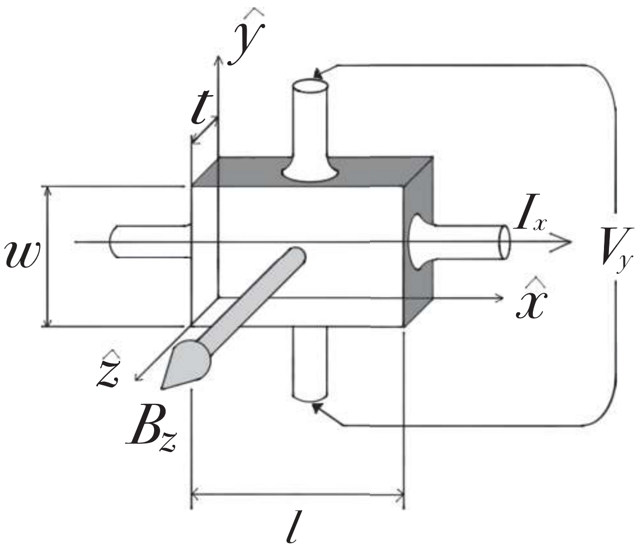

FIGURE 1.2. Typical Hall experiment set up showing the direction of the current, I_x, and the magnetic field, B_z. V_y denotes the Hall voltage.

or yields the values for the fundamental constants used in quantum physics is truly amazing and hence calls for an intense scrutiny. The effect occurs when the density of the carriers, n, is such that they are encoded in the integers (that come as proportionality constants to the Hall resistivity in terms of h/e^2) as if the charges locked their separation at some particular values. The phenomenon remains resilient to changing the carrier density by a small amount; however, changing it by a large amount does destroy the effect.

The Hall resistivity (upper curve) in Fig. 1.3 becomes constant for certain ranges of the external magnetic fields, which are called plateaus. Further, the longitudinal resistivity (lower curve) in the same plot vanishes everywhere. Although it shows peaks wherever there is a jump in the Hall resistivity from one plateau to another. Later on, it was observed that ν may take values that are rational fractions, such as $\nu = \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{3}{7}, \frac{4}{9}, \frac{5}{9}$. There are about 100 fractions (including the improper ones) that have been noted in experiments so far. The corresponding plot appears in Fig. 1.4.

1.2 General perspectives

The charge carriers being confined in 2D wells have a longer history. Since 1966 it has been known that the electrons accumulated at the surface of a silicon single

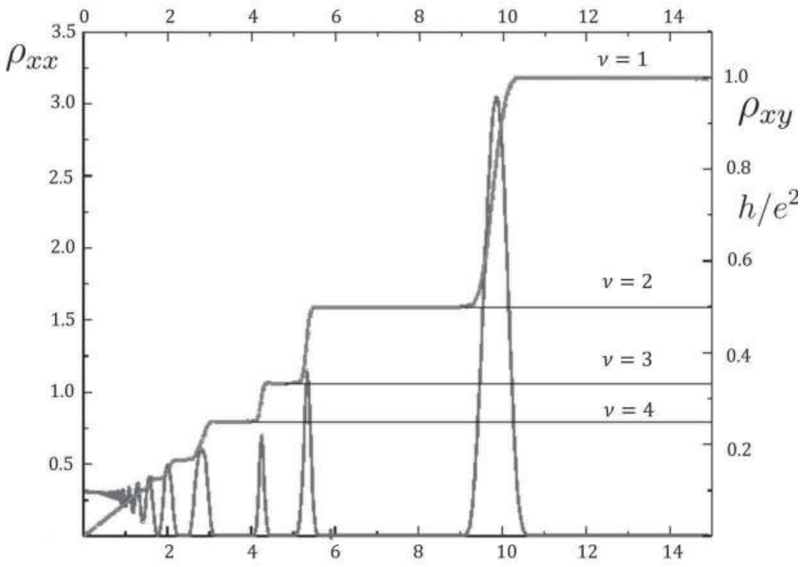

FIGURE 1.3. Schematic plot of integer quantum Hall effect (IQHE) as a function of the applied magnetic field. The plateaus in the upper curve denote quantization of Hall resistivity (ρ_{xy}), while the lower curve with spikes denote the magnetoresistivity (ρ_{xx}).

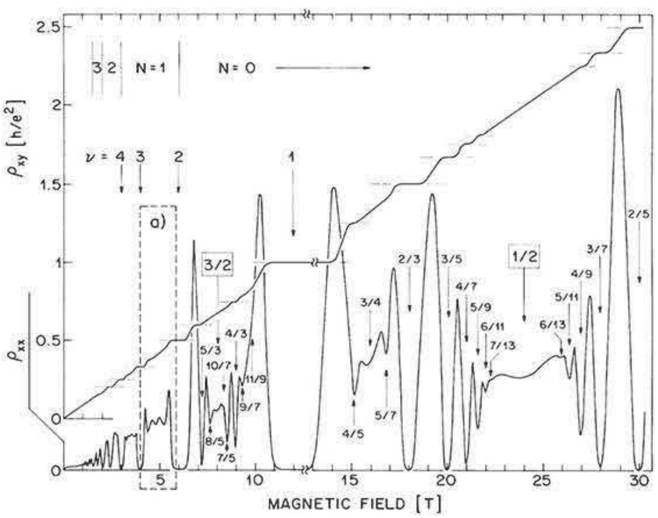

FIGURE 1.4. The plot shows fractional quantum Hall effect (FQHE). The plateaus are shown at fractional values in units of h/e^2. Taken from Ref.[62].

crystal induced by a positive gate voltage form a 2D electron gas (2DEG). The energy of the electrons that moves perpendicular to the surface is quantized (box quantization), and on top of it, the free motion of the electrons in 2D becomes quantized when a strong magnetic field is applied perpendicular to the plane

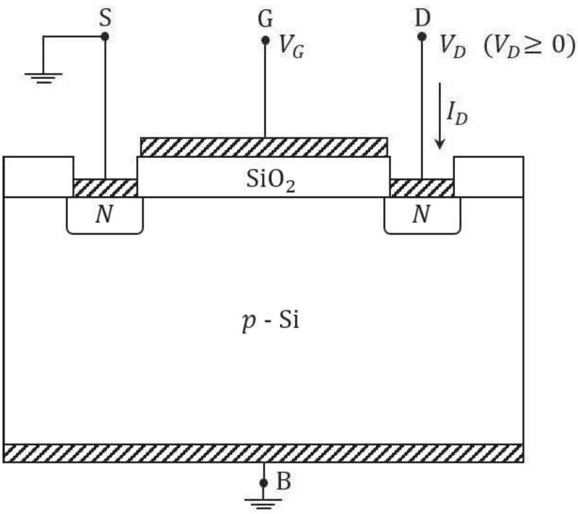

FIGURE 1.5. MOSFET structure showing base (B), gate (G), source (S) and drain (D). The substrate is a p-type Si. SiO_2 denotes the insulating oxide layer.

(Landau quantization). Thus, QHE has both the quantization phenomena built into it.

An important recent development in the study of semiconductors is the achievement of structures in which the electrons are restricted to move essentially in 2D. This immediately indicates that the carriers are prohibited to move along the direction transverse to the plane. Hence, the motion is quantized. Such 2D behaviour of the carriers can be found in metal-oxide-semiconductor (MOS) structures, quantum wells and superlattices. An excellent prototype is the metal-insulator-semiconductor (MIS) layered structure of which the insulator is usually an oxide, such as Al_2O_3 (thereby making it an MOS structure). In Fig. 1.5, we show a typical MOS device where the substrate is a doped p-type silicon (Si), which is grounded and is called a base (shown by B in Fig. 1.5). On the top, there is a metallic layer (shown by the hatched regime), followed by an insulating layer formed by SiO_2. The metallic layer is called the gate, denoted by G, which is biased by a voltage V_G. The source (grounded) and the drains are denoted by S and D, respectively, which in Fig. 1.5 consist of n-type materials. The gate voltage causes the carriers beneath the gate electrode to drift between the source and the drain. The layer of charge carriers below the oxide layer forms the 2DEG, which is central to our discussion. The energy dispersion in this case reads

$$E_n(k_x, k_y) = \frac{\hbar^2 k_x^2}{2m_{xx}^*} + \frac{\hbar^2 k_y^2}{2m_{yy}^*}, \tag{1.2.1}$$

Cambridge University Press & Assessment
978-1-316-51175-6 — Quantum Hall Effect
Saurabh Basu
Excerpt
More Information

where m^*_{xx} and m^*_{yy} are the components of the effective mass tensor defined by the inverse of the curvature of the band structure, namely

$$m^*_{\alpha\beta} = \hbar^2 \left(\frac{\partial^2 E(k)}{\partial k_\alpha \partial k_\beta} \right)^{-1}. \tag{1.2.2}$$

The physical properties of all systems are governed by their density of states (DOS), which plays a crucial role in deciding on the dependence of the temperature, density of carriers, etc. In 2D systems with a parabolic dispersion, as elaborated above in Eq. (1.2.1), the DOS is a constant and assumes a form

$$g(E) = g_{2D} = \frac{m^*}{\pi\hbar^2}. \tag{1.2.3}$$

The energy-independent DOS is very special to 2D and is in sharp contrast to three dimensions (3D), where it goes as $E^{1/2}$, and to one dimension (1D), where it goes as $E^{-1/2}$. In a general sense, and not restricted to the discussion on the Hall effect, the DOS enters while calculating the average quantities, such as the average energy or the average number of particles. For example, the average of a physical observable, O, of a fermionic system is computed using

$$\langle O \rangle = \int_0^\mu O f(E) g(E) dE, \tag{1.2.4}$$

where $f(E)$ is the Fermi distribution function given by

$$f(E) = \frac{1}{e^{\beta(E-\mu)} + 1},$$

with $\beta = \frac{1}{k_B T}$ and μ denoting the chemical potential. In general, this integral is quite challenging to compute analytically because of the Fermi distribution function ($f(E)$) present in the integrand.

Meanwhile, there is a wonderful simplification where $f(E)$ assumes a value unity at all temperatures at which the experiments are performed. Only at temperatures close to the Fermi temperature, T_F, defined via $\epsilon_F = k_B T_F$ (ϵ_F being the Fermi energy), $f(E)$ starts deviating from unity, and its exact form needs to be incorporated in the integral. However, T_F is usually of the order of tens of thousands of Kelvin for typical metals (such as Cu and Al), which is too high for them to appear in experimental situations. Moreover, the DOS only depends on energy and is independent of the temperature to a very good approximation. Thus, computation of Eq. (1.2.4) becomes independent or weakly dependent on temperature [63,64].

1.3 Why is 2D important?

In the following discussion, we mention that there is something interesting about the transport properties of 2D systems. In the linear response regime Ohm's law is valid and indicates that

$$V_\alpha = R_{\alpha\beta} I_\beta, \tag{1.3.1A}$$

where $R_{\alpha\beta}$ denotes the resistivity tensor and α, β denote spatial variables x, y, etc. One can equivalently invert this equation to write $I_\alpha = G_{\alpha\beta} V_\beta$, where $G_{\alpha\beta}$ represents the conductivity tensor with $G = R^{-1}$. Equivalent relations in terms of the components of the electric field (**E**) and current density (**j**) read as

$$E_\alpha = \rho_{\alpha\beta} j_\beta \quad \text{and} \quad j_\alpha = \sigma_{\alpha\beta} E_\beta, \tag{1.3.1B}$$

where ρ and σ denote the resistivity and the conductivity tensors, respectively.

An interesting (and useful too) artefact of 2D physics is an accidental similarity that exists in decoding some of the key features of the transport properties. For example, the resistivity, ρ (or the conductivity, σ), is a quantity that is independent of the system geometry and hence is useful for a theoretical analysis. Whereas in experiments, one measures the resistance of a sample, R (or the conductance, G). For a sample in the shape of a hypercube of sides having length, L, the resistance and the resistivity are related by

$$R = \rho L^{2-d}, \tag{1.3.2}$$

where d denotes the dimensionality. Only for $d = 2$ is the resistance a scale invariant quantity. This puts the experimentalists and the theorists on the same page, as the geometry of the sample does not enter explicitly in the analysis of its transport properties.

1.4 Why are the conductivity and the resistivity tensors antisymmetric?

As we dig more into the details of the transport properties of 2DEG in the presence of a magnetic field, further useful information emerges. The off-diagonal elements of both the conductivity and the resistivity tensors are antisymmetric with regard to the direction of the applied field, **B**. Consider a planar sample with dimensions

$L_x \times L_y$. The conductivity tensor is of the form

$$\sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{pmatrix}. \tag{1.4.1}$$

Let us try to understand the nature of the tensor, σ, in the presence of a magnetic field. A conductor in an external magnetic field obeys

$$j_\alpha = \sigma_{\alpha\beta} E_\beta. \tag{1.4.2}$$

Onsager's reciprocity principle does not hold in the presence of a magnetic field, $\mathbf{B}$, which implies [65]

$$\sigma_{\alpha\beta}(\mathbf{B}) \neq \sigma_{\beta\alpha}(\mathbf{B}). \tag{1.4.3}$$

Instead, one has $\sigma_{\alpha\beta}(\mathbf{B}) = \sigma_{\beta\alpha}(-\mathbf{B})$ to make sure that the time reversal holds only if $\mathbf{B}$ changes sign. Let us write the conductivity tensor as a sum of a symmetric and an antisymmetric tensor (note that this is always possible for a rank 2 tensor). Thus,

$$\sigma_{\alpha\beta} = S_{\alpha\beta} + A_{\alpha\beta}, \tag{1.4.4}$$

where $\mathbf{S}$ and $\mathbf{A}$ are the symmetric and the antisymmetric tensors that obey the relations

$$\begin{aligned} S_{\alpha\beta}(\mathbf{B}) &= S_{\beta\alpha}(-\mathbf{B}) = S_{\alpha\beta}(-\mathbf{B}) \\ A_{\alpha\beta}(\mathbf{B}) &= A_{\beta\alpha}(-\mathbf{B}) = -A_{\alpha\beta}(-\mathbf{B}) \end{aligned} \tag{1.4.5}$$

such that the components of $S_{\alpha\beta}$ are even functions of $\mathbf{B}$, while those of $A_{\alpha\beta}$ are odd functions of $\mathbf{B}$. Putting in Eq. (1.4.2),

$$j_\alpha = S_{\alpha\beta} E_\beta + A_{\alpha\beta} E_\beta. \tag{1.4.6}$$

But owing to the antisymmetry,

$$A_{\alpha\beta} = \epsilon_{\gamma\alpha\beta} A_\gamma = -\epsilon_{\alpha\gamma\beta} A_\gamma. \tag{1.4.7}$$

Putting it in Eq. (1.4.6),

$$\begin{aligned} j_\alpha &= S_{\alpha\beta} E_\beta - \epsilon_{\alpha\beta\gamma} A_\beta E_\gamma \\ &= S_{\alpha\beta} E_\beta - (\mathbf{A} \times \mathbf{E})_\alpha \\ &= S_{\alpha\beta} E_\beta + (\mathbf{E} \times \mathbf{A})_\alpha. \end{aligned} \tag{1.4.8}$$

Assuming that we can expand $\sigma(\mathbf{B})$ in powers of $\mathbf{B}$,[1] such that the antisymmetric part contains odd powers of $\mathbf{B}$, we can write

$$A_\alpha = \eta_{\alpha\beta} B_\beta, \tag{1.4.9}$$

and $S_{\alpha\beta}(\mathbf{B})$ consists of even powers of $\mathbf{B}$,

$$S_\alpha = (\sigma_0)_{\alpha\beta} + \zeta_{\alpha\beta\gamma\delta} B_\gamma B_\delta. \tag{1.4.10}$$

The first term is the zero field conductivity tensor. Thus, putting things together up to terms linear in $\mathbf{B}$,

$$j_\alpha = S_{\alpha\beta} E_\beta + (\mathbf{E} \times \mathbf{A})_\alpha. \tag{1.4.11}$$

The second term denotes the Hall effect, which is linear in $\mathbf{B}$. This implies that the Hall current is perpendicular to the electric field, $\mathbf{E}$, and is proportional to $\mathbf{E}$ and $\mathbf{B}$. Thus, an antisymmetric tensor is relevant to the study of the Hall effect, which is why the conductivity and the resistivity tensors are antisymmetric. This is an important result that deviates from the corresponding scenario that arises in the absence of an external magnetic field.

1.5 Translationally invariant system: Classical limit of QHE

It is quite an irony that the extreme universal signature of the transport properties of a 2DEG is characterized by the flatness of the plateaus that not only occurs but survives even in the presence of disorder, impurity and imperfection. In the absence of the magnetic field, Anderson localization would have governed the transport signatures of non-interacting electrons, which indicates that in any dimension less than three, all eigenstates of a system are exponentially localized even for an infinitesimal disorder strength. Only in 3D is there a critical disorder at which a metal-insulation transition occurs. However, the scenario is strongly altered by the presence of the magnetic field, which yields, as we shall shortly see, a series of unique phase transitions from a perfect conductor to a perfect insulator. No other system demonstrates re-occurrence of the same phases over and over again as the magnetic field is gradually ramped up.

[1]It should be valid for weak magnetic fields and is not exactly true for the quantum Hall effect, but nevertheless it serves our purpose.

To begin with, we shall consider the case which is free from disorder, or equivalently a translationally (Lorentz) invariant system that possesses no preferred frame of reference. Thus, we can think of a reference frame that is moving with a velocity $-\mathbf{v}$ with respect to the lab frame, where the current density is given by $\mathbf{j} = -ne\mathbf{v}$ (n: areal electron density, $-e$: electronic charge). In this frame, the electric and the magnetic fields are given by[2]

$$\mathbf{E} = -\mathbf{v} \times \mathbf{B} \quad \text{and} \quad \mathbf{B} = B\hat{z}. \tag{1.5.1}$$

The above transformation ensures that an electric field must exist to balance the Lorentz force $-e\mathbf{v} \times \mathbf{B}$ in order to conduct without deflection. For the electric field, this yields

$$\mathbf{E} = \frac{1}{ne}\mathbf{J} \times \mathbf{B}. \tag{1.5.2}$$

This is equivalent to the tensor equation

$$E^{\mu} = \rho_{\mu\nu} j^{\nu}, \tag{1.5.3}$$

with the resistivity tensor given by

$$\rho_{\mu\nu} = \frac{B}{ne}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{1.5.4}$$

Inverting the tensor equation, one can obtain

$$j^{\mu} = \sigma_{\mu\nu} E^{\nu}, \tag{1.5.5}$$

where the conductivity tensor $\sigma_{\mu\nu}$ is defined via

$$\sigma_{\mu\nu} = \frac{ne}{B}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{1.5.6}$$

There is an interesting paradox that states that $\sigma_{xx} = \rho_{xx} = 0$ (see above), which is of course contradictory. However, we reserve this rather interesting topic for a discussion immediately afterwards. Here we wish to point out that we get the results for a classical Hall effect, that is, $\sigma_{xy} = \frac{ne}{B}$ (or $\rho = \frac{B}{ne}$). It is important to realize that the result is an artefact of Lorentz invariance, where the

[2] Remember that $\mathbf{E} = 0$ in the lab frame, though $\mathbf{B}$ remains unchanged.