QUANTUM TRANSPORT IN THE 2D ELECTRON GAS*

E.H. HAUGE

Department of Physics, The Norwegian Institute of Technology
The University of Trondheim, N-7034 Trondheim, Norway

(Received February 15, 1995)

A quick introduction is provided to basic elements of quantum transport in a mesoscopic two-dimensional electron gas (2DEG). The introduction is designed for an audience with a background mostly in statistical physics.

PACS numbers: 05.30.Fk

1. Introduction

The following notes summarize lectures held at the Seventh Symposium on Statistical Physics during the last 10 days of September 1994, in Zakopane, Poland. The lectures were designed as a crash course on quantum transport in a mesoscopic two-dimensional electron gas (2DEG), introducing this fascinating field to an audience with a background mostly in statistical physics.

Quantum transport in the 2DEG has been a growth industry with intense activity over the last 15 years, when "$t = 0$" is conveniently defined by von Klitzing's spectacular discovery of the quantum Hall effect (QHE) in 1980. An increasingly sofisticated technology with the ability to produce, in a controlled way, a variety of structures with a linear scale of 100nm or less ("nanostructures") has opened up a playground full of exciting possibilities for experimentalists and theorists alike. It is conceivable that a deeper understanding of quantum transport in these small structures might lead to an era of commercially viable nanoscale electronics. These speculative prospects have done nothing to dampen the activity in this branch of physics.

---

* Presented at the VII Symposium on Statistical Physics, Zakopane, Poland, September 22–28, 1994.

Basic understanding of some key features of this mesoscopic world will be the focus of these lectures. For a much broader introduction, the reader is referred, for example, to the excellent review by Beenakker and van Houten [1]. In addition to providing an authorative introduction to the field, that review also serves as an extensive source of references to the literature prior to 1991. In keeping with the tutorial slant of the present lectures, the number of references will be kept close to a minimum.

2. Basics

2.1. The 2DEG

We shall not describe the technicalities involved in producing two-dimensional electron gases here. Suffice it to say that they live on interfaces in semiconductor heterostructures, with the AlGaAs/GaAs interface as the most popular example. For concreteness we shall confine ourselves to that system here. In the (z-)direction perpendicular to the interface, the electrons are trapped in a well of triangular shape, and at low temperatures only the ground state is occupied. Thus, the z-direction is frozen out, resulting in an electron gas that for most purposes can be considered as two-dimensional.

In addition, conditions in this 2DEG can be controlled in a number of ways. Typically, metallic gates in a parallel plane close to the 2DEG can be custom designed on a linear scale down to a few tens of nanometers. When a negative voltage is applied, the electron density in the 2DEG underneath the corresponding gate is depleted. In this way one can control[1] the potential energy $V(x, y)$ seen by the electrons in the 2DEG and construct narrow electron wave guides ("quantum wires"), rings, constrictions, washboards, regular lattices *etc., etc.* This flexibility in manufacturing devices has offered experimentalists and theorists a number of challenging opportunities, with no end in sight.

At temperatures $T < 1\text{K}$ the *inelastic* mean free path or, more to the point, the phase coherence length Λ_ϕ, in these structures can exceed $1\mu\text{m}$. *I.e.*, transport through the entire structure can be considered as *coherent*. Thus, one must expect all kinds of QM interference effects in these mesoscopic systems. We shall confine ourselves to coherent transport in the following. In addition to inelastic scattering, also elastic scattering

[1] As a matter of fact, the control is less than perfect. What one *can* control is the voltage on the gates. The potential seen by the electrons in the 2DEG is the result of Nature's own self-consistent solution of the Poisson equation, also involving the 2DEG electrons themselves.

from random impurities can be reduced to the extent that the elastic mean free path Λ can take values up to 1μm or more.

Much of the excitement in this area stems from the fascinating effects of a magnetic field, usually chosen perpendicular to the 2DEG. The integer and fractional QHE are star examples but they do not, by any means, exhaust the list.

It is remarkable to what extent the physics of the 2DEG can be understood in a single electron picture. This may sound discouraging to the statistical physicist whose bread and butter are interactions and collective effects. However, there is no cause for despair, many-body effects are lurking in the background, ready to be taken seriously! (The fractional QHE providing but one example.) However, it seems wise to exploit the one-particle picture for what it is worth before embarking upon more complex ventures. We shall stay strictly within the single electron picture here, and use the effective mass approximation throughout.

For easy reference, typical GaAs values of some quantities of central interest are given in Table I.

TABLE I

Typical values of crucial quantities in GaAs

$m = 0.067m_e$	effective mass
$n_s = 4 \cdot 10^{15}\mathrm{m}^{-2} = (16\mathrm{nm})^{-2}$	electron sheet density
$E_F = 14\mathrm{meV}$	Fermi energy
$\lambda_F = 40\mathrm{nm}$	Fermi wavelength
$\Lambda = 10^2 - 10^4\mathrm{nm}$	elastic mean free path
$\Lambda_\phi > \Lambda$	phase coherence length
$l_B = \sqrt{\hbar/eB} = 26(B/\mathrm{T})^{-1/2}\mathrm{nm}$	magnetic length

2.2. Density of states

In the 2DEG the density of states (DOS) is, in zero magnetic field, *independent* of energy,

$$\rho(E) = g_s \frac{m}{2\pi\hbar^2} = \frac{m}{\pi\hbar^2} \;\Rightarrow\; n_s = \frac{m}{\pi\hbar^2} E_F\,, \tag{1}$$

with spin degeneracy $g_s = 2$. As a result, the Fermi energy E_F is directly proportional to the electron sheet density n_s.

In a magnetic field (and with negligible disorder) the continuous energy spectrum collapses into discrete Landau levels,

$$E_n = \hbar\omega(n + \tfrac{1}{2}) \;\;;\;\; n = 0, 1, 2\ldots\,, \tag{2}$$

where $\omega = eB/m$ is the cyclotron frequency. As a result, the DOS becomes a sum of delta-functions,

$$\rho(E) = g_s \frac{eB}{h} \sum_n \delta(E - E_n)\,, \tag{3}$$

where we neglected the Zeeman splitting both in (2) and (3).

For transport problems the DOS in a 1D quantum wire is essential. We consider a perfect wire which gives a separable Schrödinger equation, so that the energy takes the form,

$$E = E_\alpha^t + \frac{\hbar^2 k^2}{2m} \equiv E_\alpha(k)\,, \tag{4}$$

where E_α^t is the discrete energy spectrum associated with the *transverse* eigenstates. The dispersion relation $E_\alpha(k)$ of every *subband* (corresponding to a given value of α) has the quadratic form representing free propagation in the longitudinal direction, The DOS for particles moving to the *right* in subband α is

$$\rho_\alpha^+(E) = 2\frac{1}{2\pi} \cdot \frac{dk}{dE_\alpha} = \frac{2}{hv_\alpha(E)}\,, \tag{5}$$

with $v_\alpha(E)$ the corresponding group velocity. It is important to realize that ρ_α^+ is inversely proportional to v_α *irrespective* of the nature of the transverse states. Thus, (5) is valid for any type of transverse confinement, with or without a magnetic field.

2.3. The Einstein relation

The electron current density can be driven by two 'forces', gradients in the electron sheet density and external electric fields,

$$\vec{j} = -\overrightarrow{\overrightarrow{D}} \cdot \nabla n_s + \overrightarrow{\overrightarrow{\sigma}} \cdot \vec{E}/(-e)\,. \tag{6}$$

In equilibrium (under circumstances when persistent currents can be ignored) $\vec{j}$ vanishes. On the other hand, the general equilibrium condition at $T = 0$ says that the (electro-)chemical potential must be the same everywhere,

$$\mu = -eV(\vec{r}) + E_F(n_s(\vec{r})) = \text{const.} \tag{7}$$

Taking the gradient of (7), using that $\vec{E} = -\nabla V$ and that the DOS given by (1) is $\rho(E_F) = dn_s/dE_F$, one finds from (6),

$$\overrightarrow{\overrightarrow{\sigma}} = e^2 \rho(E_F) \overrightarrow{\overrightarrow{D}}\,. \tag{8}$$

This relation between the conductivity and diffusivity tensors is nothing else than the familiar Einstein relation, adapted to the situation at hand.

The Einstein relation allows us, in the context of the theory of *linear* transport, to replace field-generated conduction by gradient-generated diffusion.

2.4. The Landauer formula

Consider coherent transport, at $T = 0$, from one reservoir at chemical potential $\mu + \delta\mu$, to another at chemical potential $\mu = E_F$. The electrons are transported through some mesoscopic "device", coupled to each of the two reservoirs by ideal quantum wires. We are interested in the conductance G of the device, as measured from one reservoir to the other. Or, by the Einstein relation (8), we are interested in its "diffusance" $\tilde{D}$. This trick makes it is abundantly clear that we do not have to worry about real current distributions (which are self-consistently given by the Poisson equation) when calculating[2] G. Focusing on the transport in the quantum wires (let them, for simplicity, be equal), we can think of the right-moving states of the left wire as being filled up to $\mu + \delta\mu$, the left-moving states in the right wire filled up to μ. Clearly, transport processes to the right and to the left cancel one another up to energy $\mu = E_F$. Net transport, at $T = 0$, happens in the narrow energy interval $(E_F, E_F + \delta\mu)$. The net current to the right, in subband (*i.e.*, in transverse mode) α, is given by,

$$J_\alpha = \int_{E_F}^{E_F+\delta\mu} dE\, \rho_\alpha^+(E) v_\alpha(E) T_\alpha(E) \approx \delta\mu \cdot \frac{2}{h} T_\alpha(E_F)\,. \tag{9}$$

Here $T_\alpha(E)$ is the total transmission probability to the right reservoir from incoming mode α. In (9) we used (5) and kept the term linear in $\delta\mu$ only. Summation over all modes propagating into the device gives the diffusance: $J = \sum_\alpha J_\alpha \equiv \tilde{D}\delta n_s$. Since $\tilde{D}$ is related to G in precisely the same way that the diffusivity tensor is related to the conductivity tensor one has, by the Einstein relation (8) and the DOS (1),

$$G = \frac{2e^2}{h} \sum_{\alpha=1}^{N} T_\alpha(E_F)\,; \quad \frac{h}{2e^2} = 12.9\mathrm{k}\Omega\,.$$

---

[2] This says basically nothing beyond the generally true statement that all coefficients of *linear* transport are properties of the *equilibrium* state. Equilibrium time correlation functions as in the Kubo formulas. Or equilibrium scattering matrices at the Fermi energy, as in the Landauer formula to be derived here.

For the case of spin degeneracy equal to 2, this is the (linear[3]) Landauer formula [2].

Remarks:

— The Landauer formula reduces the problem of calculating the conduction to the solution of a standard QM scattering problem. No more, but also no less.

— When the Kubo formula for the conductance is specialized to the present circumstances (with or without a magnetic field), it is *equivalent* to the Landauer formula.

— A perfect quantum wire sustaining N transverse modes propagating between the reservoirs has, according to (11), the conductance $G = (2e^2/h)N$. This looks paradoxical: How can a perfect conductor have a finite conductance? Or, put differently: Where does *dissipation* take place in this case? The answer is: In the resevoirs. Remember that the linear Landauer formula (10) gives the conductance from reservoir to reservoir. Moving an electron from a reservoir at $\mu + \delta\mu$ to one at μ increases the total entropy of the system.

— In the derivation of the Landauer formula we added incoming currents, *not* incoming amplitudes. In other words, basic to the Landauer formula is the assumption that there are no correlations between modes incoming from the source reservoir. But correlations will, in general, *develop* between the modes during the passage through the device. It is essential to keep track of the phase relationships caused by scattering between the modes.

In conjunction with the last remark, we note that it is customary to write

$$T_\alpha = \sum_\beta |t_{\beta\alpha}|^2 \;\Rightarrow\; G = \frac{2e^2}{h}\mathrm{Tr}(tt^\dagger)\,. \tag{11}$$

This needs interpretation. With $A_{\beta\alpha}$ the amplitude of transmitted mode (often called 'channel') β resulting from incoming mode (channel) α, current conservation implies, for every α, that $|e^{ik_\alpha x}|^2 v_\alpha = \sum_\beta |A_{\beta\alpha}|^2 v_\beta$. Let us define $t_{\beta\alpha} \equiv \sqrt{v_\beta/v_\alpha}A_{\beta\alpha}$. We can then write particle conservation in the convenient form,

$$1 = \sum_\beta |t_{\beta\alpha}|^2\,;\, \forall\alpha\,, \tag{12}$$

where α is an incoming mode and $\sum_\beta$ goes over *all* outgoing modes, including those reflected into the initial reservoir (the source). In the Landauer

[3] The formula originally proposed by Landauer was nonlinear. Although aspects of the nonlinearity are interesting in themselves, it is the *linear* version that relates to typical experimental situations.

expresssion (11) for the conductance on the other hand, these same $|t_{\beta\alpha}|^2$'s are summed, not over all β, but over those β's that correspond to transmission into the final reservoir (the drain).

2.5. The Landauer–Büttiker formalism

The Landauer–Büttiker formalism generalizes the linear Landauer formula to a situation with many ideal leads connecting the central device to external reservoirs with given chemical potentials (see Fig. 1)

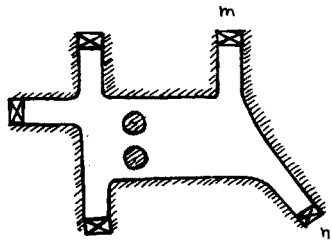

Fig. 1. A "device" of some sort connected by ideal leads to reservoirs with given chemical potentials.

It is convenient to generalize our previous definition of the modified scattering amplitudes to

$$t^{nm}_{\beta\alpha} = \sqrt{\frac{v^n_\beta}{v^m_\alpha}} \cdot A^{nm}_{\beta\alpha}\,, \tag{13}$$

with $A^{nm}_{\beta\alpha}$ the scattering amplitude of mode β in lead n originating from the incoming mode α in lead m. Note that $t^{mm}_{\beta\alpha}$ are the modified amplitudes for *reflection* back into lead m.

The full t-matrix, including reflections, and with labels running over all (α, m), has the following symmetries in a magnetic field B,

$$\begin{aligned} t(B)t^\dagger(B) = t^\dagger(B)t(B) = I\,; &\quad \text{particle conservation} \\ t^*(-B)t(B) = I\,; &\quad \text{time reflection invariance} \\ t(-B) = t^T(B)\,; &\quad \text{(Onsager type symmetry)}. \end{aligned} \tag{14}$$

Here t^T stands for the transposed matrix, t^* for the complex conjugate, $t^\dagger = t^{T*}$, and I is the unit matrix. Note that the third line in (14) is a consequence of the first two.

Now define $T_{nm} = \sum\limits_{\alpha\beta} |t^{nm}_{\beta\alpha}|^2$, which sums all the probabilities for transmission from every incoming mode in lead m to every outgoing mode in lead

n. In terms of these quantities, the Landauer formula (11) generalizes to the Landauer–Büttiker equations [3],

$$I_m = \frac{2e^2}{h}\left(\sum_n{}' T_{nm}V_m - \sum_n{}' T_{mn}V_n\right) \; ; \; m = 1, 2, ...M \, . \qquad (15)$$

Here I_m is the net current into lead m, V_n the voltage of the reservoir connected to lead n, M is the number of reservoirs, and the prime on the sums indicates that the term $n = m$ should be excluded from the sum. Remarks:

— By particle conservation, one must have $\sum_m I_m = 0$. This requirement is automatically satisfied. *I.e.*, one of the equations (15) is redundant.
— One of the voltages can arbitrarily be set to zero, since only voltage *differences* count. This is again automatically satisfied: With all voltages set equal, all currents vanish.
— The symmetry relations (15) imply that $T_{nm}(-B) = T_{mn}(B)$.

From the symmetry of the T-matrices, Büttiker deduced [3] the following Onsager symmetries for multi-terminal resistances,

$$R_{mn,pq}(-B) = R_{pq,mn}(B) \, . \qquad (16)$$

By definition $R_{mn,pq}$ is the resistance (*i.e.*, ratio of voltage to current) when the current flows from lead m to lead n, and the voltage is measured between the reservoirs connected to leads p and q. As a special case it follows that the two-terminal resistance, where $p = m, q = n$, is symmetric under reversal of the magnetic field. The two-terminal resistance must, by the second law, be positive. When voltage and current probes are crossed in a 4-terminal measurement there is, in contrast, no fundamental reason why the corresponding 4-terminal resistance should be positive (the example of Hall resistances demonstrates this).

As an illustration, the episode almost 10 years ago leading to the derivation of (16), should be of interest to the statistical physicist. In a series of clever experiments to uncover Aharonov-Bohm oscillations in mesoscopic metal rings, Webb *et al.* [4] considered the configuration shown in Fig. 2.

The resistance of a ring, with diameter $\sim$ 800nm and connecting reservoirs 1 and 2, was measured, at $T = 0.01$K, as a function of the (perpendicular) magnetic field B. The current was measured from 1 to 2, and the voltage across the ring from a to b. The resistance $R = (V_a - V_b)/I$ plotted as a function of B showed characteristic Aharonov-Bohm oscillations with a period corresponding to the flux quantum h/e. In addition, Webb *et al.* found to their surprise that $R(-B) \neq R(B)$! What happened to the Onsager symmetry? Büttiker solved this puzzle by pointing out that the

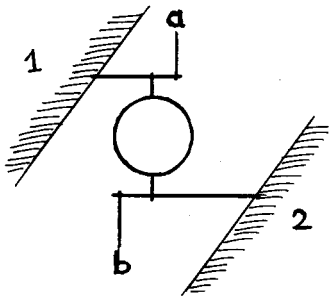

Fig. 2. The metal ring configuration of Webb *et al.*

experimental setup is not a 2-terminal, but a 4-terminal measurement. The resistance $R_{12,ab}$ should have an antisymmetric as well as a symmetric part. One can disentangle these parts by combining different measurements with current and voltage probes in different positions. Subsequent experiments [5] beautifully confirmed Büttiker's explanation.

3. Point contacts

A generic illustration of mesoscopic transport is provided by the quantum point contact (QPC). A narrow constriction connects two wide regions in the 2DEG, these regions being sufficiently wide to be considered as reservoirs (see Fig. 3).

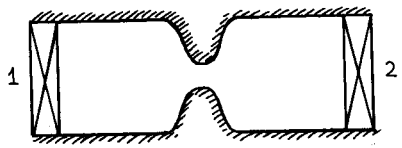

Fig. 3. A simple point contact between reservoirs 1 and 2.

Clearly the energy spacing of the transverse modes is much larger in the narrow constriction than elsewhere. With a given energy E_F only a small number $N(E_F)$ of transverse modes can be accomodated in the constriction or, to put it differently, only a small number of channels are open. Neglecting tunneling and internal reflections we find from the Landauer formula (10) that the conductance through the constriction is the staircase function $G = (2e^2/h)N(E_F)$. Experimentally one finds [6] something qualitatively similar, as shown in Fig. 4.

This simplest of explanations leaves something to be desired. One would like to understand what determines the slope between the steps or, equivalently, the width of the conductance steps. The tendency to oscillatory

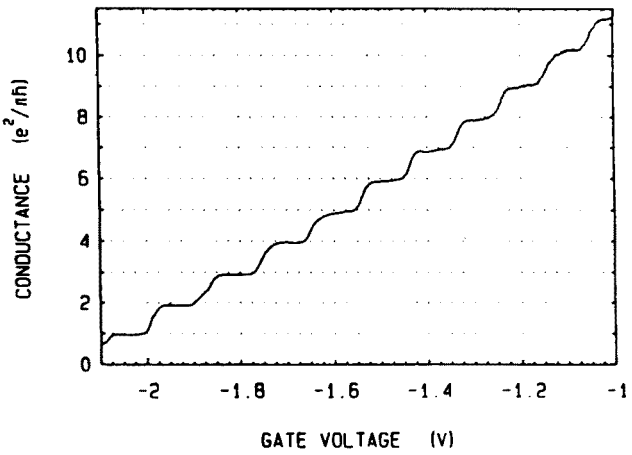

Fig. 4. The conductance as a function of gate voltage, or better, as a function of $E_F - E_0$, where E_0 is linear in the gate voltage, according to van Wees *et al.*

behavior occasionally found, should also be explained. In short, one would like to dig underneath the surface of the simple phenomenological treatment of 'full transmission or no transmission'. The Landauer formula gives the recipe: Solve the scattering problem, calculate the transmisson amplitudes. We shall not pause here to go into this kind of detail [7].

What we would like to do is to emphasize the dual value of the Landauer formula. On the one hand it gives, essentially without calculations, an outline of the principle features of the phenomenon at hand. On the other, it provides the framework within which real calculations can be performed and more detailed questions addressed.

4. Magnetotransport

In the context of magnetotransport, generalization of the Landauer formula to the Landauer–Büttiker equations becomes essential. The dual function survives: With transition probabilities estimated as 0 or 1, the formalism provides an overview of the physics. A case in point is the integer quantum Hall effect (IQHE) in mesoscopic systems which, as we shall see, can be simply understood in this fashion, with direct links between the theoretical formulation and the experimental setup. On the other hand, the formalism provides a framework within which (now quite complicated) scattering calculations can be performed from which the widths of the quantization plateaus etc. can be obtained for the specific case of interest.

Before sketching this, let us consider a simple quantum state from which some basic features of magnetotransport can be understood.

4.1. A simple QM magnetotransport state

Take a quantum wire with harmonic confinement in the transverse y-direction, subject to a magnetic field in the z-direction. The Hamiltonian reads,

$$H = \frac{(\vec{p} + e\vec{A})^2}{2m} + \tfrac{1}{2}m\omega_0^2 y^2 \,. \tag{17}$$

With the Landau gauge for the vector potential $\vec{A} = (-yB, 0, 0)$, the problem is translationally invariant in the x-direction, and the solution of the Schrödinger equation must have the form

$$\psi_{k,\alpha}(x, y) = e^{ikx}\phi_\alpha(y) \,. \tag{18}$$

This ansatz gives a harmonic oscillator problem for the transverse eigenfunctions ϕ_α, with the spectrum,

$$E_\alpha(k) = \hbar\Omega(\alpha - \tfrac{1}{2}) + \frac{\omega_0^2}{\Omega^2}\frac{\hbar^2 k^2}{2m}; \qquad \alpha = 1, 2, 3\cdots, \tag{19}$$

where $\Omega^2 = \omega_0^2 + \omega^2$, and $\omega = eB/m$. The corresponding eigenfunctions are scaled and shifted versions of the standard harmonic oscillator eigenfunctions $\chi_\alpha(\xi)$, namely

$$\phi_\alpha(y) = \chi_\alpha\left(\frac{y - \lambda_B^2 k}{\lambda_B}\right), \tag{20}$$

where $\lambda_B = (\omega/\Omega)l_B$ and where $l_B = \sqrt{\hbar/eB}$ is the magnetic length.

The group velocity in subband α is,

$$v_\alpha = \frac{1}{\hbar}\frac{dE_\alpha(k)}{dk} = \frac{\omega_0^2}{\Omega^2}\frac{\hbar k_\alpha}{m} \overset{\omega_0 \to 0}{\longrightarrow} 0 \,. \tag{21}$$

In (21) k_α is given by the solution of the equation $E_F = E_\alpha(k_\alpha)$. Note that without confinement, there is no transport. Without an external potential the electrons will circulate in states corresponding to cyclotron orbits and they will get nowhere. The presence of some kind of transverse force, here from the parabolic confinement, is necessary to get a non-zero group velocity.

From the form (18) and (20) of the state $\psi_{k,\alpha}(x, y)$ we can, since $\chi_\alpha^2(\xi)$ is symmetric in ξ for all α, read off the center of gravity $\langle y\rangle_\alpha$,

$$\langle y\rangle_\alpha = \lambda_B^2 k_\alpha = \frac{\omega^2}{\Omega^2} l_B^2 k_\alpha = \frac{\omega}{\omega_0^2} v_\alpha \,. \tag{22}$$

These results are illustrated in Fig. 5.

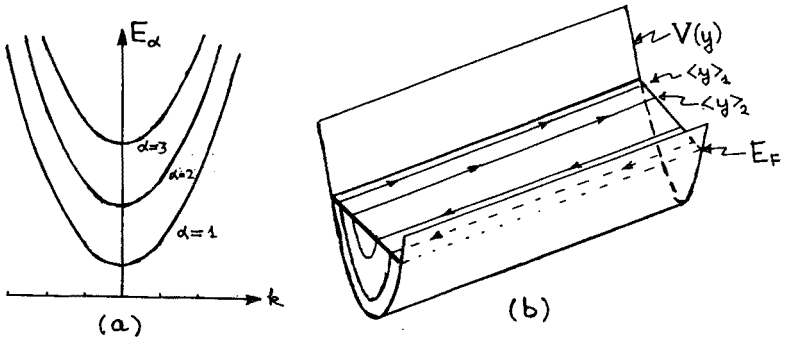

Fig. 5. (a) — The dispersion relations of the Landau subbands. (b) — The corresponding mean positions of the propagating states at E_F.

With $\langle y\rangle_\alpha$ proportional to k_α, which in turn follows from $E_F = E_\alpha(k_\alpha)$, the states propagating in the positive x-direction are ordered so that $\alpha = 1$ gives the state with the highest $\langle y\rangle_\alpha$ (*i.e.*, "closest to the upper edge"), with $\langle y\rangle_\alpha$ smaller the higher α. The states propagating in the negative x-direction are similarly ordered at the lower edge. The number of these pairs of propagating states is equal to the number of Landau levels where $V = 0$, *i.e.*, in the bulk.

The Ehrenfest theorem shows that with harmonic potentials the correspondence between quantum and classical physics is particularly close. It is interesting to check just how close this correspondence is here. For example, the current density (which has both signs!) associated with a single quantum state, mimics the current associated with classical trajectories corresponding to cyclotron orbits drifting in the x-direction, when these trajectories, at the same energy, are averaged over initial positions.

The details of the states discussed above are, of course, special to the idealized model considered. However, many of the qualitative aspects are representative for magnetotransport in realistic quantum wires: In high magnetic fields edge states are responsible for the transport, their number (in each direction) is equal to the number of bulk Landau levels, and they are ordered in space with the state corresponding to the lowest Landau level closest to the edge.

4.2. Adiabatic transport in strong B-fields

The lateral extent of the wavefunction in a propagating state is, as shown in the previous subsection, of the order of the magnetic length, $l_B = \sqrt{\hbar/eB}$. When this length becomes small on the scale set by the

geometry of the system (*i.e.*, by $V(x, y)$), on enters the regime of adiabatic magnetotransport. Characteristic for this regime is that the overlap between different propagating states decreases exponentially with increasing B, so that scattering between states is suppressed. In particular, the probability for scattering across the the entire device, from a forward to a backward propagating state, becomes very small. Let us consider some qualitative consequences of these facts.

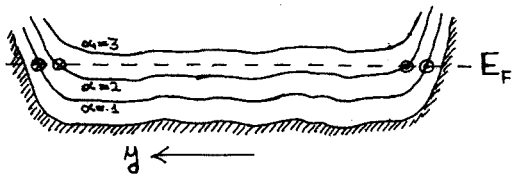

Fig. 6. Crossection of a typical quantum wire.

Fig. 6 shows the crossection of a quantum wire[4] with a transverse confinement potential $V(y)$. The idealized version of this potential is approximately parabolic close to the edges, and flat in the middle. Deviations from this idealization depend, of course, on the circumstances. Let us now make the not unreasonable assumption that these deviations can be described by a random set of low-k Fourier components with small amplitudes. One can think of this stochastic potential as resulting from the randomly distributed donors in a layer parallel to the 2DEG, but respectfully removed from it.

The simplest possible picture is then that of Fig. 6: The equally spaced Landau levels dutifully follow the gently undulating $V(x, y)$, which locally defines the zero point of energy. The lateral compression of the Landau states close to the edges further contributes to raising their energies there. Thus every Landau level, in this generalized sense, will intersect the Fermi energy at some point. The lowest subband will intersect closest to the edge. The number of propagating "edge states" is, in each direction, clearly equal to the number of bulk Landau levels below E_F.

First think of the two-terminal conductance, as given by the Landauer formula (10), for such a quantum wire. When the Fermi energy lies in the gap between two bulk Landau levels, one has a definite number $N(E_F)$ of states propagating in the positive x-direction, close to the upper edge. The same number of states, propagating in the opposite direction is found close to the lower edge. With the width of the device much larger than the magnetic length, $W \gg l_B$, scattering from forward to backward moving states is highly improbable.

---

[4] A wide wire this time. There is no conceptual distinction between a wire and a wider bar.

When, on the other hand, The Fermi level cuts through an undulating Landau level, the situation is qualitatively different. The last "edge state" is now all over the volume and, in addition, there are many current carrying localized states trapped around local hills or in local valleys in the landscape $V(x, y)$. Scattering from the forward, via localized states in the bulk, to the corresponding backward propagating state is now quite likely. A quantitative treatment of this scenario will provide predictions for the *transition* region between two quantized values of G. For a small system the details of these steps will depend on the particular realization of the random potential. In mesoscopic systems, therefore, one should expect the details of $G(E_F)$ or $G(B)$ in the transition regions to be *non-universal.*

4.3. The integer quantum Hall effect

The qualitative features of adiabatic transport in high B-fields are sufficient to provide an explanation [8] of the integer quantum Hall effect

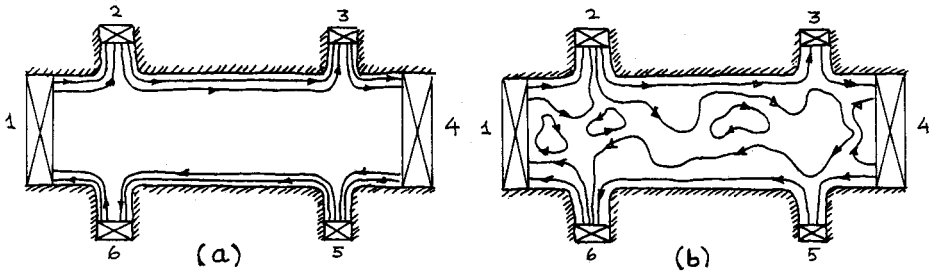

Fig. 7. A Hall bar with propagating states at the Fermi level (a) when E_F lies in the gap between Landau levels, and (b) when E_F cuts through an undulating Landau level.

(IQHE). Fig. 7 shows a Hall bar in a 6-terminal configuration typical for such measurements. The net current enters from reservoir 1 (the source) and ends up in reservoir 4 (the drain). *I.e.*, $I_1 = -I_4 = I$. The remaining four "reservoirs" are voltage probes with vanishing net current. At the high B-fields contemplated here, one should treat the different spin states as separate channels. With this larger state space explicitly taken into account, the spin degeneracy factor 2 in the Landauer–Büttiker equations (15) should be replaced by unity. In Fig. 7(a), E_F lies in the gap between spin split Landau levels, whereas in Fig. 7(b) E_F cuts through an undulating level.

From Fig. 7(a) one can immediately write down excellent approximations for the T_{nm} needed in the Landauer–Büttiker equations. With N spin

split states propagating along the edges of the bar one has,

$$T_{nm} = \begin{cases} N & \text{if } (n,m) = (m+1,m) \\ 0 & \text{otherwise} \end{cases} . \tag{23}$$

Let us define the Hall conductance as $G_H = I/(V_2 - V_6)$ (or any other combination with one probe along the upper, and one along the lower edge). Then (15) with $2e^2/h$ replaced by e^2/h gives,

$$G_H = R_H^{-1} = \frac{I}{V_2 - V_6} = \frac{e^2}{h} N . \tag{24}$$

Similarly, from (15) and (23), the longitudinal resistance R_L vanishes,

$$R_L = \frac{V_2 - V_3}{I} = 0 .$$

This simple picture immediately shows that when E_F lies in the gap between Landau levels, the Hall conductance is quantized and given by the number of spin-split channels propagating along the edges. It also directly correlates this quantization with a vanishing longitudinal resistance.

The transition between two levels, say $N \to N - 1$, is caused by E_F moving through the Landau level no. N from above or, equivalently, by the Landau level moving through E_F from below. The latter case corresponds to B being increased. Thus, a simplistic interpretation of (24) yields G_H as a descending staircase function for increasing B. On can go beyond this zeroth approximation and study the shape of the steps by detailed calculations based on the scenario sketched in Fig. 7.

Note that in this mesoscopic setting a random potential with localized states is *not* required to explain the finite width of the plateaus. The reason for this is basically that the chemical potential is fixed, not by the mesoscopic Hall bar itself, but by the ("infinite") reservoirs. This is in contrast to the essentially macroscopic Hall bars on which, for example, von Klitzing's original experiments were done.

On the other hand, by reintroducing the gently undulating potential $V(x,y)$, one can approach the macroscopic limit by considering scattering across the width W, from forward to backward propagating states. This leads, in (extremely) high magnetic fields, to an interesting type of percolation problem in which electrons essentially drift along equipotential lines in the undulating landscape. This percolation problem is, presumably, characterized by universal exponents as $W \to \infty$. The correlation length exponent ν has, in fact, been determined experimentally to $\nu = 2.3 \pm 0.1$, consistent with $\nu = \frac{4}{3} + 1$. The exact result for classical percolation in 2D is $\nu_{cl} = \frac{4}{3}$, so the extra contribution of unity must be of quantum origin. What precisely

does QM do to the percolation problem? Explanations relying on tunneling alone [9], and on tunneling with crucial contributions from interference [10], have both been put forward. The final verdict is, in my opinion, not yet in.

5. A classical model

With all the current activity in the field of quantum transport, it is interesting that there are still surprises in store in *classical* magnetotransport. The good old Lorentz model, in which a single electron moves with constant speed v in a random array of short range scatterers, can be studied in 2D and with a perpendicular magnetic field. Amazingly enough, the Boltzmann equation is *not* correct in this case, even in the Grad limit, $n \to \infty, a \to 0, \Lambda = (2an)^{-1} =$ finite, where n is the number density of hard disk scatterers with radius a, and Λ is the mean free path. The physical reason for the failure of the Boltzmann equation is simple, and *special to* 2D. In a magnetic field classical electrons move in circular orbits with the cyclotron radius R. Thus, in 2D, there is finite probability that the electron will complete the circle without scattering. In the Grad limit this probability is $P_0 = \mathrm{e}^{-2\pi R/\Lambda}$. The recurrence has the effect that some electrons keep circling in a cyclotron orbit forever, whereas others suffer many recollisions with the same scatterer before hitting a different one. Thus, non-Markovian effects are introduced and the Stosszahlansatz, basic to the standard Boltzmann equation, fails.

At first sight the corresponding complications look formidable, but as the smoke clears [11], the generalized Boltzmann equation appears (for the spatially homogeneous case) in the form,

$$\left(\frac{\partial}{\partial t} + \omega \frac{\partial}{\partial \phi}\right) f(\phi, t) = \sum_{s=0}^{[t/T]} P_0^s B_{-sT} f(\phi, t - sT)\,, \tag{26}$$

with the retarded Boltzmann operator defined by,

$$B_{-sT} f(\phi, t - sT) = \nu \int_{-\pi}^{\pi} d\psi g(\psi) \left[f(\phi - (s+1)\psi, t - sT) - f(\phi - s\psi, t - sT)\right]. \tag{27}$$

Here ϕ defines the direction of the velocity, ω is the cyclotron frequency, $T = 2\pi/\omega$ the corresponding period, ν the collision frequency and $g(\psi)$ the dimensionless differential crossection. This generalized Boltzmann equation is, properly interpreted, *exact* in the Grad limit. The corresponding initial value problem can even be solved exactly, leading to a diffusion tensor D_{ij}

given by the complex $\mathcal{D} = D_1 + iD_2$, with $D_1 = D_{xx} = D_{yy}$ and $D_2 = -D_{xy} = D_{yx}$, as

$$\mathcal{D} = \frac{\frac{1}{2}v^2\tau_D(x)(1-x^2)}{1-i\omega\tau_D(x)}, \tag{28}$$

where $x = \exp(-\frac{1}{2}\nu T) = \exp(-\pi\nu m/eB)$ and

$$\tau_D^{-1}(x) = \nu\left[1 - \frac{1-x^2}{2x^2}\left(\frac{1-x^2}{2x}\ln\frac{1+x}{1-x}a - 1\right)\right]. \tag{29}$$

Now move away from the Grad limit, but with $na^2 \ll 1, a/R \ll 1$. A new parameter of arbitrary size can then be defined as $r \equiv R\sqrt{n} = R/\sqrt{2\Lambda a}$. In the Grad limit, $r \to \infty$. For finite r a percolation problem arises. Below the percolation threshold $r_p = 0.5995 \pm 0.005$, no diffusion is possible, the electrons are trapped in localized clusters of scatterers. A numerical study [11] of the corresponding percolation exponents gives results that are *almost* consistent with exactly known values for lattice percolation. More work is needed.

I could not refrain from quoting these tantalizing results here, but do not give further details since our focus is on *quantum* transport in the 2DEG. Whether the classical Lorentz model provides a reasonable *caricature* of some remote region in this basically quantum world (a random array of quantum antidots with judiciously chosen parameters comes to mind) remains to be seen. As a frivolous fling into classical kinetic theory, study of the Lorentz model certainly has its own rewards!

6. Final remarks

Within the limited scope of these lectures it has been impossible to do justice to the variety of phenomena and problems encountered in the study of the 2DEG. A couple of final pointers:

We have stayed strictly within the single-electron picture here. With interactions basically of Coulomb character, this has its limitations. For example, a number of subtle effects from Poisson self-consistency add important nuance to the simple one-electron pictures sketched here, the character of edge states being a case in point. Moreover, the whole vigorous subfield of Coulomb blockade effects [12] (of both fundamental and (potentially) applied interest!), hinges on taking the Coulomb interaction seriously. The crucial point is that, for sufficiently small tunnel junctions, the effective capacitance C can be made so small that the charging energy of a single electron, $e^2/2C$, becomes a dominant energy (for example, $e^2/2C \gg k_BT$). The charging energy is clearly a nonlinear effect, basically of many-body origin. However, simple approximate theories based on classical electrostatics

have been remarkably successful in describing the main features of Coulomb blockade phenomena. Subtle effects of the environment, on the other hand, require more elaborate methods.

In the mesoscopic world of coherent transport, described in these lectures, the integer quantum Hall effect is almost reduced to a triviality. However, the IQHE is also observed *outside* the realm of mesoscopics. In fact, von Klitzing's original measurements were done on an essentially macroscopic Hall bar. A whole body of literature on the QHE has macroscopic systems with incoherent transport in mind. The connections between these two views are only partly understood. Similar remarks apply to the fractional quantum Hall effect, not covered in these lectures. Ideally, one would like to start from a unified theory which, from the outset, can encompass both interactions and edge effects, and from which mesoscopic and macroscopic theories on the integer and fractional Hall effects follow as special cases. Also, current distributions, dissipation, 'critical' exponents, and limitations of the linear theory should all follow from this grand unified viewpoint. We are not yet there.

REFERENCES

[1] C.W.J. Beenakker, H. van Houten in *Solid State Physics: Advances in Research and Applications*, Vol. 44, edited by H. Ehrenreich and D. Turnbull, Academic Press, Boston 1991.

[2] R. Landauer, *IBM J. Res. Dev.* **1**, 223 (1957).

[3] M. Büttiker, *Phys. Rev. Lett.* **57**, 1761 (1986).

[4] R.A. Webb, S. Washburn, C.P. Umbach and R.B. Laibowitz, *Phys. Rev. Lett.* **54**, 2996 (1985).

[5] A.D.Benoit, S. Washburn, C.P. Umbach, R.B. Laibowitz, R.A. Webb, *Phys. Rev. Lett.* **57**, 1765 (1986).

[6] B.J. van Wees, H. van Houten, C.W.J. Beenakker, J.G. Williamson, L.P. Kouwenhoven, D. van der Marel, C.T. Foxon, *Phys. Rev. Lett.* **60**, 848 (1988).

[7] The literature on point contacts is large. For orientation, see A. Yacoby, Y. Imry, *Phys. Rev.* **B41**, 5341 (1990); Frank A. Maaø, I.V. Zozulenko, E.H. Hauge, *Phys. Rev.* **B50**, 17320 (1994).

[8] M. Büttiker, *Phys. Rev.* **B38**, 9375 (1988).

[9] G.V. Mil'nikov, M. Sokolov, *Pis'ma Zh. Eksp. Teor. Fiz.* **48**, 494 (1988) [*JETP Lett.* **48**, 536 (1988)]; Alex Hansen, C.A. Lütken, *Phys. Rev.* **B51**, 5566 (1995).

[10] J.T.Chalker, P.D.Coddington, *J.Phys.C* **21**, 2665 (1988); Dung-Hai Lee, Ziqiang Wang, Steven Kivelson, *Phys. Rev. Lett.* **70**, 4130 (1993).

[11] A.V. Bobylev, Frank A. Maaø, Alex Hansen, E.H. Hauge, to be published.

[12] *Single Charge Tunneling, Coulomb Blockade Phenomena in Nanostructures*, edited by H. Grabert and M.B. Devoret, NATO ASI, Series B, Physics, Vol. 294, Plenum, New York 1992.