5 Localization and Quantum Transport

A. MacKinnon

5.1 Introduction

Traditional solid-state physics is based on the concept of the perfect crystalline solid, sometimes with a relatively low density of defects. This perfect crystallinity has played a crucial role in the development of the subject, with Bloch's theorem providing the central conceptual base. Concepts that arise from this theorem, such as bands, Brillouin zones, vertical transitions, effective mass and heavy and light holes, are really only well-defined in a perfect infinite crystal. In the absence of crystallinity none of these concepts is strictly valid, though in some cases it provides a useful starting point. In general, however, a new approach is required to characterize electrons and phonons in disordered solids.

When we consider low-dimensional structures Bloch's theorem may or may not be valid. There is nothing intrinsic to low dimensionality which invalidates it. Many of the simple examples in quantum mechanics and solid-state physics textbooks are, in fact, one-dimensional (e.g. the particle in a box, the Kronig–Penney model). Indeed, in a quantum well prepared by any of the standard growth methods (Chapter 1), much of the physics can be understood by using first-year undergraduate quantum mechanics and the effective mass approximation (Chapter 2). This is because a region of adjacent GaAs layers in $Al_xGa_{1-x}As$ can, for many purposes, be regarded as a perfect potential well. By doping the AlGaAs, the electrons in the well can be spatially separated from the scattering due to the ionized donor atoms (Chapter 3). Thus, in many respects, the electrons in this system can be treated as particles in a one-dimensional box.

The quantum well is, however, a very special quasi-two-dimensional system, albeit a very important one. As discussed in Chapter 1, it is very difficult to prepare low-dimensional samples of high quality for other than lattice-matched planar heterostructures. Thus, most heterojunctions, such as those with a significant lattice mismatch, metal-oxide-semiconductor field-effect transistors (MOSFETs), narrow quantum wells and quantum wires, etc., are in practice highly disordered with an effective density of scatterers which can approach the density of atoms. Clearly, in such systems, it cannot be valid to treat the effect of scatterers with perturbation theory using the perfect crystalline case as a starting point.

As we shall see later, there is one sense in which low-dimensional systems are intrinsically different from three-dimensional systems. The amount of scattering

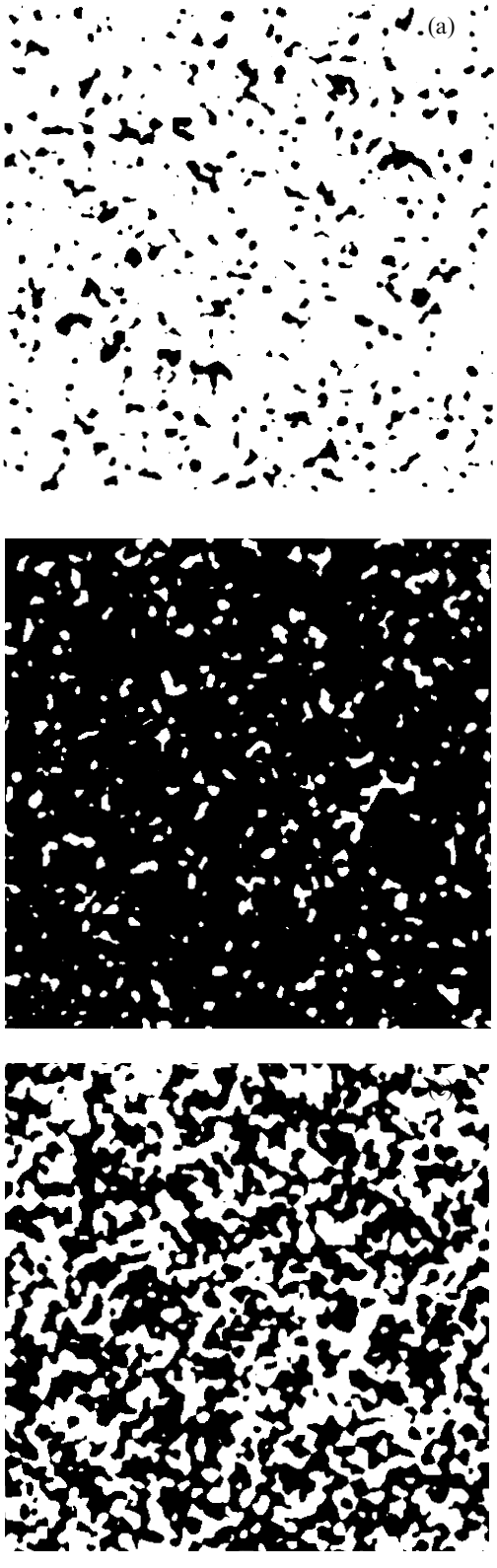

Figure 5.1. Percolation diagrams, with (a) low water level with a few lakes, (b) high water level with a few islands, (c) intermediate (critical) water level.

required to produce dramatic changes in the behaviour can sometimes be so small that perturbation theory may *never* be valid.

5.2 Localization

5.2.1 Percolation

Let us start with a simple classical problem. How does a fluid flow through a random medium? This is a problem of considerable practical importance in its own right: the extraction of oil from porous rock strata.

Consider a random landscape which is being slowly filled with water. At first there will be a continuous land mass with a few lakes (Fig. 5.1(a)). When the water level is very high we have islands in a sea (Fig. 5.1(b)). Let us now suppose there is a dam at the edge of the area which requires large quantities of water to drive a power station. When the water level is low only the lake next to the dam can be used and it will soon run out. As the level is raised this lake becomes larger but still finite. The power station will run longer but will still eventually drain the lake and have to stop. At a critical water level (Fig. 5.1(c)) the system changes from a lake district to an archipelago. This is analogous to the *percolation transition* (Stauffer and Aharony, 1994), where the water first forms a continuous network through the landscape. After this the power station can run indefinitely without fear of running out of water.

This phenomenon has much in common with more conventional phase transitions. There is a characteristic length scale which diverges at the transition: the size of the lakes or islands. There is a well-defined critical water level, rather like the critical temperature of the freezing transition or the ferromagnetic-to-paramagnetic transition in iron. If we think in terms of the density of blockages rather than the water level we see that there is a critical density above which the flow of water stops.

The one-dimensional version of this problem is special. Any blockage of the channel is enough to prevent the flow of water. The critical density is zero. This is an example of a problem which cannot be solved by perturbation theory. There is a discontinuous jump in the behaviour between a system with no blockages and one with a single blockage. In higher dimensions, in contrast, water can flow *around* the blockage.

5.2.2 The Anderson Transition and the Mobility Edge

The concept of the localization of electrons caused by disorder is due to Anderson (1958). He argued that an electron which starts at a particular site cannot completely diffuse away from that site if the disorder is greater than some critical value. Anderson thus introduced the concept of *localized* and *extended* states. The

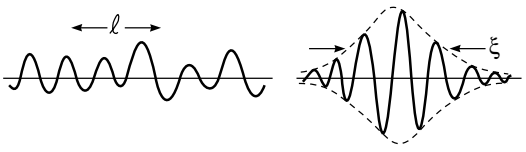

Figure 5.2. Schematic diagrams of (a) extended and (b) localized states, showing the correlation length, ℓ, and the localization length, ξ.

characteristics of these states can be summarized as follows (Fig. 5.2):

(a) extended	(i)	spread over the entire sample
	(ii)	not normalizable
	(iii)	contributes to transport
(b) localized	(i)	confined to a finite region
	(ii)	normalizable
	(iii)	does not contribute to transport

It is worth noting at this point that the phenomenon of localization is not confined to electrons, but can also be observed in other wave phenomena in random media, such as acoustic and optical waves (Section 4.2.2), as well as water waves (Fig. 5.3).

Mott (1968) later introduced the concept of a *mobility edge* (Fig. 5.4). He argued that it is meaningless to consider localized and extended states which are degenerate since any linear combination of a localized and an extended state must be extended. Thus, the concept of localization can only be meaningful if there are separate energy regions of localized and extended states, rather like bands and gaps. These regions are separated by a mobility edge. Mott further argued that the states close to a band edge are more likely to be localized than those in the middle of a band. Since the localized states do not take part in conduction, electrons in a disordered semiconductor must be activated to beyond the mobility edge rather than simply to the band edge to contribute to the conductivity. This activated process would be manifested in a conductivity σ of the form

$$\sigma = \sigma_0 \exp\left(-\frac{E_\mu - E_F}{k_B T}\right) \tag{5.1}$$

where E_μ and E_F are the mobility edge and Fermi energy, respectively. This form should reveal itself as the slope in an *Arrhenius plot* of the conductivity, i.e. a plot of $\ln\sigma$ vs. $1/T$:

$$\ln\sigma = \ln\sigma_0 - \frac{E_\mu - E_F}{k_B T} \tag{5.2}$$

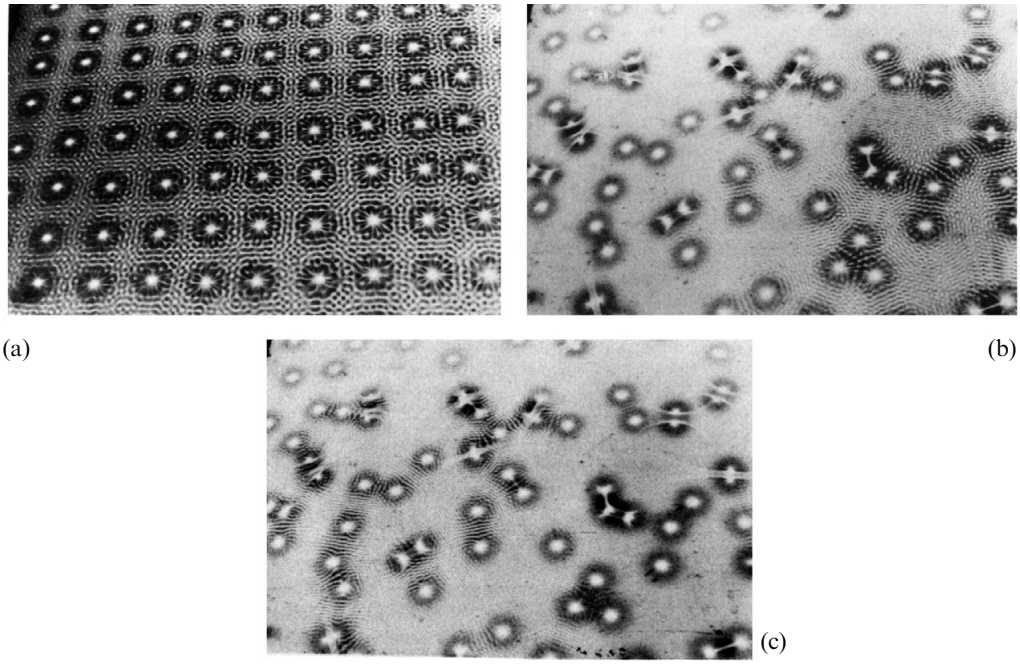

Figure 5.3. Three photographs of a water bath exposed to an audio-frequency oscillation (not stroboscopic). (a) shows a situation where the obstacles sit in a regular quadratic lattice (frequency 76 Hz). We see strong Bragg reflection corresponding to standing waves. (b) and (c) show randomly spaced obstacles exposed to two different audio frequencies (105 Hz and 76 Hz). Both (b) and (c) show standing wave patterns, but localized in different areas. (With the permission of the authors from Lindelof *et al.* (1986).)

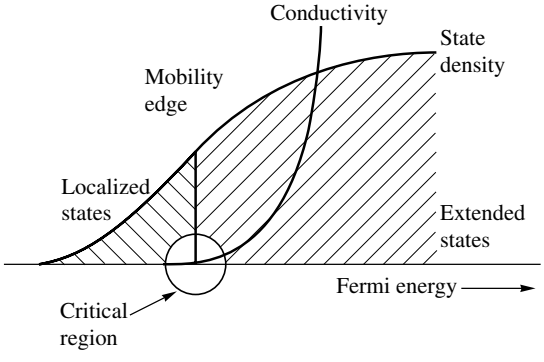

Figure 5.4. Schematic illustration of the density of states and the conductivity near a band edge in a disordered system. The regions of localized states (no conductivity) and extended states (finite conductivity) are indicated.

5.2.3 Variable Range Hopping

At very low temperatures, when activated conductivity is not significant and the Fermi level is in a region of localized states (e.g. an amorphous semiconductor), transport takes place by *hopping* between localized states. The electrons can gain or lose energy of order $k_B T$ from interactions with phonons and other excitations. In more than one dimension the electron is more likely to find a state in this energy range the further it hops.

On the other hand, the exponential envelopes of the localized states must overlap for the phonon to couple them. In this way Mott (Mott and Davis, 1979) found the famous $T^{-1/4}$ law. More precisely, Mott's law is written as

$$\sigma = \sigma_0 \exp\left[-\left(\frac{T_0}{T}\right)^{1/(d+1)}\right] \tag{5.3}$$

where d is the number of spatial dimensions. This result has been verified many times in different systems. Or has it? To make an accurate measurement of the exponent, $1/(d+1)$, the conductivity must be measured over several decades of temperature while still remaining below the onset of activated transport. Thus, the exponent cannot be determined very precisely. In addition, the measured value of the pre-exponential factor, σ_0, often disagrees with that obtained from theory by several orders of magnitude.

5.2.4 Minimum Metallic Conductivity

Yet another idea from Nevill Mott: the semi-classical conductivity can be written in the form

$$\sigma = \frac{ne^2\tau}{m} = \frac{ne^2\ell}{mv_F} = \frac{ne^2\ell}{\hbar k_F} \tag{5.4}$$

where n is the density of conduction electrons, m and e are the electron mass and charge, respectively, τ is a scattering time, ℓ is the mean free path, and v_F and k_F are the Fermi velocity and wave vector. The density n of electrons is proportional to k_F^d (Ashcroft and Mermin, 1976). The Ioffe–Regel (Ioffe and Regel, 1960) criterion states that the de Broglie wavelength of an electron cannot be greater than the mean free path ℓ (essentially, this is the Heisenberg uncertainty principle) and, in any case, neither can be less than the interatomic distance, a. Hence, the conductivity cannot be less than

$$\sigma_{\min} \propto \frac{e^2}{h} k_F^{d-1}\ell \geq \frac{e^2}{h} k_F^{d-2} \geq \frac{e^2}{h} a^{2-d} \tag{5.5}$$

In two dimensions, the material-dependent quantities k_F and a do not appear and $\sigma_{\min}$ may be a *universal* quantity: $\sigma_{\min} = e^2/h$. There have been many experiments which purported to measure a value for $\sigma_{\min}$ corresponding to 25 813 Ω in

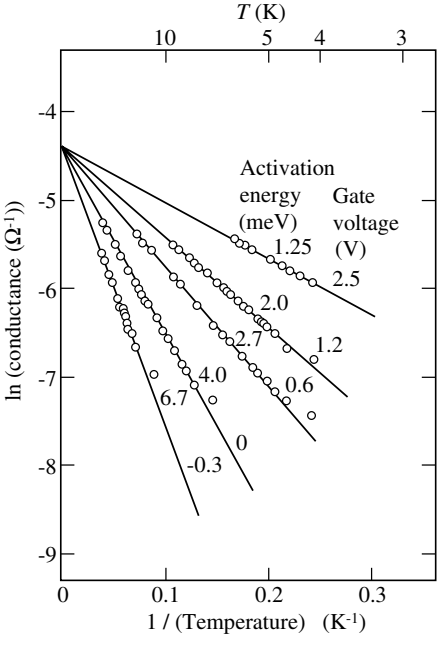

Figure 5.5. Arrhenius plot on a Silicon MOSFET, for various gate voltages V_g/V. Note the common intercept of the lines (Pepper, 1978). (Courtesy SUSSP Publications)

two-dimensional (2D) systems. By plotting equation (5.1) for several different systems there appears to be a common value of $\ln \sigma_0$ (see, e.g. Fig. 5.5). This is now believed to be characteristic of an intermediate regime and not a proof of the existence of $\sigma_{\min}$.

5.3 Scaling Theory and Quantum Interference

5.3.1 The Gang of Four

A decisive breakthrough in the theory of the metal-insulator transition was made when Abrahams, Anderson, Licciardello and Ramakrishnan (1979) – the 'gang of four' – published their scaling theory. The essence of their idea is that the transport properties of a disordered system can be expressed in terms of a single extensive variable, which can be chosen to be the conductance G (i.e. the inverse of the resistance, *not* the conductivity, which is the inverse of the resistivity). In particular, consider a block of disordered d-dimensional material. What is the change in the conductance when we join 2^d such blocks together to form a new block with all d of its dimensions doubled? The assumption (which can be justified) is that we can write

$$G(2L) = f(G(L)) \tag{5.6}$$

or, as a differential equation,

$$\frac{\mathrm{d}\ln g}{\mathrm{d}\ln L} = \beta(\ln g) \tag{5.7}$$

where $g = (\hbar/e^2)G$ is the dimensionless conductance, L is the length of an edge of a d-dimensional hypercube and β is some function of $\ln g$.

What are the properties of the function $\beta(\ln g)$? For strong disorder, the states are highly localized and we expect g to fall exponentially with the size of the system:

$$g = g_0 \exp(-\alpha L) \tag{5.8}$$

where α is an inverse decay length. Substituting this relation into (5.7) yields

$$\beta(\ln g) = \ln g - \ln g_0 \tag{5.9}$$

Thus, for $g \ll 1$, $\beta(\ln g)$ is always negative.

In the case of weak disorder the classical behaviour of the conductance (Exercise 7) should be valid:

$$g = \frac{\hbar}{e^2}\sigma L^{d-2} \tag{5.10}$$

and we obtain

$$\beta(\ln g) = d - 2 \tag{5.11}$$

Thus, for $g \gg 1$, $\beta(\ln g)$ is positive for $d = 3$, negative for $d = 1$ and zero for $d = 2$. Since g and β should be smooth functions of disorder, energy, etc., β must change sign for $d = 3$, may change sign for $d = 2$ and probably does not change sign for $d = 1$ (Fig. 5.6).

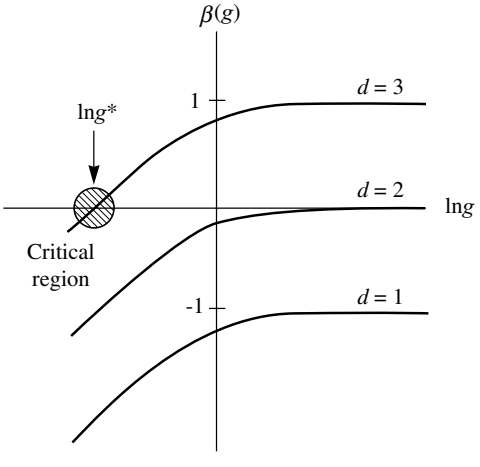

Figure 5.6. The β function (5.7) for the conductance g in 1, 2 and 3 dimensions.

What then is the meaning of the sign of β? If β is negative, then according to (5.7), g is decreasing with increasing L, so β becomes even more negative, and so on until g vanishes. That is, there is no conductance. When β is positive, by contrast, g is increasing with increasing L and will eventually approach the classical behaviour in (5.10) and (5.11). Thus, there is no metal-insulator transition for $d = 1$. All states are localized, just as in the percolation example discussed in Section 5.2.1. For $d = 3$, there is always a *metal-insulator transition*, which corresponds to the point $\beta = 0$.

What about $d = 2$? This is the marginal case. The existence of a transition depends on whether $\beta(\ln g)$ approaches zero from above or from below as $g \to \infty$. Abrahams *et al.* (1979) were able to show that the leading term in an expansion of (5.7) in $1/g$ has the form

$$\beta(\ln g) = -\frac{a}{g} \tag{5.12}$$

Thus, β is always negative and all states are localized. In other words, there is no true metallic conductivity in two dimensions.

While it is true in a strict mathematical sense that all states are localized in two dimensions, this result requires some interpretation. After all, there is no shortage of experimental evidence that there is considerable conductivity in some 2D systems (e.g. HEMTs, as discussed in Chapter 10). Moreover, there have been very few attempts to calculate the actual numerical value of the localization length for a real 2D system. Results of computer simulations (MacKinnon and Kramer, 1983a) suggest that it can be of the order of *centimetres*, even when the fluctuations in the potential are of the same order of magnitude as the band width. If we ask for the localization length in a very pure sample, then numbers larger than the length scale of the universe (e.g. $10^{10^{30}}$) tend to emerge (MacKinnon and Kramer, 1983b). Is it meaningful then to talk about localization in this case, when the localization length is often much larger than the sample size?

5.3.2 Experiments on Weak Localization

In practice, although localization often cannot be observed directly there are various precursor effects which can be observed fairly easily. These are collectively referred to as *weak localization*. By substituting (5.12) into (5.7) and integrating, we obtain a formula for the conductivity in two dimensions:

$$\sigma = \sigma_0 - a \ln L \tag{5.13}$$

This is still not of much use to us. The sample size is not an easily varied quantity. However, the discussion so far has been only in terms of disorder effects or elastic scattering. Inelastic effects such as scattering by phonons and by other electrons must also be considered. There is one important distinction between elastic and inelastic scattering. In elastic scattering there is a well-defined phase relationship between an incident and a scattered wave, whereas inelastic scattering destroys

such phase coherence. Since localization is really an interference phenomenon, it can be destroyed by inelastic scattering. This can be built into our picture in a simple way. Equation (5.13) is valid until the electron is scattered inelastically, so we can identify the length L with the inelastic scattering length L_{inel}. In general, the inelastic scattering length scales with the temperature as $L_{\text{inel}} \propto T^{-\alpha}$, so that by substituting this into (5.13) we obtain

$$\sigma = \sigma_0 + \alpha a \ln T \tag{5.14}$$

This logarithmic temperature dependence of the conductivity has been observed in a number of systems (MOSFETs, thin films, etc.) and was considered a confirmation of the concept of weak localization in two dimensions.

5.3.3 Quantum Interference

What is the origin of the negative coefficient in equation (5.12)? In order to understand this we first consider a simpler problem: that of quantum interference between two waves travelling in opposite directions around a ring. Our bulk system will then be treated as an ensemble of such rings.

Consider a disordered ring with contacts at two points diagonally opposite each other. We assume that the circumference of the ring is large compared to the electronic mean free path but small compared to the localization length. There are many scattering events but no localization. An electron which starts at one contact can travel to the other contact by one of two routes (Fig. 5.7(a)). The two waves which arrive at the other side have been scattered differently. There is, therefore, no particular phase relationship between them. On the other hand, if we follow the two waves all the way around the ring and consider the effect back at the origin, then the two waves have been scattered *identically* (Fig. 5.7(b)). One is the time-reversed case of the other. They thus arrive back at the origin with the same phase. The probability of returning to the origin is twice what it would have

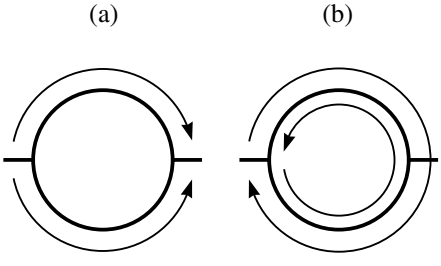

Figure 5.7. Scattering paths round a ring. (a) Waves round opposite sides of the ring interfere randomly on the other side as the paths travelled are different. (b) Waves round the whole ring but in opposite directions interfere constructively back at the origin as the paths travelled are the same.

been had we ignored the interference between the waves and simply added their intensities.

If we consider the probability that a particle in a general disordered system will eventually return to the origin, then the expression we obtain will contain terms which refer to pairs of waves which go around rings in opposite directions, in addition to lots of other terms. These other terms, at least for weak scattering, do not produce any non-classical effects, but the ring interference terms still give an enhanced probability of returning to the origin, and thus a tendency towards localization.

The same interference phenomenon can also be described in **k**-space. In this case a wave which starts with a wavevector **k** has a higher probability of being scattered to $-\mathbf{k}$ than to any other direction. This *correlated backscattering* has given rise to a number of experiments that have looked for optical analogies of weak localization. To avoid any misunderstanding it should perhaps be pointed out that backscattering is different from specular reflection. In the latter case only the component of **k** perpendicular to the surface is reversed, whereas in the weak localization effect all the components of the wavevector are reversed.

5.3.4 Negative Magnetoresistance

What happens when we introduce a magnetic field into this system? Consider again a single ring of radius R. When a magnetic flux, $\Phi = \pi R^2 B$, is fed through this ring the momentum term in the Schrödinger equation† is changed in the usual way (Goldstein, 1950) according to the replacement

$$p^2 \rightarrow (\mathbf{p} - e\mathbf{A})^2$$
$$= (\mathbf{p} - \tfrac{1}{2}eBR\hat{\theta})^2 \tag{5.15}$$

where $\mathbf{A} = \frac{1}{2}BR\hat{\theta}$ is the magnetic vector potential. This breaks the symmetry between $+\mathbf{k}$ and $-\mathbf{k}$. Thus, the two opposite paths around the ring are no longer equivalent, the probability of a particle returning to the origin is no longer enhanced and there is no enhanced backscattering. Translated into the language of a solid rather than a single ring, we see that the chief mechanism which leads to weak localization is no longer active. Hence, the resistance decreases in a magnetic field, i.e. negative magnetoresistance (Fig. 5.8).

Returning though to the single ring, it so happens that it is now possible to make small rings or cylinders with dimensions such that these phenomena can be observed directly. Firstly, consider again the interference between two waves which cross to the opposite side of the ring by different routes. The phase difference between the two waves will be

$$\theta = \theta_0 + \left(2\pi R \times \frac{eBR}{2\hbar}\right) = \theta_0 + \frac{e}{\hbar}\Phi \tag{5.16}$$

† Strictly speaking, we are referring here to the tangential component of the momentum, which is the only important component in a simple ring.

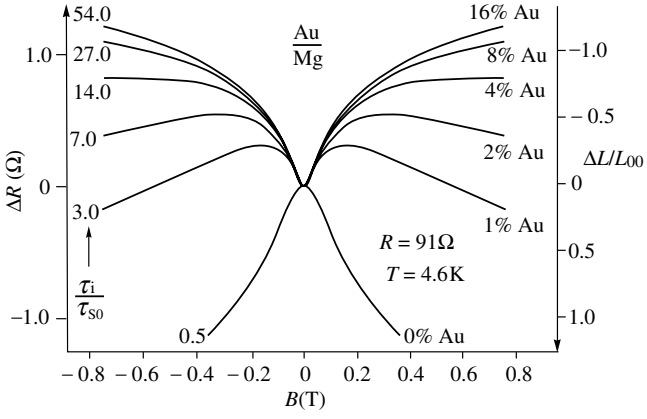

Figure 5.8. The magnetoresistance ΔR of thin Mg–films. The clean film shows a negative magnetoresistance indicating localization. When the film is covered with a small number of gold atoms the magnetoresistance becomes positive due to increasing spin-orbit scattering. The right scale shows the magnetoconductance ΔL. On the left, the ratio of the inelastic scattering time and the spin-orbit scattering time is indicated (Adapted from G. Bergmann (1984), *Phys. Rep.* **107**, with permission of Elsevier Science.

This has the value $\theta = \theta_0 + 2n\pi$ whenever $\Phi = n(h/e)$. Therefore, we expect the current through the ring to vary with a period of one flux quantum.[†] This is one variant of the well known Aharonov–Bohm effect (Aharonov and Bohm, 1959). In fact, when Sharvin and Sharvin (1981) performed the experiment on a hollow magnesium cylinder the period was found to be $2h/e$ (Fig. 5.9). Why? It certainly does not indicate pairing of electrons as might be suspected by anyone familiar with flux quantization in superconducting rings.

In fact, Altshuler, Aronov and Spivak (1981) showed that there is an alternative interference process with precisely this period. As before, we must consider interference back at the origin. In this case the total flux enclosed is doubled and the period is, therefore, $2h/e$. However, since the two paths are identical in the absence of a magnetic field, the term θ_0 vanishes. By contrast, in the Aharonov–Bohm effect, θ_0 is non-zero with a random value from sample to sample. Since Sharvin and Sharvin's cylinder may be considered as an ensemble of such rings, the phase is randomized between samples and no oscillation of the resistance with period h/e is observed.

5.3.5 Single Rings and Non-local Transport

More recently, it has become possible to etch very fine patterns on metal films, leaving very fine wires. An example is shown in Fig. 5.10. Note the periodic

[†] The flux quantum is h/e rather than $h/2e$, as we are dealing with a single electron effect rather than one due to pairs of electrons, such as superconductivity.

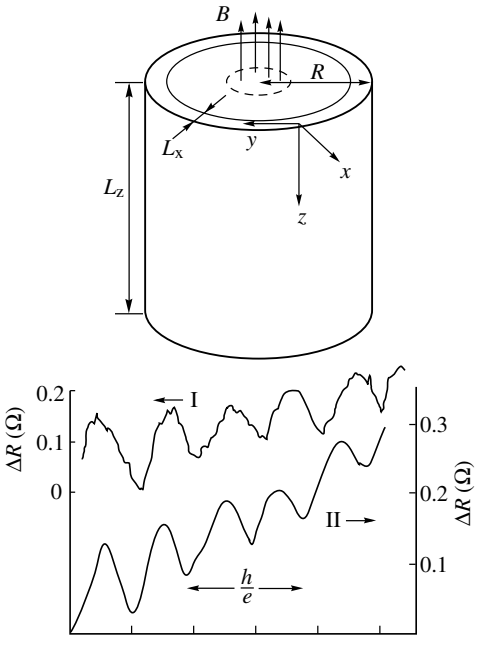

Figure 5.9. Aharonov–Bohm like magnetoconductance oscillations observed in normally conducting Mg cylinders of diameter 1.5 μm by Sharvin and Sharvin (1981). Left and right resistance scales correspond to samples 1 and 2, respectively. The periodicity of the oscillations corresponds to $\Delta\Phi = h/2e$. (Courtesy V. Gantmakher)

oscillations of the current as a function of the magnetic fields with the period of a single flux quantum (h/e), in contrast with the Sharvin and Sharvin experiment discussed in the preceding section.

An even more dramatic example of quantum interference effects on transport in microstructures can be seen in Fig. 5.11. Here the figure without the ring (or head) shows universal conductance fluctuations (Section 9.2.3) whereas, in the second figure, with a ring, a periodic oscillation is superimposed. Note that, classically, the ring is irrelevant, as it constitutes a dead end for the current, but that quantum interference between different paths around the ring can still contribute, leading to the oscillations. This is the first of a range of phenomena involving *non-local* transport, in which classically irrelevant interference paths can contribute to transport.

Consider a sample with several different leads in which a current is sent between leads k and l and a potential difference is measured between leads i and j. The result may be defined in terms of a generalized resistance $R_{ij;kl}$ such that

$$V_{ij} = R_{ij;kl} I_{kl} \tag{5.17}$$

This is a very general notation for describing most common transport measurements, and is often used to represent non-local effects.

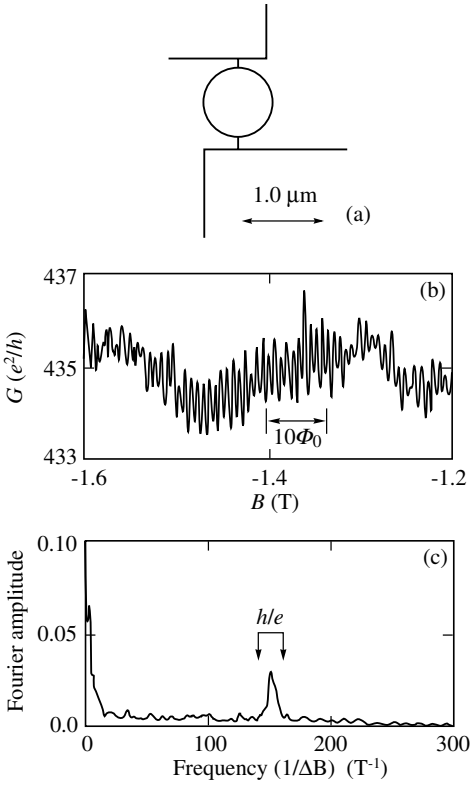

Figure 5.10. Aharonov–Bohm oscillations in a small ring. The period of the oscillations is one flux quantum (i.e. $\Delta\Phi = h/e$) (Umbach *et al.*, 1987). (Courtesy C. Van Haesendonck)

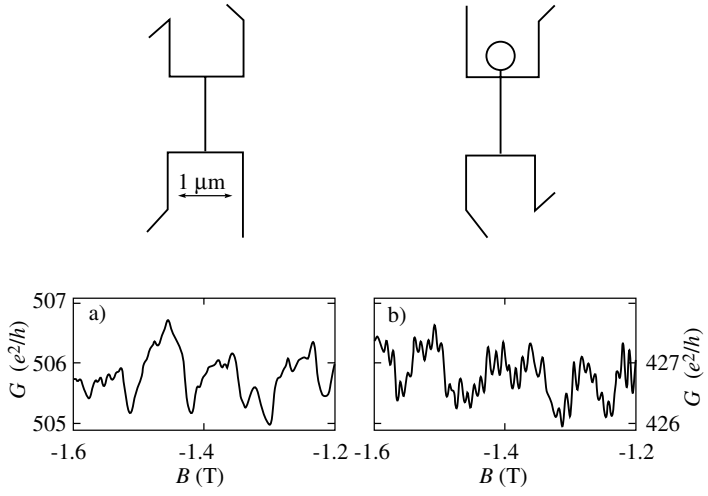

Figure 5.11. Non-local transport in thin wires. In (a) only random fluctuations are observed. In (b), however, interference round the 'head' contributes a periodic oscillation (Umbach *et al.*, 1987). (Courtesy C. Van Haesendonck)

5.3.6 Spin-orbit Coupling, Magnetic Impurities, etc.

Soon after the discovery of weak localization the quantum Hall effect (Section 5.5) was discovered. It rapidly became clear that equation (5.12) is not universally valid. In fact, there are three important exceptions. Besides high magnetic fields (see below), spin-orbit coupling and magnetic impurities will also give deviations.

In the case of spin-orbit coupling the deviation from the classical behaviour is positive. In fact, it has not been shown analytically that there can be any localized states for purely spin-orbit scattering. In a beautiful set of experiments on Au doped Mg, Bergmann (1984) was able to demonstrate the validity of the perturbation theory for such systems (Fig. 5.8) (note that spin-orbit scattering rises as Z^4). The spin-orbit effect is sometimes termed *weak anti-localization*.

Magnetic impurities destroy the weak localization effect by destroying the time reversal symmetry. The β-function has a leading term of $-a/g^2$, and the localization is even weaker than before. Magnetic impurities can also give rise to a divergence of resistance at low temperatures known as the Kondo (1964) effect.

5.3.7 Universal Conductance Fluctuations

So far in our discussion of the conductivity of disordered systems, we have implicitly assumed that the transport properties are *self-averaging*. As is usual in statistical mechanics, we have assumed that the sample is so big that the distribution of possible values of the resistance is very narrow, essentially a δ-function in the case of a large enough sample. In fact, this assumption is invalid. As long as we are working in a regime where the inelastic scattering length is larger than the sample size, the sample cannot be considered as made up of a large number of statistically independent systems. A small change in one place may have consequences for the entire sample. In the regime discussed earlier in this section, the conductance is a number of order e^2/h. In fact, as $T \to 0$ the standard deviation $(\delta g)^2$ is also of order e^2/h. Note that this behaviour is exactly what would be expected if the conductance takes the values e^2/h or zero randomly.

Experimentally this can be measured not by comparing different samples, but by looking at the way the conductance depends on such quantities as gate voltage or magnetic field. In the former case the Fermi level is moved through the spectrum, alternately seeing regions of allowed states and gaps. In the latter the spectrum is moved in a systematic way. The results look like noise (Fig. 5.11(a)). Unlike noise, however, the structure is reproducible. Different samples behave qualitatively similarly but differ in details. These *universal conductance fluctuations* are discussed at length in Section 9.2.3.

5.3.8 Ballistic Transport

By using a so-called *split gate* it is possible to study the transition from 2D to one-dimensional behaviour. If the scattering is weak and the one-dimensional channel

short enough, it is possible to measure the remarkable phenomenon discussed in Section 3.4.1. When the conductance of such a channel is measured as a function of electron density (i.e. conventional gate voltage) it is found to be quantized as (Section 3.4.1)

$$G = \frac{2e^2}{h} i_{\max}$$

5.4 Interaction Effects

5.4.1 The ln *T* Correction

Although the ln T term in equation (5.14) seems to constitute a proof of weak localization, there is unfortunately another effect which gives rise to a similar term. If we consider interacting free electrons and treat the disorder with perturbation theory, the term α in (5.14) becomes $1 - F$, where F depends on the details of the particle–particle scattering. F is difficult to estimate, but is probably of order unity. Given the uncertainties in the coefficients it is impossible to distinguish between the two effects on the simple basis of resistance measurements alone.

However, the interaction effect leads to a conventional *positive* magnetoresistance. Thus the two effects can be distinguished by studying the influence of a magnetic field on the resistance. In fact, the interaction effect depends on the electron density, such that Uren *et al.* (1980) were able to measure a change in the sign of the magnetoresistance as the gate voltage is varied on a MOSFET.

5.4.2 Wigner Crystallization

A gas of electrons behaves very differently from a gas composed of neutral weakly interacting particles. One of the most striking differences is the behaviour of these two types of gases as a function of the density. At large densities, interactions between the particles in atomic and polyatomic gases become increasingly important. But for an electron gas, the phenomenon of *screening* leads to behaviour that for many purposes may be regarded as that of free electrons. Thus, a high-density electron gas behaves essentially like an ideal gas of fermions. As the density of an atomic or polyatomic gas is lowered, the interactions diminish in importance and the gas approaches ideal behaviour. For an electron gas, however, decreasing the density *increases* the effect of the Coulomb potential because the screening effect becomes much less effective.

These observations led Eugene Wigner (Wigner, 1934, 1938) to propose the existence of a lattice of electrons as the ground state of an interacting gas – what is now called a *Wigner crystal* (Mellor, 1992). Wigner argued that below a certain critical density the kinetic energy will be negligible in comparison to the potential energy. Thus, at low enough temperatures the energy of a system of

electrons would be dominated by the pair-wise Coulomb potential between the particles and the behaviour of the gas will be determined by the configuration that minimizes this potential energy. Since the potential of a random array is higher than that of an ordered array, electrons in this regime will form a *crystal*. In three dimensions, the case that Wigner considered, the lowest potential energy is obtained for a body-centred cubic crystal.

In two dimensions there are two regimes to consider: the *quantum* regime, where $k_B T \ll E_F$, and the *classical* regime, where $k_B T \gg E_F$. The classical regime of Wigner crystallization is relatively easy to achieve when the density n_s of electrons is small, since $E_F \propto n_s$. The potential energy V per electron can then be estimated by $V \approx e^2/4\pi\varepsilon_0 r \propto n_s^{1/2}$ The average kinetic energy can be obtained from the equipartition theorem, so the cross-over temperature where the kinetic and potential energies are of comparable magnitude is $T \propto n_s^{1/2}$. The first observation of a Wigner crystal was, in fact, in the classical regime for electrons on the surface of liquid helium (Grimes and Adams, 1979).

The higher densities n_s (and lower effective masses) of 2DEGs in semiconductors means that $E_F \gg k_B T$. In this (quantum) regime, the kinetic energy of the electrons remains non-zero down to the lowest temperatures, being of order E_F, which leads to the kinetic and potential energies being of comparable magnitude. Thus, electrons in most semiconductors remain in a 'liquid' state even at the lowest temperatures. Achieving lower densities is technically very demanding, so an alternative approach has been to apply a large ($\sim 10\,\mathrm{T}$) magnetic field perpendicular to the 2DEG which has the effect of confining electrons to small ($\sim 5\,\mathrm{nm}$) orbits. This makes the 2DEG easier to solidify and there have been a number of experiments carried out that support the notion that 2DEGs in GaAs crystallize in very high magnetic fields and low temperatures (Goldman *et al.*, 1990).

5.5 The Quantum Hall Effect

5.5.1 General

The classical picture of carrier mobility in a magnetic field is modified substantially for the corresponding measurement on a 2DEG. For a 2DEG in the (x, y)-plane in a uniform magnetic field along the z-direction, $\mathbf{B} = B\mathbf{k}$, the Schrödinger equation may be written in the form

$$\frac{1}{2m}\left[p_x^2 + (p_y - eBx)^2\right]\psi = E\psi \tag{5.18}$$

where $\mathbf{A} = Bx\,\mathbf{j}$ is the magnetic vector potential. By using the substitution

$$\psi(x, y) = \varphi(x)\exp(\mathrm{i}ky) \tag{5.19}$$

this equation can be transformed into the Schrödinger equation for a harmonic oscillator,

$$\left[\frac{p_x^2}{2m}+\frac{e^2B^2}{2m}\left(x-\frac{\hbar k}{eB}\right)^2\right]\varphi = E\varphi \tag{5.20}$$

where $\omega_c = eB/m$ is the cyclotron frequency. If we focus only on the orbital motion of the electrons, then the allowed energies are those of a quantum harmonic oscillator:

$$E_n = (n+\tfrac{1}{2})\hbar\omega_c \tag{5.21}$$

The states corresponding to different n are called *Landau levels*. The transformation (5.19) shows that the effect of the magnetic field is to change only the motion along the x-direction; the motion along the y-direction corresponds to free electrons. The solutions φ of the Schrödinger equation (5.20) are harmonic oscillator wavefunctions centred at x_0, which is given by

$$x_0 = \frac{\hbar k}{eB} \tag{5.22}$$

where k is the wavevector associated with the motion of the electron along the y-direction. Thus, x_0 is seen to be a good quantum number.

The energy spectrum of this system is thus a regularly spaced sequence of Landau levels, each separated by an energy $\hbar\omega_c$. To distribute the original density of states among these discrete levels requires each level to have an enormous degeneracy. This degeneracy can be calculated from the number of centre coordinates x_0 that can be accommodated within the sample subject to the Pauli exclusion principle. For a sample of dimensions $L_x \times L_y$, the centre coordinates are separated along the y-direction by

$$\Delta x_0 = \frac{\hbar}{eB}\Delta k = \frac{\hbar}{eB}\times\frac{2\pi}{L_y} = \frac{h}{eBL_y} \tag{5.23}$$

which corresponds to a degeneracy N_0 of

$$N_0 = \frac{L_x}{\Delta x_0} = \frac{L_xL_yeB}{h} \tag{5.24}$$

The degeneracy N per unit area is therefore given by

$$N = \frac{N_0}{L_xL_y} = \frac{eB}{h} \tag{5.25}$$

Note that this degeneracy depends on two fundamental constants (e and h) and on an experimentally controllable quantity (B). In particular, there is no dependence on any parameters associated with the particular material, such as the effective mass of the electrons. When disorder or impurities are included this degeneracy is broken, but as long as the cyclotron energy is large compared with the potential fluctuations the basic structure remains.

Let us now consider the behaviour of the conductivity as the electron density is varied (e.g. in a MOSFET, as discussed in Chapter 10). Whenever the 2D electron density n_s is varied by more than the degeneracy of a Landau level, the Fermi level jumps from one Landau level to another. When the Landau level is broadened we expect the conductivity roughly to follow the density of states, with a maximum in the longitudinal conductivity σ_{xx} when each level is half-filled. The conductivity will therefore vary periodically as the density is varied. The Fermi level E_F is in a Landau level when

$$E_F = (n + \tfrac{1}{2})\hbar\omega_c = (n + \tfrac{1}{2})\hbar\frac{eB}{m} \tag{5.26}$$

In the middle of the Landau level the index n can be related to n_s by using the degeneracy N. To do so we must ignore the contribution to the energy due to the interaction of electron spin with the magnetic field, in effect doubling the degeneracy. Thus

$$n_s = 2N(n + \tfrac{1}{2}) = 2\frac{eB}{h}(n + \tfrac{1}{2}) \tag{5.27}$$

Alternatively we can use the filling factor $i = 2n + 1$ to obtain

$$n_s = iN = i\frac{eB}{h} \tag{5.28}$$

It is instructive to relate this to the Fermi energy E_F by using (5.26) and (5.27):

$$n_s = 2\frac{eB}{h} \times \frac{E_F}{\hbar eB/m} = \frac{E_F m}{\pi\hbar^2} \tag{5.29}$$

which is exactly the relationship one obtains from the density of states in (2.14) in the absence of the magnetic field. If the density of states at the Fermi energy of the 2DEG is zero, i.e. if the Fermi lies between filled and unfilled Landau levels, the carriers cannot be scattered and the cyclotron orbit drifts in a direction perpendicular to the electric and magnetic fields. In this case, the conductivity σ_{xx} (current flow in the direction of the electric field) becomes zero, since the electrons are moving like free particles perpendicular to the electric field with no diffusion (originating from scattering) in the direction of the field. Experimentally, σ_{xx} is never precisely zero, but becomes unmeasurably small at high magnetic fields and low temperatures.

In the Shubnikov–de Haas effect we measure the longitudinal resistivity ϱ_{xx} of the sample as the magnetic field is varied. At low temperatures, $k_B T \ll \hbar\omega_c$, where $\hbar\omega_c$ is the Landau level separation, $\sigma_{xx} \to 0$ and $\varrho_{xx} \to 0$ whenever the Landau levels are full, i.e. equation (5.28) holds. By plotting ϱ_{xx} against $1/B$ and looking for the periodicity it is possible to measure the electron density. Complications occur in real systems due to the additional spin splitting and to valley effects in semiconductors. The reason why both σ_{xx} and ϱ_{xx} vanish for full Landau levels can be appreciated from Exercise 9.

5.5.2 The Quantum Hall Effect Measurements

Probably the most remarkable effect observed in low-dimensional systems was first discovered by von Klitzing, Dorda and Pepper in 1979 (von Klitzing *et al.*, 1980). Their results are illustrated in Fig. 5.12. A constant current I_x is imposed on the 2DEG between source and drain (see inset Fig. 5.12), the longitudinal voltage U_{pp} measured between two probes along the sample and the Hall voltage U_{H} measured between two probes across the sample, as the magnetic field B perpendicular to the 2DEG is varied. The voltage measurements are usually interpreted in terms of the component of a resistivity tensor ϱ with

$$\varrho_{xy} = \frac{U_H}{I_x} \quad \text{and} \quad \varrho_{xx} = f\,\frac{U_{pp}}{I_x} \tag{5.30}$$

where f is a geometrical factor. The components of the conductivity tensor $\sigma = \varrho^{-1}$ are then related to those of ϱ by (Exercise 9)

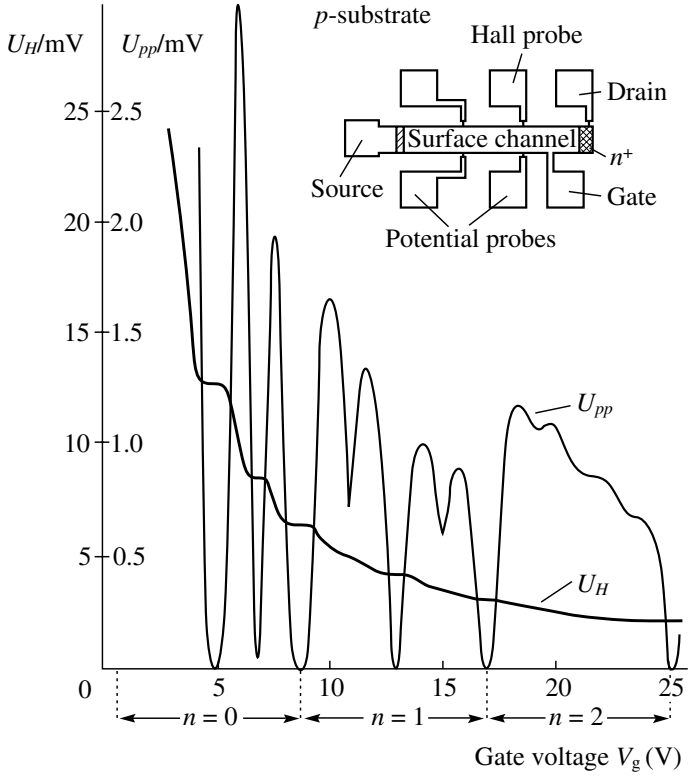

Figure 5.12. Normal U_{pp} and Hall U_{H} potential versus gate voltage, V_{g}, in a silicon MOSFET at $T = 1.5$ K with a magnetic field of 18 T and a source drain current of 1 μA (von Klitzing *et al.*, 1980). The inset shows a diagram of the sample. (Courtesy K. von Klitzing)

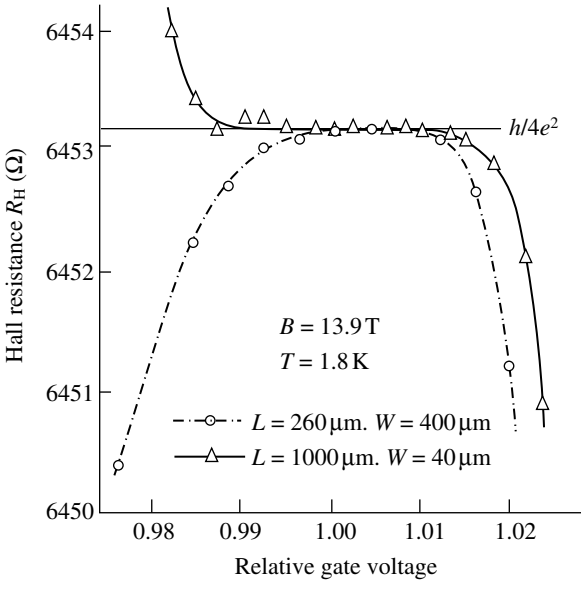

Figure 5.13. Close up of a single plateau for two samples of different geometry. Note the scale and the similarity between the plateau values for the two samples (von Klitzing *et al.*, 1980). (Courtesy K. von Klitzing)

$$\varrho_{xx} = \varrho_{yy} = \frac{\sigma_{xx}}{\sigma_{xx}^2 + \sigma_{xy}^2}, \quad \varrho_{xy} = -\varrho_{yx} = -\frac{\sigma_{xy}}{\sigma_{xx}^2 + \sigma_{xy}^2} \tag{5.31}$$

For a 2DEG with electron density per unit area n_s given by (5.28), one expects (Exercise 10)

$$\varrho_{xy} = \frac{h}{ie^2} \tag{5.32}$$

But the real surprise comes on closer inspection of the plateau, as in Fig. 5.13. The value of ρ_{xy}, the Hall resistivity, for this plateau is $h/4e^2$ (1 ± 10^{-6}). The measurement can now be almost routinely carried out with an accuracy of around 1 part in 10^8. This accuracy is astonishing. Typically, agreement between theory and experiment of 1% or even 10% is considered good in condensed matter physics. The only other phenomenon which comes anywhere close to this accuracy is the AC Josephson effect in superconductors. In fact, the accuracy is such that the quantum Hall effect has now been internationally adopted as the standard of resistance.

5.5.3 The Semiclassical Theory

In order to gain some understanding of the effect we consider a simple approximate picture. It should be borne in mind, however, that a full explanation may not contain any significant approximations as the result appears to be exact.

The equation of motion for a classical particle in a magnetic field is

$$m\frac{\partial \mathbf{v}}{\partial t} = e\mathbf{B}\times\mathbf{v} + e\mathbf{E} \tag{5.33}$$

where $\mathbf{v}$ is the velocity of the particle and $\mathbf{B}$ and $\mathbf{E}$ the magnetic and electric fields, respectively. The solution of this equation has the form

$$\mathbf{v} = \omega_c\left(A_x\cos\omega_c t, A_y\sin\omega_c t\right) + \frac{1}{B^2}\,\mathbf{B}\times\mathbf{E} \tag{5.34a}$$

$$\mathbf{r} = \mathbf{r}_0 + \left(A_x\sin\omega_c t, -A_y\cos\omega_c t\right) + \frac{t}{B^2}\,\mathbf{B}\times\mathbf{E} \tag{5.34b}$$

This behaviour consists of two parts: a circular motion with radius A and frequency $\omega_c = eB/m$ (the cyclotron frequency), and a drift velocity perpendicular to both fields with magnitude $v_\mathrm{d} = E/B$:

$$v_\mathrm{d} = \frac{1}{B^2}\mathbf{B}\times\mathbf{E} \tag{5.34c}$$

In our case, where we define the magnetic field to be perpendicular to the 2D plane and the electric field to be in this plane, it is important that v_d is perpendicular to the electric field.

If we return to the general problem of an electron in a disordered system in a magnetic field and assume now that the radius A is small compared with the rate at which the potential changes, then the electron sees a constant local potential gradient. This is of course indistinguishable from an electric field. With this idea in mind we substitute the local potential gradient for the electric field in the drift velocity to obtain

$$v_\mathrm{d} = \frac{1}{\mathrm{B}^2}\mathbf{B}\times\nabla\phi \tag{5.35}$$

Hence, the drift velocity is always perpendicular to the local potential gradient, $\nabla\phi$.

Let us return now to our analogy with a random landscape. The drift velocity follows the contour lines rather than going up or down the mountains. It is as though it had misread its map and mistaken the contour lines for paths. Using Fig. 5.1 again, we see that most contours are around the tops of mountains or the bottoms of valleys. In fact, for a random landscape there is only one level where a contour crosses the whole system. This is the situation depicted in Fig. 5.1(c). Using our previous picture of varying water levels, the electron is now constrained to follow the shoreline. It is only at this one level that it can manage to cross.

Translating this result into the language of localized and extended states we have the result that all states are localized except for those at a single level. Thus, almost none of the states can contribute to the current. Clearly this gives us a simple mechanism by which it is possible to vary the number of electrons in the system while the current remains unchanged.

Now let us consider the current through a cross-section of the system, as illustrated in Fig. 5.14. Firstly, consider a system with a uniform charge density, $n_s e$ corresponding to a full Landau level. This condition will be relaxed later. The current density through the cross-section is

$$J_x = n_s e v_\mathrm{d} = \frac{n_s e}{B}\frac{\partial \phi}{\partial x} \tag{5.36}$$

and the total current is given by the integral from one side to the other

$$I_x = \frac{n_s e}{B}\int \frac{\partial \phi}{\partial x}\,\mathrm{d}x = \frac{n_s e}{B}\Delta\phi \tag{5.37}$$

Hence, the total current is independent of the details of the potential but only depends on the potential difference, $\Delta\phi$, across the sample. Since, however, we already know that only the electrons at a particular level can contribute to the current, it must be possible to remove the electrons from the localized states at the tops of mountains or the bottoms of valleys without altering the total current.

Since we have a well-defined current which does not change when electrons are added or removed, we have a mechanism for the plateaux in the quantum Hall effect. Note, however, that we do not yet have a value for the Hall conductivity associated with each plateau. For that we need quantum mechanics. The result e^2/h contains Planck's constant, after all.

We return therefore to (5.34) and look at the part describing circular motion. The root mean square (rms) deviation of the electron from the orbit centre is

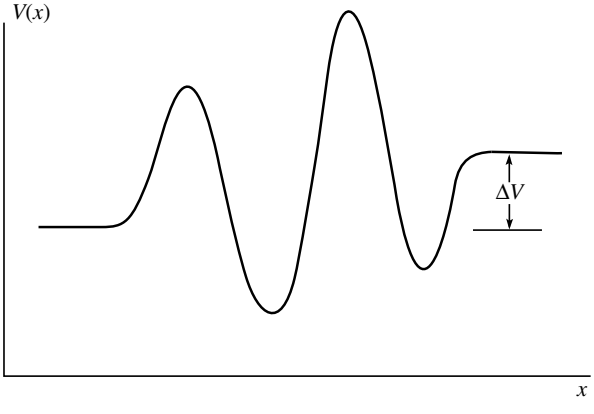

Figure 5.14. Cross-section through a random potential with a net potential difference ΔV between the two sides.

$\Delta x = A/\sqrt{2}$. Since the motion is circular, this immediately gives us the rms deviation of the momentum, namely $\Delta p = m\omega_c \Delta x$, where $\omega_c = eB/m$, the cyclotron frequency. Using the Heisenberg uncertainty principle, $\Delta x \Delta p_x = \hbar$, gives us a formula for the radius of the orbit:

$$A^2 = \frac{2\hbar}{m\omega} = \frac{2\hbar}{eB} \tag{5.38}$$

The Pauli exclusion principle reserves the area of a circle of radius A for each electron, so that the maximum uniform charge density must be

$$n_s e = e\frac{1}{\pi A^2} = \frac{e^2 B}{2\pi\hbar} = \frac{e^2}{h}B \tag{5.39}$$

Substituting this in (5.37) gives us the result we have been seeking:

$$I_y = \frac{e^2}{h}\Delta\phi \tag{5.40a}$$

or

$$J_y = \sigma_{xy} E_x \tag{5.40b}$$

which serves as a definition of the Hall conductivity, σ_{xy}. Our analysis in terms of a semiclassical picture of electron motion in a magnetic field and a slowly varying potential illustrates several important aspects of the physics of the quantum Hall effect.

- The presence of disorder and, hence, of localized states is vital to the explanation of the effect and its precision.
- The density corresponding to a full Landau level and, hence, the quantized state corresponds to that of an incompressible fluid. The Pauli exclusion principle forbids us to compress the system.
- Higher Landau levels contain the same density of electrons, so a full Landau level corresponds to a Hall conductivity of $\sigma_{xy} = ie^2/h$.
- The number of open contours in the above semiclassical argument is proportional to the potential difference across the sample. In fact, that number is independent of the disorder. Hence, there is no reason to worry that a macroscopic current is being carried by a single state.

5.5.4 The Fractional Quantum Hall Effect

Just as the quantum Hall effect seemed to be understood, at least at a basic level, Tsui, Störmer and Gossard (1982) of Bell Laboratories in Murray Hill, New Jersey published some remarkable experiments on $GaAs/Al_{0.3}Ga_{0.7}As$ prepared by MBE. The experiment produced the startling result that the Hall conductivity not only has steps in integer multiples of $\sigma_{xy} = n(e^2/h)$, where n is an integer, but they also observed steps at $n = \frac{1}{3}$ and $n = \frac{2}{3}$. Since then features have been observed at $n = \frac{1}{3}, \frac{2}{3}, \frac{2}{5}, \frac{3}{5}, \frac{3}{7}, \frac{4}{7}, \frac{4}{9}, \frac{5}{9}, \frac{4}{3}, \frac{5}{3}$ and many more rational fractions,

all with odd denominators (Chang *et al.*, 1984). Typically, these have an accuracy of 10^{-5} for the $\frac{1}{3}$ feature and lower accuracy on the others. For some fractions the feature is only observed as a peak in σ_{xx}. Fig. 5.15 shows a typical spectrum.

There have been many attempts to explain these results. Some of the theories rely on some very exotic mathematics and many are still controversial. Here, we will try to give a simple explanation based on those aspects on which there is general agreement.

Consider first the wavefunction for 2D electrons in a perpendicular magnetic field expressed in the so-called symmetric gauge, where the magnetic vector potential in (5.19) takes the form $\mathbf{A} = \frac{1}{2}(-By, Bx, 0)$. The wavefunctions in the lowest Landau level are

$$\psi(z) \propto z^m \exp\left(-\frac{|z|^2}{4l_c^2}\right) \tag{5.41}$$

where $z = x + \mathrm{i}y$ and l_c is the cyclotron radius. Laughlin (1983) proposed the following generalization of this to N electrons

$$\Psi_m(z_1, z_2, \ldots, z_N) \propto \prod_{j<k} (z_j - z_k)^m \prod_i \exp\left(-\frac{|z_i|^2}{4l_c^2}\right) \tag{5.42}$$

A useful insight into the meaning of this wavefunction can be obtained by thinking of $|\Psi|^2$ as a probability distribution of a classical plasma, obeying a partition

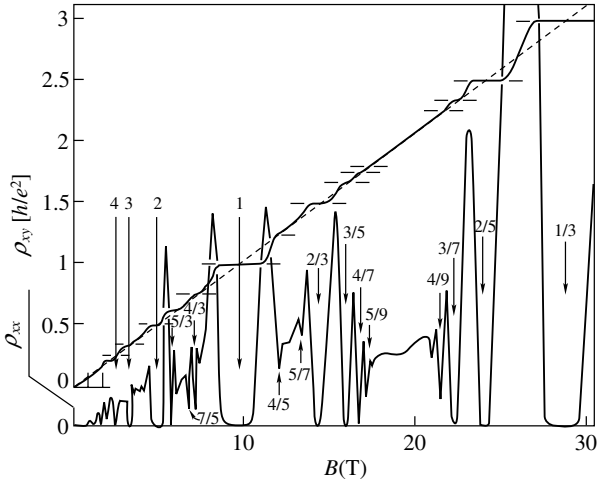

Figure 5.15. A recent example of the Hall resistivity, ρ_{xy}, and normal resistivity, ρ_{xx}, as a function of magnetic field, in Tesla (Willet *et al.,* 1987). Note the large number of fractional features marked. (Courtesy A. C. Gossard)

function $Z = \exp(-\beta\Phi)$ in which $\beta = 1/m$ and

$$\Phi = -2m^2 \sum_{jk} \ln(z_j - z_k) + \tfrac{1}{2} m \sum_i \frac{|z_i|^2}{l_c^2} \tag{5.43}$$

This corresponds to a system of particles which repel each other logarithmically and a uniform background density $\rho_m \propto 1/m$.[†]

Laughlin's wavefunction has several properties which are useful for understanding the effect.

- Ψ_m corresponds to $1/m$th occupation of the lowest Landau level.
- When m is odd, the wavefunction is odd under exchange of electrons.
- This restriction to odd m explains why fractional densities with odd denominators are special.
- At the special densities it can be shown that the system has an energy gap (Laughlin, 1983): more energy is required to add an electron than is gained by removing one. As with the gap in a semiconductor, or, more accurately in this case, a superconductor, the state is stable unless excitations are possible across the gap, due to finite temperature, radiation, etc.

Interesting though Laughlin's wavefunction certainly is, in its simplest form it is restricted to fractions of the form $1/m$. How can we understand the occurrence of other fractions and the order in which they appear? A very simple semiclassical picture is illustrated in Fig. 5.16. Here, we see a ring of states with one-third of the states occupied (Fig. 5.16(a)). When one particle is added we find a situation as in Fig. 5.16(b). This is not the ground state, however. If the particles mutually repel one another the ground state will be one in which their density is most uniform, as shown in Fig. 5.16(c). Note the peculiar feature of the relaxed system: instead of a single *defect* due to the additional particle there are now three defects which are as far apart as possible. Thus, the additional particle is behaving as if it is three *quasi-particles*, each with charge of $\frac{1}{3}$. A similar result, with one particle removed, is shown in Figs. 5.16(d,e). Again there are three *defects*.

To a semiconductor physicist there should be no difficulty in understanding the concept of quasi-particles illustrated here. We are all well used to the concept of a *hole*, which is used to describe an almost full band of electrons. Indeed we often tend to forget that it is not a true particle.

We now require a leap of the imagination. It has been shown that a $\frac{1}{3}$ charged Landau level behaves as if it is *full* of particles of charge $\frac{1}{3}$. Similarly for all $1/m$th full levels. Consider now a level which is $1/m$th full of $\frac{1}{3}$ charged quasi-particles. These could now form a special state in which a new generation of *quasi*-quasi-particles is formed. These can then condense into another special state and form *quasi*-quasi-

[†] The solution of Poisson's equation in two dimensions is logarithmic, so that *truly* 2D particles would repel each other logarithmically. Alternatively, a system of charged rods in three-dimensional space also interacts logarithmically.

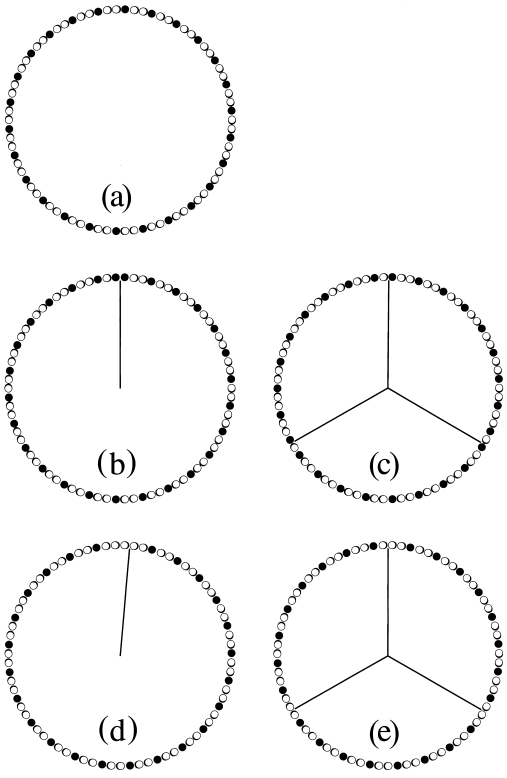

Figure 5.16. Mutually repulsive classical particles restricted to specific ordered sites on a ring. (a) Density is $\frac{1}{3}$; (b) density $\frac{1}{3}$ with one additional particle (unrelaxed); (c) as (b) (relaxed); (d) density $\frac{1}{3}$ with one particle removed (unrelaxed); (e) as (d) (relaxed).

quasi-particles, and so on. A whole hierarchy of different states involving different generations of quasi-particles can arise, each corresponding to a different fractional density and each giving rise to a feature in the transport measurements.

EXERCISES

1. The quantity e^2/h is sometimes called the *quantum of conductance* because it arises as the unit of quantization in, e.g., the quantum Hall effect, ballistic transport and universal conductance fluctuations.

(a) Show that the quantum conductance has the correct units for a conductance (i.e. inverse ohms).

(b) Show that a single quasi-1D channel in a crystalline system carries a current $I = (e^2/h)\Delta V$, where ΔV is the voltage drop along the length of the channel.

(c) Use the Ioffe–Regel criterion to show that in two dimensions the minimum metallic conductivity is $\sigma_{\min} \propto e^2/h$.

2. Define the terms 'localized,' extended' and 'mobility edge' as used to describe the behaviour of electrons in a non-crystalline solid.

3. Describe briefly the two main modes of electronic transport in a disordered system at low temperatures when the chemical potential lies in an energy range in which the states are localized.

4. A measurement of the conductivity of a silicon MOSFET gives a straight line when the conductivity is plotted against the logarithm of the temperature. Describe the physical process which gives rise to the logarithmic temperature dependence and explain how this is related to the behaviour of electrons in small rings.

5. The result of a 4-probe measurement of resistivity is often written in the form

$$V_{ij} = R_{ij,kl} I_{kl}$$

where the voltage difference, V_{ij}, between probes i and j is associated with the current, I_{kl}, between probes k and l.

(a) Two samples of wires of submicron dimensions are prepared as in Fig. 5.17. Explain (quantitatively where possible) the behaviour you would expect for the temperature and magnetic field dependence of $R_{12,34}$ and $R_{13,24}$ in these samples.

(b) A third sample is prepared in which many rings, like the one in Fig. 5.17(a) and nominally all of the same diameter, are arranged in parallel between the two horizontal wires. What differences would you expect to observe in this sample compared with the sample in Fig. 5.17(b)?

6. The metal-insulator transition is often described in terms of a β function

$$\frac{\mathrm{d}\ln g}{\mathrm{d}\ln L} = \beta(\ln g)$$

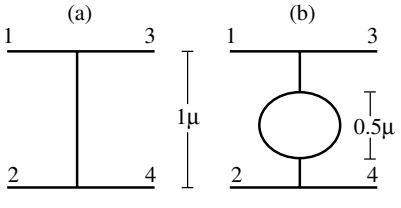

Figure 5.17. Figure for Exercise 5.

where g is the dimensionless conductance, $g = (e^2/h)G$, and L is the length of a side of a cubic sample.

(a) By assuming that g is a smooth, i.e. differentiable, function of the Fermi energy, E_F for all finite L, show that g depends on L as $g(L/\xi)$, where ξ is independent of L.

(b) Close to the transition g can be written in the form

$$\ln g = \ln g_c + A(E_F - E_c)L^{\alpha}$$

where $\alpha = \mathrm{d}\beta/\mathrm{d}\ln L$ when $\beta = 0$. Hence, show that $\xi \propto |E_F - E_c|^{-\nu}$.

7. Verify equation (5.10) in the form

$$G = \sigma L^{d-2}$$

for the particular cases $d = 3$ and $d = 2$ starting from Ohm's law:

$$\mathbf{J} = \sigma \mathbf{E}$$

(a) In $d = 3$, consider a conductor of rectangular cross-section ac with the current flow parallel to a side of length b.

(b) In $d = 2$, consider a planar conductor of width a and current flow parallel to a side of length b.

8. The Schrödinger equation for electrons in the (x, y)-plane subject to a magnetic field B in the z-direction and an electric field E in the x-direction may be written, in the Landau gauge, as

$$\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + \frac{1}{2m}\left(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial y} - eBx\right)^2\psi + eEx\psi = \epsilon\psi$$

(a) Using the substitution $\psi \to \phi(x)\exp(\mathrm{i}ky)$, show that the Schrödinger equation may be transformed into that of a harmonic oscillator centred around

$$X = \frac{\hbar k}{eB} - \frac{meE}{e^2B^2}$$

with a spring constant $K = e^2B^2/m$, plus some additional constants.

(b) Using your knowledge of the harmonic oscillator, confirm that the eigenenergies are given by

$$\epsilon_n = (n + \tfrac{1}{2})\hbar\frac{eB}{m} + \text{additional terms}$$

and give an interpretation of the additional terms.

9. The relationship between the conductivity tensor σ and the resistivity tensor ϱ is $\sigma = \varrho^{-1}$.

(a) Derive the expressions for ϱ_{xx} and ϱ_{xy} in equation (5.30) starting from the conductivity tensor

$$\sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ -\sigma_{xy} & \sigma_{xx} \end{pmatrix}$$

(b) At high B-field in a high-mobility 2DEG, $\sigma_{xy} \gg \sigma_{xx}$. In this limit, what is the relationship between ϱ_{xy} and σ_{xy}?
(c) In the limit in (b), what happens to ϱ_{xx} if $\sigma_{xx} \to 0$?

10. Assuming that the Hall bar in the inset to Fig. 5.12 has width a and that the distance between the longitudinal probes is b, show that if

$$\varrho_{xx} = f\frac{U_{pp}}{I_x} = fR_L$$

in the Hall condition ($J_y = 0$), then $f = a/b$. Furthermore, show that if

$$\varrho_{xy} = \frac{U_{\mathrm{H}}}{I_x}$$

in the Hall condition, then

$$R_{\mathrm{H}} = \frac{E_y}{J_x B_z} = \frac{\varrho_{xy}}{B_z}$$

Hence, using equations (5.34c) and (5.36), show that if (5.39) holds, then

$$\varrho_{xy} = \frac{h}{ie^2}$$

11. Shubnikov–de Haas oscillations were studied in a high-mobility 2DEG in a heterostructure at 4.2 K. Minima were observed in the longitudinal resistance at certain values of the perpendicular **B**-field and the filling factors i at the various minima were identified. A graph of i against $1/B$ was plotted and the data gave a straight line of gradient 8 T. Use this information to find n_s, the electron density per unit area in the 2DEG.

In the same experiment the longitudinal resistance at zero **B**-field was measured to be 100 Ω. The distance between the potential probes along the length of the Hall bar (b) was found to be 7.5 times the width of the bar (a). Use this information and the value of n_s calculated above to estimate the mobility of the 2DEG.

References

E. Abrahams, P. W. Anderson, D. C. Licciardello and T. V. Ramakrishnan, *Phys. Rev. Lett.* **42**, 673 (1979).
Y. Aharonov and D. Bohm, *Phys. Rev.* **115**, 485 (1959).
B. L. Altshuler, A. G. Aronov and B. Z. Spivak, *Sov. Phys. JETP Lett.* **33**, 94 (1981).
P. W. Anderson, *Phys. Rev.* **109**, 1492 (1958).
N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976)

G. Bergmann, *Phys. Rep.* **107**, 1 (1984).
A. M. Chang, P. Berglund, D. C, Tsui, H. L. Störmer and J. C. M. Hwang, *Phys. Rev. Lett.* **53**, 997 (1984).
V. J. Goldman, M. Santos, M. Shayegan and J. E. Cunningham, *Phys. Rev. Lett.* **65**, 2189 (1990).
H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1950).
C. C. Grimes and G. Adams, *Phys. Rev. Lett.* **42**, 795 (1979).
A. F. Ioffe and A. R. Regel, *Prog. Semicond.* **4**, 237 (1960).
S. Kawaji, T. Igarashi and J. Wakabayashi, *Prog. Theor. Phys. Suppl.* **57**, 176 (1975).
K. von Klitzing, G. Dorda and M. Pepper, *Phys. Rev. Lett.* **45**, 494 (1980).
J. Kondo, *Prog. Theoret. Phys.* **32**, 37 (1964).
R. B. Laughlin, *Phys. Rev. Lett.* **50**, 1395 (1983).
P. E. Lindelof, J. Nørregaard and J. Hanberg, *Phys. Scr.* **T14**, 17 (1986).
A. MacKinnon and B. Kramer, *Z. Physik* **B53**, 1 (1983a).
A. MacKinnon and B. Kramer, in *High Magnetic Fields in Semiconductors*, G. Landwehr, ed., Lecture Notes in Physics, Vol. 177 (Springer, Berlin, 1983b) p. 74.
C. Mellor, *New Scientist* **135**(1833), 36 (1992).
N. F. Mott, *J. Non–Cryst. Sol.* **1**, 1 (1968).
N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials*, 2nd edn (Clarendon Press, Oxford, 1979).
M. Pepper, in *The Metal Non-Metal Transition in Disordered Solids*, L.R. Friedman and D.P. Tunstall, eds. (SUSSP, Edinburgh, 1978) pp. 285.
D. Y. Sharvin and Y. V. Sharvin, *Sov. Phys. JETP Lett.* **34**, 272 (1981).
D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor and Francis, London, 1994).
D. C. Tsui, H. L. Störmer and A. C. Gossard, *Phys. Rev. Lett.* **48**, 1559 (1982).
C. P. Umbach, P. Santhanam, C. Van Haesendonck and R. A. Webb, *Appl. Phys. Lett.* **50**, 1289 (1987).
M. J. Uren, R. A. Davies and M. Pepper, *J. Phys. C* **13**, L985 (1980).
E. Wigner, *Phys. Rev.* **46**, 1004 (1934).
E. Wigner, *Trans. Farad. Soc.* **34**, 678 (1938).
R. Willet, J. P. Eisenstein, H. L. Stormer, D. C. Tsui, A. C. Gossard and J. H. English, *Phys. Rev. Lett.* **59**, 1776 (1987).