

Hall Effect and Resistivity in a two-dimensional electron gas

(Written and updated by M. Gurvitch, 1-30-2001; 11-4-2002; 09-23-2011)

1. Introduction and History

Hall Effect is observed when electrical current in a sample flows perpendicular to an applied magnetic field. This results in the appearance of transverse voltage in the direction perpendicular to both the current and the magnetic field. This voltage, which is proportional both to the current and the field, can be easily measured with appropriate probe arrangement. This simple effect was discovered by Edwin Hall in 1879 [1], 18 years before the discovery of an electron. At the time of the discovery Hall was a graduate student in Johns Hopkins University in Maryland (on the photograph below he is much older, a Harvard professor in his later life).



Edwin Hall (1855-1938)

The effect can be easily understood on the basis of elementary physics of moving electric charge in the crossed E and B fields (Sec. 3). Its significance in the Physics of solids stems from the fact that it allows direct determination of the carrier concentration and, coupled with resistivity measurement, of carrier mobility. In many solids the situation is somewhat complicated by the presence of two types of carriers, electrons and holes (see below), but still these two measurements allow determination of concentrations and mobilities.

The familiar Hall Effect measurement, when performed at low temperatures on certain types of semiconductors containing *two-dimensional electron gas* (abbreviated as *2DEG*), quite unexpectedly led in the late 70th - early 80th to the discovery of remarkable *Quantum Hall Effects (QHE)*. So far, two Nobel Prizes were awarded for *integral* [2] and *fractional* [3] QHE discoveries: in 1985 to Klaus von Klitzing (Germany) and in 1998 to Horst Stormer, Daniel Tsui and Robert Laughlin (USA, Bell Labs and Lawrence-Livermore National Lab). From 1990, QHE is the basis of the official resistance standard (*Quantum Ohm*, or 1 *klitzing*) maintained in the US by the National Institute for Standards and Technology (NIST). Its value of $R_K = h/e^2 = 25812.807557(18) \Omega$ is independent of material parameters within an uncertainty of better than 1 part per billion, 10^{-9} .

The discovery of these and other interesting effects [4] could not be made without a well-developed technology of semiconductor samples containing high mobility 2DEG. One way of making such samples is called *Molecular Beam Epitaxy (MBE)*; it was invented in the early 1970-s at Bell Labs by J. R. Arthur and Alfred Y. Cho [5] and further developed by A. Cho and many other workers. A modern MBE deposition system is a sophisticated (and very expensive – typically costing millions of dollars) ultra-high vacuum “mini-factory” producing perfectly ordered atomic layers of various materials; most notably, of III-V semiconductors. The roman numerals III and V stand for the Groups in the Periodic System of Elements; for example GaAs is a III-V semiconductor. One very popular system consists of AlGaAs and GaAs layers. Such systems are also called *heterostructures* or *superlattices*.

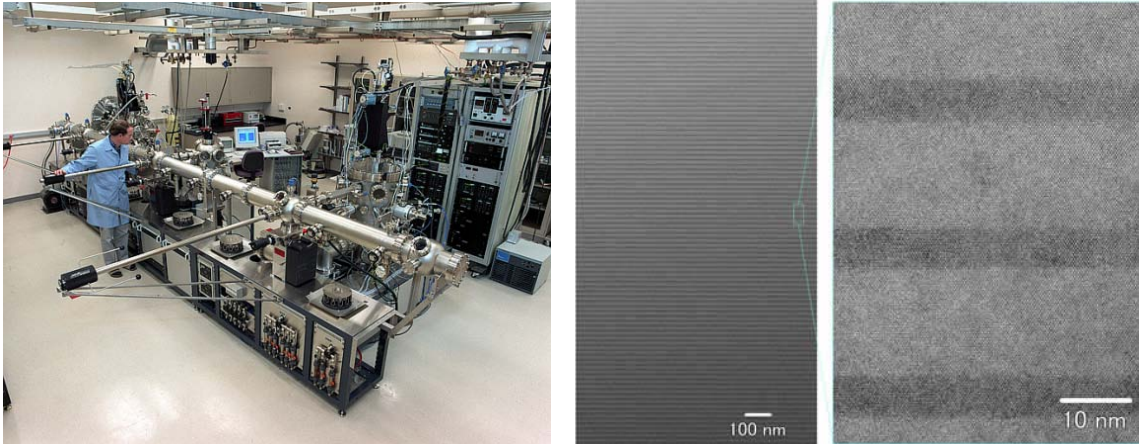


Fig. 1. An MBE system and electron microscope images of a III-V semiconductor superlattice; note that in the last high-resolution image on the right side one can just depict a number of individual atomic planes within each layer.

Such a *superlattice* is grown so that each subsequent layer (atomic plane) is in a crystallographic reference to an underlying layer (which is the meaning of the word *epitaxy*), so that each atom is at its proper place in the structure, forming a layered artificial single crystal [6]. Some layers or interfaces between layers are intended as 2-D conductors (for example, an interface between GaAs and a wider band gap GaAlAs), while others as doping layers or insulators (for example, AlGaAs is an insulator). The physical separation of donors in doping layers from the mobile carriers in the conducting layers (called *modulation doping*) is a clever idea, which allows obtaining high mobility 2DEG [7]. In today's semiconductor electronics most of the novel electronic device ideas stem from this technological ability to layer-by-layer engineer the most intricate semiconductor materials. Of course, since the early 1970s, a vast literature (books, articles, conference proceedings, etc.) has been written on all these subjects.

In this Lab you will indeed work with such samples containing 2DEG. Sample description is given in Section 13. The samples were prepared by MBE at Stony Brook, by Bent Nielsen, in Prof. Emilio Mendez's Lab. Such samples represent cutting-edge technology, and we hope that you will learn about this technology and the peculiar physics of 2DEG by reading the provided references, in addition to learning about Hall Effect and resistivity measurements.

Finally, you will use the *four-probe* technique for the measurement of your sample's resistivity (or, in case of a 2DEG, of resistance per square). Understanding four-probe measurement is essential in experimental Solid State Physics.

The main purpose of this experiment is to learn about standard DC resistivity and Hall Effect measurements, including the 4-probe method, and, from these measurements performed at 300 K and at 77 K on two different heterostructure AlGaAs/GaAs samples, to determine the following macroscopic and microscopic parameters of the two samples:

- a) The resistance per square
- b) The Hall coefficient
- c) The carrier sign (electrons or holes?)
- d) The carrier density
- e) The mobility
- f) The carrier relaxation time (mean time between collisions)
- g) The *mean free path* (mfp) of the carriers (mean distance traveled between collisions).

Additionally, we will ask you to discuss various relationships between these microscopic parameters and parameters of the sample, extracting, hopefully, some Physics understanding from your measurements.

2. Carriers in E field: Drude Model, Ohm's Law; mobility

Let us start by reviewing a simple description of electrical conduction in the electric field known as Drude Model, introduced by Paul Drude in 1900; we will add magnetic field to it and discuss Hall Effect in Section 3.

Consider a macroscopic solid conductor containing a very large number of charged particles which are free to move around this solid; for now, consider them to be electrons. These *free electrons*, moving around with their *intrinsic* velocities, can be thought of as forming *electron gas* or *electron cloud* inside the conductor. The negatively charged electron gas is moving amongst stationary positively charged ions (the lattice), and, of course, the solid is on the whole neutral.

Classical Physics (and Paul Drude in 1900) naturally assumed that in an electron gas, as in any gas, these intrinsic velocities would be thermal in nature, so that kinetic energy of each free electron would be $mv^2/2 = 3/2 k_B T$, where T is absolute temperature. Quantum Physics discovered that, because of the Pauli Exclusion Principle, their intrinsic velocities are much higher than that, being in metals and highly-doped (*degenerate*) semiconductors close to the maximum velocity of electrons in an electron gas at the absolute zero: Fermi velocity v_F [8].

In the absence of electric field, electron's motion is random, meaning that if we, at any instant of time, will average all their velocities, we will find zero. In other words, in zero field $E = 0$, the center of mass of electron gas must be at rest. While moving inside the conductor, electrons are scattering off structural imperfections, impurities, thermal vibrations of the atoms, or off each other. Let the average time between collisions, called *relaxation time*, be τ .

Suppose we turn on a uniform electric field \mathbf{E} in x direction inside the conductor. Now each electron experiences a force in the $-x$ direction, and therefore individual electron's trajectories between collisions will change, and electron cloud as a whole will start drifting in the direction opposite to the direction of that field. In a constant E field, were it not for those scattering events, the velocity of electron cloud's drift would keep increasing indefinitely, with constant acceleration $a = eE/m$. However, experiment tells us that this does not happen. Indeed, if this were the case, the current would also keep increasing, which is not observed. This can only mean one thing: an electron cloud experiences a drag (or friction) as it moves through the solid. In other words, those collisions we mentioned above, on the whole must be *inelastic*, serving to remove the additional momentum and energy which were acquired by the electron gas in the field. Instead of accelerating indefinitely, electrons acquire some finite, constant "terminal" velocity, which is called *drift velocity*, v_d . The energy transferred to the ionic lattice produces Joule heat. Obviously, if now, in case of a non-zero electric field, we will again average all electronic velocities, instead of zero we will find v_d ; all the random intrinsic velocities – whether classical or quantum -- will still cancel out.

Let us assume that each collision removes that extra acquired momentum in the direction of \mathbf{E} , and see what this simple picture predicts. Following elementary textbooks, we will write $v_d = a\tau$, where $a = eE/m$ [Note however, that $(a\tau)$ is the maximum velocity acquired in time τ ; for the average velocity v_d we should be really writing $v_d = (1/2)a\tau$ here, which will introduce an incorrect factor of $1/2$ into the final result. See Ashcroft and Mermin in [8] and also [9] for a more sophisticated treatment of the distribution of times between collisions which removes this factor]. The current density $j = env_d$, where n is electron concentration, and thus we obtain a familiar relation

$$j = E/\rho \quad (1)$$

This is the *Drude relation*, which also explains Ohm's Law, see below. Here we introduced *resistivity*:

$$\rho = m^*/(e^2 n \tau) \quad (2)$$

In the last formula, instead of using free electron mass m , we wrote m^* , which is the effective mass of electrons inside a solid.

The concept of an effective mass is developed in the *band theory* of solids [8]. In essence, the free electron mass is modified when electron moves through the periodic potential in a lattice; m^* can be smaller or larger than free electron mass $m = 9.1 \times 10^{-31}$ kg. For example, in GaAs electronic effective mass is only 0.067 m (6.7% of the free electron mass). Effective mass m^* may also depend on the direction in the crystal lattice.

The inverse of resistivity is called conductivity σ , and formula $\sigma = e^2 n \tau / m^*$ is called *Drude formula*.

Let us now introduce a quantity much used in the physics of electrical conduction, especially when discussing semiconductors. It is called carrier *mobility* and defined as

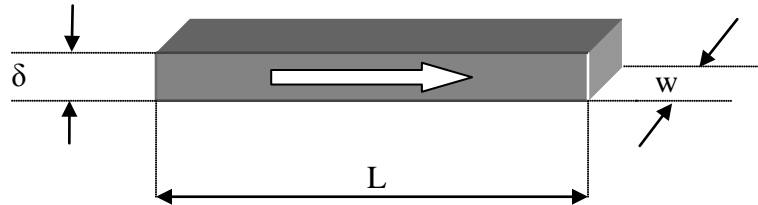
$$\mu = v_d / E \quad (3)$$

As can be seen, "mobility" is indeed an appropriate name for this quantity: it is the magnitude of drift velocity per unit electric field. From (1), (2) and $j = en v_d$ we see that, expressed through microscopic quantities, mobility is

$$\mu = e \tau / m^* \quad (4)$$

The unit of mobility in SI is m^2/Vs , although it is more common to see it given in cm^2/Vs (i.e. 10^4 times larger number). For example, at room temperature (300 K) in pure Si, mobility of electrons $\mu_e = 0.14 m^2/Vs = 1.4 \times 10^3 cm^2/Vs$, and that of holes $\mu_p = 0.045 m^2/Vs = 450 cm^2/Vs$. High mobility is desirable in many experiments with bulk (3-D) semiconductors and in most experiments with 2DEG semiconductors, since many interesting phenomena, such as fractional QHE, escape observation in low mobility samples. High mobility is also beneficial in a number of devices, such as Field Effect and hot-electron transistors. The history of MBE and other technologies aiming at creating sophisticated semiconductor structures [5-7] is, in some sense, a history of increasing electron mobility. In today's III-V 2DEG, at low temperatures, mobility can exceed $3 \times 10^6 cm^2/Vs$ [9]. It is interesting that a recently-discovered atomically-thin (i.e. essentially 2-D) carbon allotrope *graphene* exhibits room-temperature mobility in excess of $1.5 \times 10^4 cm^2/Vs$ [10], which is 10 times higher than room-temperature mobility in Si.

Let us now recall a relation between resistivity and resistance. For a rectangular sample of length L , width w and thickness δ ,



with current flowing along L , if we multiply both sides of (1) by a sample volume $Lw\delta$, we will get $(jw\delta)L = (EL)w\delta/\rho$, or $I = Vw\delta/\rho$, where I is total current and V is potential drop (voltage) across L . Defining resistance $R = V/I$, we obtain Ohm's Law in the usual "electrician's" form:

$$I = V/R, \text{ where } R = \rho(L/w\delta) \quad (5)$$

The unit of resistance is V/A called *ohm* (also denoted by a capital Greek letter omega: Ω); the unit of resistivity is Ωm .

If we consider resistance of a sample which is shaped as a square of uniform thickness, so that $L = w$, we see from the last formula that it will be

$$R_{\square} = \rho/\delta \quad (6)$$

where we introduced “R per square”. We realize that resistance of a square does not depend on its planar size. R_{\square} is often alternatively called *sheet resistance*, or R_s . We will return to this topic when discussing 2DEG samples in section 7; we will see that in a 2-D case, R_{\square} is the only parameter which is needed to characterize the material, on the par with resistivity in a 3-D case.

3. Carriers in E and B fields; Hall Effect and magnetoresistance for single type carriers

Consider now a charged carrier (charge e will be assumed to be negative for an electron and positive for a hole) moving with velocity \mathbf{v} in simultaneously present electric and magnetic fields \mathbf{E} and \mathbf{B} . It will experience Lorentz force $\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B})$. The electric part of this force, $\mathbf{F}_e = e\mathbf{E}$, will accelerate the carrier in the direction of \mathbf{E} , changing its velocity magnitude, just as was discussed in Section 2, while the magnetic part of the force, $\mathbf{F}_m = e\mathbf{v} \times \mathbf{B}$, being always perpendicular to \mathbf{v} , will change the direction of \mathbf{v} without affecting its magnitude.

Let us again assume, as in section 2, that each collision completely resets a carrier to its initial random state: there is no memory of the direction in which it was accelerated by $e\mathbf{E}$ and turned by $e\mathbf{v} \times \mathbf{B}$, and it loses its additional momentum and drift velocity previously acquired from the field. By Newton’s Second Law, carrier’s linear momentum \mathbf{P} is changing according to the equation $d\mathbf{P}/dt = e(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, so that infinitesimal change during dt is $d\mathbf{P} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B})dt$. If we average the last expression over all the carriers, then we need to replace \mathbf{v} by \mathbf{v}_d : random velocities average to zero and leave only \mathbf{v}_d , as in Section 2. The average momentum change during τ can therefore be written as

$$\Delta\mathbf{P} = e(\mathbf{E} + \mathbf{v}_d \times \mathbf{B})\tau \quad (7)$$

But with the complete “resetting”, the initial drift velocity after each collision is zero, and so $\Delta\mathbf{P} = m^*\mathbf{v}_d - 0 = m^*\mathbf{v}_d$. Then we will get from (7)

$$\mathbf{v}_d = \mu(\mathbf{E} + \mathbf{v}_d \times \mathbf{B}) \quad (8)$$

where we used mobility formula (4). As we said, the $\mathbf{v}_d \times \mathbf{B}$ term acts to turn \mathbf{v}_d , but cannot change the magnitude of \mathbf{v}_d , the latter therefore being entirely due to the action of an electric field \mathbf{E} . Thus, looking at (8), we recognize that mobility can be also written as $\mu = v_d/E$, which is indeed its definition given in Section 2 in formula (3).

Formula (8) is applicable to any orientation of \mathbf{E} and \mathbf{B} . Let us now consider a specific orientation of \mathbf{B} field with respect to a sample. Consider what happens in a conducting layer (slab) of thickness δ and width w , when current I is set up to flow along the length L of the sample, in x -direction, while a uniform magnetic field \mathbf{B} is applied perpendicular to its surface, in z -direction, as shown below:

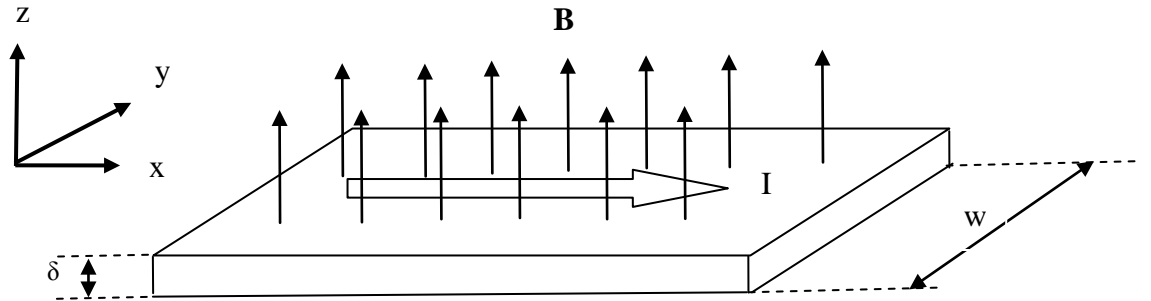


Fig. 2 Schematic view of Hall Effect arrangement for the current and the magnetic field

Taking $\mathbf{B} = B_z\mathbf{k}$ in (8) (magnetic field in the positive z -direction), the vector product $\mathbf{v}_d \times \mathbf{B}$ can be represented as a matrix

$$\mathbf{v_d} \times \mathbf{B} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ v_{dx} & v_{dy} & v_{dz} \\ 0 & 0 & B_z \end{vmatrix}$$

and, therefore, from (8), the results for $\mathbf{v_d}$ in components are

$$v_{dx} = \mu (E_x + v_{dy} B_z);$$

$$v_{dy} = \mu (E_y - v_{dx} B_z);$$

$$v_{dz} = \mu E_z ;$$

Substituting v_{dy} into v_{dx} , requiring that transverse velocity in a steady state is equal to zero: $v_{dy} = 0$ (this condition is equivalent to a requirement that charge does not indefinitely continue to build up at sample edges), and recalling that current density can be expressed as $\mathbf{j_x} = en\mathbf{v_{dx}}$, we will obtain Drude relation (or Ohm's Law) $\mathbf{j_x} = E_x/\rho$, where resistivity is given by formula (2), and we will further obtain something new: a Hall relationship

$$E_y = j_x B_z / ne \quad (9)$$

We recognize that while the initial electric field must have been applied in x-direction to produce the current in x-direction, interaction of that current with the magnetic field works so as to produce electric field component in y-direction.

The fact that, starting from (8), we again obtained Ohm's law, even in the presence of magnetic field, means that for a single type carriers, ρ does not depend on the magnetic field. The ratio of the resistivity change in the magnetic field, $\Delta\rho = \rho(B) - \rho_0$ to the initial, zero field value ρ_0 , $\Delta\rho/\rho_0$, is called *magnetoresistance*. Therefore, we can state that in conductors with single type carriers there is zero magnetoresistance -- a result which we will discuss more in the next section.

Uniform transverse field E_y across a sample of width w will produce transverse Hall voltage drop $V_H = E_y w$, which can be measured (see below). Relating current density j_x to the total current $I = j_x \delta w$, we find from (9)

$$V_H/I = R_H B_z / \delta \quad (10)$$

where

$$R_H = 1/(en) \quad (11)$$

is called *Hall coefficient*. We see that, for one type of carriers, measurement of R_H directly leads to determination of carrier density n . Combined with measurement of resistivity, Hall measurement allows determination of carrier mobility.

The ratio of Hall voltage V_H to the longitudinal current I , V_H/I , is called *Hall resistance* (not to be confused with regular longitudinal resistance, nor with Hall coefficient R_H). As we see, it is proportional to the magnetic field. It is this dependence, Hall resistance vs. B , that was found by Klaus von Klitzing in 1980 to have peculiar flat plateaus determined only by integers times fundamental combination of constants h/e^2 , which marked the discovery of integer Quantum Hall Effect [2].

4. Hall Effect and Magnetoresistance for two types of carriers (electrons and holes) with equal concentrations

Electrical conduction in solids occurs via movement of electrons, holes, or both. Appearance of positively charged *holes* is explained in the band theory of solids [8] -- they resemble bubbles in water in a

gravitational field. While a stone will sink towards the bottom, the bubble will rise towards the surface, against the direction of the gravitational field. The deeper is the bubble, the greater is its potential energy. In reality, only water moves around an empty (gas filled) space called a "bubble", yet a bubble surely looks like an object moving up against the force of gravity. Similarly, electron vacancies may act as if they had a positive charge. As electrons move in the direction opposite to the direction of \mathbf{E} field, they move passed such a vacancy, or "hole", which heads in the opposite direction, resembling a positively charged particle. The deeper is the hole in the energy band, the greater is its potential energy. Electrons and holes may have different effective masses and different mobilities. Typically, electrons have higher mobilities than holes. They may also have different concentrations, the difference made up for by charged impurities. However, in *intrinsic* (impurity-free) semiconductors and in *semimetals*, such as Bi or Sb, the number of free electrons equals the number of holes. We will only consider that case here, and only to site the end result which can be obtained in a way similar to the derivation of Sec. 3. To obtain the results given below, instead of $v_{dy} = 0$, we require that $j_y = 0$ (if you think about it, it makes perfect sense for two types of carriers). Defining electron and hole mobilities μ_n and μ_p with respective m^* 's and τ 's for electrons and holes, we will find

$$\rho = \rho(0) (1 + \mu_n \mu_p B^2) \quad (12)$$

$$R_H = [1/(ne)] (\mu_p - \mu_n) / (\mu_p + \mu_n) \quad (13)$$

So, in this case resistivity is found to be a function of the field B , i.e. now there is non-zero magnetoresistance, which turned out to be proportional to the square of the field. The Hall coefficient is modified by a combination of mobilities. We can still find n and the two mobilities if we measure resistivity, Hall coefficient and magnetoresistance.

In the next section we will give a more intuitive view of some of the relations obtained in Sec. 3 and here.

5. A simplified view of Hall Effect and Magnetoresistance

Some (but not all) of the relations so far described can be very easily understood on the basis of elementary physics, without much math. Indeed, for the case of a single type carrier, either an electron (Fig. 3A) or a hole (Fig. 3B), Lorentz force will initially tend to deviate current to a side of a sample:

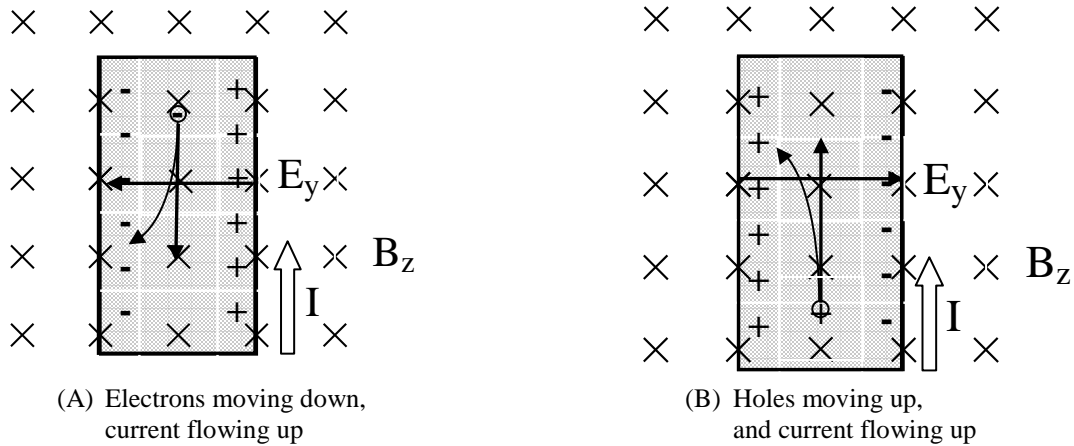


Fig. 3 Hall Effect for electrons and holes

This current-curving action of the magnetic field will lead to building up of an electric charge at the sample edges and, therefore, to a field \mathbf{E}_y across the sample. This build-up of an edge charge will cease once the electric force compensates the magnetic force (just as in a familiar *velocity selector* in E&M):

$$eE_y = ev_d B_z \quad (14)$$

We recognize that from that point on, when the two forces are equal and opposite, all carriers comprising the current will move along the x-direction without curving. Thus the magnetic field influence in a steady state will be completely balanced by an electric force, while E_y will be set up in y-direction. Thus mfp in x-direction will be unchanged, and drift velocity will be directed only along x. We conclude that in this case there is no magnetoresistance.

Multiplying (14) by n , we find $enE_y = (env_d)B_z = j_x B_z$, and with $V_H = E_y w$ and $I = j_x w \delta$, $V_H/I = R_H B_z/\delta$. Thus we obtained our formula (10) very simply, using physical arguments, without considering a vector product or doing any calculations.

Note that the polarity of Hall voltage V_H will be opposite for electrons and holes, allowing clear determination of the carrier type in Hall Effect measurements performed on materials with a single type carrier. Note also that the current direction is defined as the direction of motion of positive charge, which is taken into account in Fig.3.

Finally, if we have both electrons and holes, the current is carried by both species. Suppose further that electrons and holes are in equal concentrations and with equal mobilities, and thus equal drift velocities. In that special case we can see that equal number of electrons and holes will deviate to each side of a sample, and they will compensate each other. The transverse field E_y will not be set up -- there will be no Hall voltage; Hall Effect disappears in this case. However, magnetoresistance does not disappear: all mfp are modified, both for electrons and for holes. This is indeed predicted by formulas (12) and (13). One can even go a bit further and see that indeed resistivity will be increasing as a square of B , as in formula (12). We leave this as an exercise for the reader.

6. Two-dimensional conductors

In this lab you will be measuring two-dimensional conductors -- electrons (or holes) are confined to a quantum well, so that they can only move in two dimensions, the third being quantized. There is no longer a meaningful parameter of thickness δ relating to carrier transport that can be ascribed to the sample. How the results of previous sections change in that case?

Let us first derive Ohm's Law and define a quantity which will replace 3-D resistivity in 2-D. In 2-D, as in 3-D, we still have electric field \mathbf{E} , electron charge e and effective mass m^* , and acceleration in the field will still be $a = eE/m^*$. If average time between collisions (*relaxation* time) is τ , then $v_d = a\tau$ (Note: the same comments about distribution of times between collisions apply here as in Sec. 2). A 2-D current density $j = env_d$, where n is the density of carriers in 2-D, i.e. per meter square, m^{-2} , and j is related to current as $j = I/w$, where w is the sample width, units of j being A/m. Combining our formulas, we therefore obtain $j = (e^2 n \tau / m^*) E$, the familiar Ohm's Law, except for the reduced dimensionalities of n and j . Let us now define a quantity we will call *R per square* and denote as R_{\square} :

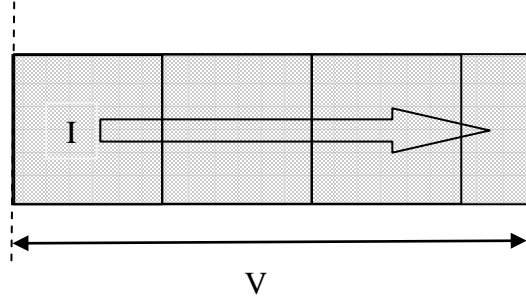
$$R_{\square} = m^* / (e^2 n \tau) \quad (15)$$

We will therefore have $j = E / R_{\square}$, a 2-D form of Ohm's Law.

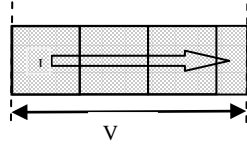
Now let us establish a relation between the last formula and "electrician's" Ohm's Law $I = V/R$. Suppose we have a rectangular 2-D sample of length L and width w , with current flowing along L . Let us multiply both sides of $j = E / R_{\square}$ by the area wL ; we will have $(jw)L = (EL)w/R_{\square}$, or $IL = Vw/R_{\square}$, or $V/I = (L/w)R_{\square}$. But V/I is the regular resistance R ; so we found that $R = (L/w) R_{\square}$. It is clear that L/w represents the number of squares that fit into a rectangular sample; let us call it $\gamma = L/w$; then

$$R = \gamma R_{\square} \quad (16)$$

Once more: the resistance $R = V/I$ you will be measuring between a set of leads (see below) is related to R_{\square} as $R = \gamma R_{\square}$, where γ is the number of squares between the leads; γ is not necessarily an integer. For example, in the picture below $\gamma = 3.44$:



And, it must be realized that resistance R will be the same independent of the size of these squares: it will be the same for the sample shown above and the one below, or any other made of the same material and with the same number of squares γ :



Now, the Hall Effect formulas in 2-D will also lack sample thickness δ . It is easy to see that instead of (10) we will have

$$V_H/I = R_H B_z \quad (17)$$

7. Calculating microscopic parameters from measured ρ , R_{\square} and R_H

As can be seen from equation (11), in case of a single type carrier, measuring Hall coefficient R_H immediately gives carrier density n -- one of the most important characteristics of a conductor. High n corresponds to small R_H and small Hall voltage V_H . Accordingly, in good metals with typical bulk concentrations $n \sim 10^{29} \text{ m}^{-3}$, measuring very small Hall voltages may be a challenge. In semiconductors, where n is orders of magnitude lower, V_H is usually in the mV range, making measurements easy. Of course, increasing B field also helps to make the measurement of V_H (and eventually n) easier.

Measuring R_H in the Hall measurement and determining either ρ in 3-D or R_{\square} in 2-D from a resistance R measurement, we can determine mobility as

$$\mu = R_H/\rho \quad (3\text{-D}) \quad (18)$$

Or

$$\mu = R_H/ R_{\square} \quad (2\text{-D}) \quad (19)$$

Knowing mobility and m^* , we can further calculate τ from formula (4).

How do we find m^* ? Sometimes m^* is known from band theory; sometimes from other measurements. In the absence of any idea, putting it equal to free electron mass m one can at least roughly estimate τ . However, such an estimate may be a factor of 10 – 20 off for a material such as GaAs. As we said in Section 2, in our material, GaAs, $m^* = 0.067 m$.

Knowing τ , if we knew the intrinsic carrier velocity (which we called Fermi velocity v_F), we could further calculate the average length a carrier travels between collisions: the *mean free path* (mfp) $l = v_F \tau$. Note that it is the intrinsic velocity v_F this time, and not the much smaller drift velocity v_d that gives us the mean free path; indeed, the additional velocity acquired in the field, $v_d \ll v_F$, changes the distance carriers travel very

little. It is clear that high mobility (large τ and small m^*) corresponds to long mfp. The mfp may be compared to some characteristic lengths associated with our sample: the distance between atoms, de Broglie wavelength at the Fermi momentum, the potential well thickness in a 2DEG sample, the cyclotron radius in the highest B field we can apply, etc.

Do what you can to calculate and fully utilize the knowledge of microscopic parameters: this is the primary interest in doing this experiment.

8. The 4-probe method

A simple measurement of a resistance we learned about in a first E&M course is performed either by measuring both the current through the resistor and the voltage across the resistor (I and V measurement) (Fig.4):

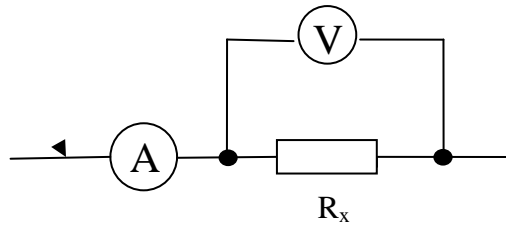


Fig. 4 Two-terminal resistance measurement

or by measuring a current, supplied by a battery, which flows through an adjustable resistor and the unknown resistor connected in series, with the scale being set as to read the unknown resistance (an *ohmmeter* measurement, which we do not show here). Either way, it is a *two-terminal* measurement: the unknown resistor is connected to the rest of the circuit at two points, or two terminals, shown in Fig. 4 as black dots. In this measurement it is implicitly assumed that *contact resistance* and resistance of the wires is negligible compared to the resistance R_x . In many cases, especially when dealing with commercially-made resistors, this is an excellent assumption.

There are however many experimental situations in which this measurement is complicated by the presence of finite, non-negligible *contact resistance*, and sometimes also by the presence of *lead resistance* (such as resistance of wires). Consider for example a measurement performed on a sample shaped as a uniform-width strip (for example, it may be a thin film sample). When performing a two-terminal measurement, we place two metal contacts at the ends of a film strip and attach to them current and voltage leads, as shown in Fig. 5:

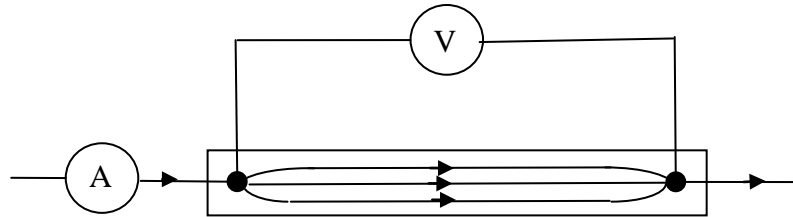


Fig. 5 Two-terminal measurement on a sample shaped as a strip, with two contacts

There are at least two problems with this arrangement:

1). If there is appreciable contact resistance between the film and the contacts (black circles in Fig. 4), the equivalent electrical circuit describing this measurement will look like the one depicted in Fig. 6, where resistors R_{cont} indicate the presence of said contact resistance:

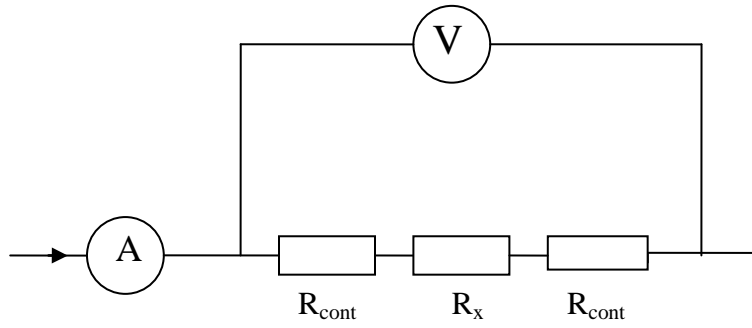


Fig. 6. Equivalent circuit showing presence of contact resistance in a two-terminal measurement

If R_{cont} is not negligibly small compared to R_x , voltage drops across them, $2IR_{\text{cont}}$ will be included in the voltmeter measurement. Note that here the full sample current I flows through these R_{cont} .

2). Current flow lines in Fig. 5 may be non-uniform around the contacts. This makes the length of a sample ill-determined, and thus will introduce an error in the calculation of resistivity, even if contact resistances would be negligible.

We note that the same un-avoidable problems exist in an *ohmmeter* two-terminal measurement.

These problems with the two-terminal connection can be overcome by making a *four-terminal measurement* described below. This measurement is known under different names such as *4-probe method*, a *4-terminal sensing*, a *4-probe geometry*, or as *Kelvin sensing* (first described by Lord Kelvin a very long time ago). This method is designed to exclude from the measurement contact and wire resistances. It is universally used in Solid State Physics to perform accurate measurements of true material parameters, primarily of resistivity ρ and R_{\square} , as well as in other situations when contact resistance needs to be excluded from the measurement.

In order to perform this measurement correctly, the material we want to study must be delineated into a specific geometrical shape. First, it will be a uniform-thickness layer of thickness δ (if our sample contains 2DEG, we will not consider this thickness in our calculations, but, nonetheless, a heterostructure containing this 2DEG will have some uniform thickness). Second, it has to be shaped in the xy -plane in a certain way. One version of 4-probe geometry is that containing a long “spine” and protruding “fingers”, as shown:

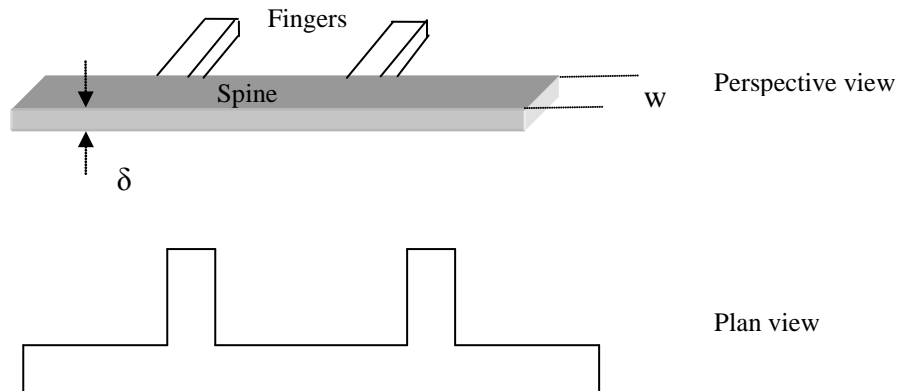


Fig. 7. Spine and fingers 4-probe geometry

Our samples are of the first type, shown in Fig. 7, with “fingers” (actually, with two sets of fingers on the two sides, for Hall measurement, see below). Next, electrical contacts must be placed on a sample. We will pass a known current along our sample using two outer current contacts **a** and **b**, and measure a potential drop across a certain known length of a sample L with another pair of inner voltage contacts, **c** and **d**, as shown in Fig. 8:

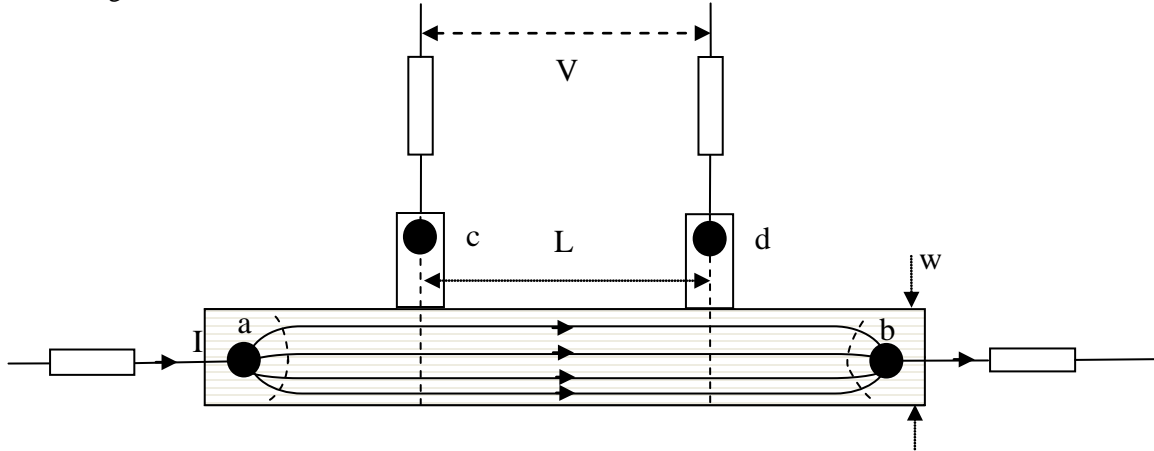


Fig. 8. 4-probe measurement

In Fig. 8 the current is shown to flow along the spine only, not flowing into the fingers. This is a good approximation in most cases (however, see below). The current *flow lines* in the *spine* can be thought of as lines drawn in the direction of local electric field. The *equipotential surfaces* (represented here in cross-section by the dashed lines) must be perpendicular to local electric fields; therefore they are curved near the contacts (see also Fig.5 above), while becoming straight away from the contacts, in the middle part of a sample, as shown. It is significant that the sample geometry and contact placement are chosen so as to guarantee that between the voltage contacts **c** and **d** current flow be uniform; corresponding electric field is also uniform, and, therefore, corresponding equipotential lines separated by a distance L are straight, as shown. Equipotential lines will continue through the fingers, touching metal contacts **c** and **d**; by virtue of being metallic, the contacts **c** and **d** will be at the potentials dictated by these lines.

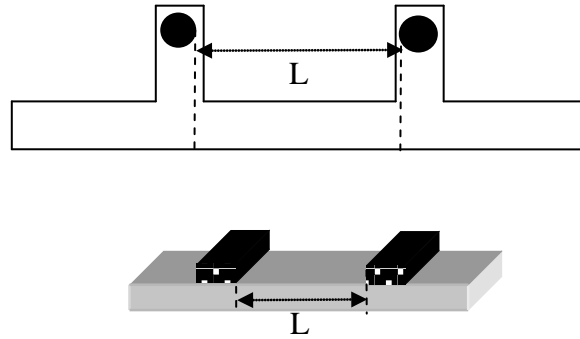
Why do we need side *fingers*? Why not place the voltage leads right on the *spine*? If we placed small voltage leads onto the *spine* like that:



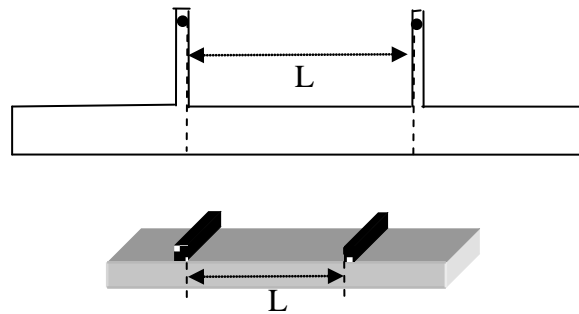
they could disturb the current flow, and thus distort the equipotentials. Wide voltage contacts all the way across the spine, however, will not distort the current flow, and this “top contact” geometry is also being used in 4-probe measurements [12]:



We note that, depending on the details of the voltage contact placement, the exact sample geometry, and the conductivity of the material under study, the position of these equipotential lines -- which in turn define distance L -- may change somewhat. Clearly, the metal contact itself (the black dot) will be all at one potential. The extended, large size metal contact may “short out” the width of a finger, or the width of a spine under the contact in the other geometry, so that equipotential lines corresponding to contacts **c** and **d** will shift towards the inner edges of said contacts, as shown below for the two geometries:



A detailed analysis of various scenarios of this type can be found in the literature (as an example, see Ref. 12). However, in practice, one prefers to avoid such complications by making side fingers or top contacts relatively narrow (approximating ideal “zero-width contacts”), so that a mistake made in the determination of true L is minimized:



In Fig. 8 we also show resistors symbolizing contact and wire resistance of current and voltage leads, those R_{cont} we discussed above when talking about a two-terminal measurement. Considering current leads, these resistances do not influence our measurement, because we are either dialing a known current from a current source, or measuring the actual current that flows in the sample. Considering voltage leads, there is (almost) no current flowing through them, and so these resistances also do not enter into a consideration in a 4-probe measurement. This is the main point: we are measuring a potential drop between voltage contacts, practically without a current, and those contact and wire resistances therefore became unimportant to us!

A note of caution is in order: said resistances in the voltage leads will however become important if they are comparable or exceeding the impedance of our voltmeter (which is usually 10 Mohm or higher). So, practically, there is a limit to just how bad our contacts can be. To understand why this is so, recall that a real voltmeter does need a tiny current for its operation. The last statement applies to both analog and digital voltmeters: whether the current is needed to move the *d'Arsonval galvanometer* in an old analog meter, or to charge a capacitor in a digital one, it is fundamentally needed nonetheless. (The only voltage-measuring instrument which does not need a current is a *potentiometer*, which is not practical in the type of measurements discussed here). That is why all voltmeters have finite – not infinite – impedance R_{imp} (internal resistance): they must admit some current. So, a voltmeter is essentially a current-sensing instrument converted to measure a voltage. It fundamentally comprises a current-sensing element (such as a galvanometer G) in series with a very large resistor R_{imp} , as depicted in Fig. 9:

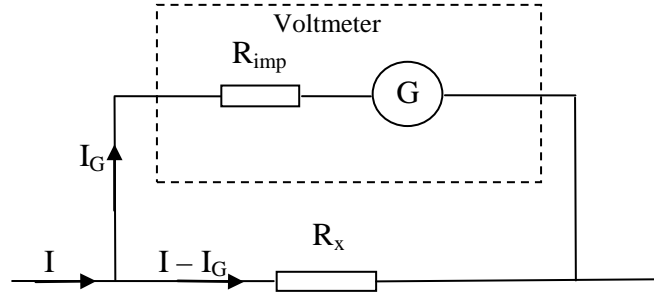


Fig. 9. Recalling how a voltmeter works

We can see that $(I - I_G)R_x = I_G R_{imp}$, and thus $I_G = (I - I_G)R_x / R_{imp} \approx IR_x / R_{imp} = V_x / R_{imp}$, where we neglected $I_G \ll I$. The instrument, while measuring I_G , is calibrated to read voltage across the sample V_x , which is proportional to I_G .

Now, suppose that contact resistances are comparable to R_{imp} . The corresponding electrical circuit showing these unknown contact resistances R_{cont} is shown in Fig. 10:

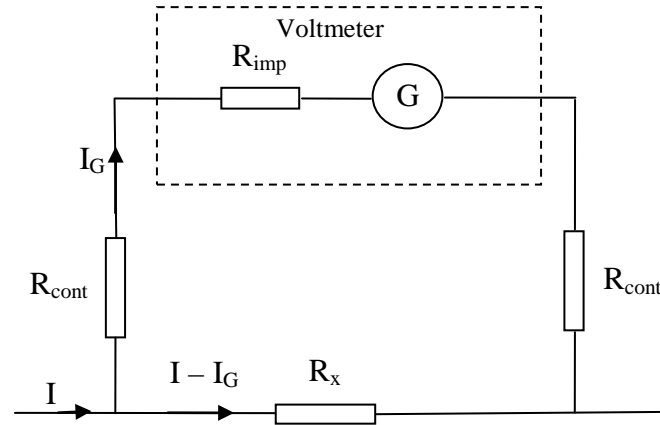
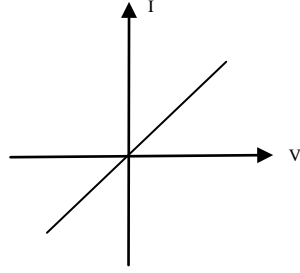


Fig. 10. A voltmeter measuring a sample with large contact resistances

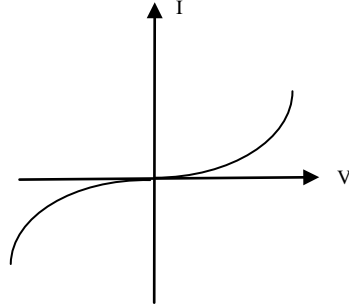
Here we have $(I - I_G)R_x = I_G (R_{imp} + 2R_{cont})$, or, again neglecting $I_G \ll I$, $IR_x \approx I_G (R_{imp} + 2R_{cont})$, or $I_G = V_x / (R_{imp} + 2R_{cont})$. Since R_{cont} is not known, the voltmeter will measure incorrect I_G and thus it will display an incorrect V_x .

Note the difference between this circuit and the one depicted in Fig. 6: here a very small current I_G flows through contact resistances; there, in a 2-terminal measurement, the full current I was flowing through these resistances. This is the main advantage of a 4-terminal over 2-terminal measurement.

Generally it is rather difficult to make good contacts to high-resistivity semiconductors. Contacts to semiconductors and avoidance of contact problems constitute an active field of study, with many publications devoted just to this subject (for example, Ref 12). One characteristic of a good contact is that it is *ohmic*, that is, it has a linear I-V curve:



One often finds that high-resistance “bad” contacts are usually also *non-ohmic*, having a non-linear I-V curve, for example like this one:



As one can see, here not only R_{cont} is unknown, but it also is not well-defined, depending on the current I_G that is being passed through it. If in doubt about a contact, measure its I-V characteristic, and stay away from the ones with a non-linear I-V.

Returning to 4-probe method as depicted in Fig. 8, if we will now take the ratio of V_{cd} to I_{ab} , it will give us true resistance R of the part of our sample of length L , and true resistivity ρ and true R_{\square} of our material will be found as $\rho = (wt/L) R$ and $R_{\square} = (L/w)R$.

9. Resistivity, R_{\square} and Hall measurements in a 4-probe geometry

When measuring both resistivity and Hall voltage, we will need additional *fingers* placed symmetrically on the other side of a spine, as shown below in Fig. 11, which depicts the so-called “Hall geometry”. Also shown in Fig. 11 is a DC electrical circuit and instrument connections for resistance and Hall measurements:

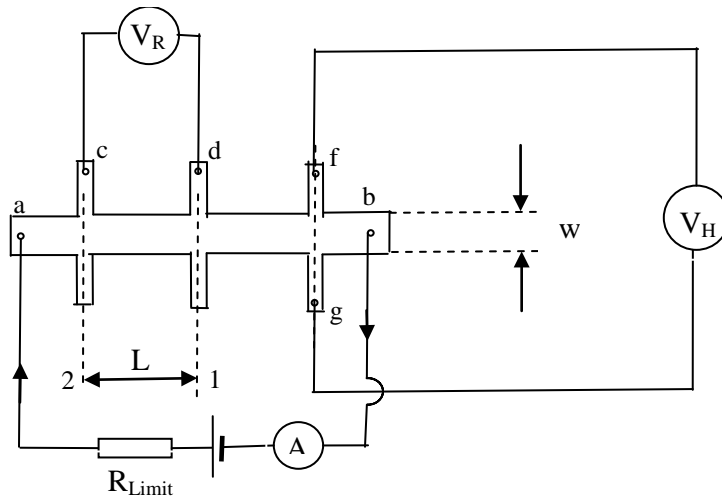


Fig. 11. Resistance and Hall measurement DC circuit

Knowing the current I_{ab} and measuring the voltage V_{cd} between fingers **c** and **d** (or voltage V_{df} between **d** and **f**, or a voltage between any other pair of contacts along the *spine*) we can find resistance of a corresponding part of a sample $R = V/I$. As was said above, in 3-D we can then find resistivity $\rho = Rw\delta/L$, where w is the width of a spine, δ is sample thickness, and L is the distance between fingers. In 2-D, we can find $R_{\square} = R/\gamma = Rw/L$. The current may be kept constant by using a current source, or by having a large resistor R_{Limit} in series with the sample. Alternatively, instead of keeping current constant, one can measure current using an ammeter A simultaneously with voltage, and take their ratio to find R .

NOTE: large current can damage 2-D sample; the current through your 2-D sample should not exceed 10 μA .

As was explained in Sections 3 and 4, zero magnetoresistance signifies that we are dealing with a single type carrier, while non-zero magnetoresistance indicates presence of two types of carriers. Thus the first thing to do is to turn on magnetic field and to check resistance between longitudinal contacts such as **c** and **d**. If resistance is not changing in the magnetic field, i.e. if magnetoresistance is zero (practically, within experimental error from zero), one can be assured about the type of analysis that will be needed to find carrier type (sign) and density.

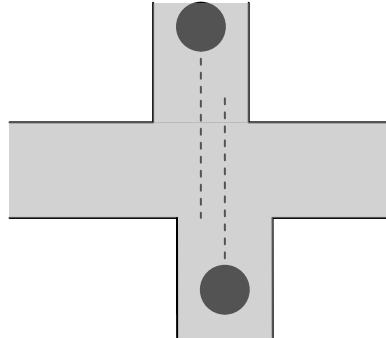
Next one can determine carrier sign (see Section 5); to do that we need to know all the relevant directions: that of magnetic field, of the current flow, and the polarity of Hall voltage.

Finally, in the magnetic field, opposing leads such as **f**, **g** will measure Hall voltage across the sample, and from that we find Hall coefficient $R_H = V_H\delta/(IB_z)$ in 3-D and $R_H = V_H/(IB_z)$ in 2-D.

10. Some sources of systematic errors

Geometrical offset:

When measuring Hall voltage between leads **f**, **g** or any other geometrically-opposite pair of leads, we do not want to include any potential difference due to sample resistivity; they should ideally be at the same equipotential. So, in the absence of magnetic field, but with the current flowing in the sample, the voltage between such opposite Hall leads should be zero. In reality one often finds finite offset voltage between nominally symmetrical opposite fingers. This offset may originate from a mismatch between the positions of the opposite fingers (imperfect geometry), as shown here:



Clearly, such an offset voltage will change sign upon current reversal, and will scale with the value of the current, becoming zero at zero current. Thus it can be identified (at zero magnetic field) and subtracted from the measured Hall voltage.

Another way of getting rid of such an offset is to measure V_H vs. B ; according to formula (17), the slope of such dependence will be equal to IR_H . We can measure this slope, even if there is a zero B -field intercept, and, knowing the current I , find R_H .

Thermal emf:

There is yet another source of a systematic error in a Hall measurement, as well as in a resistance measurement. This one is very common, and also needs to be identified and removed. It is the so-called “thermal emf”. This emf (or voltage) is current-independent: it does not scale with the current, and exists even at zero current, which allows for its identification. It appears due to *thermoelectric effect*, which manifests itself as the direct conversion of temperature differences to electric voltage. It appears in the presence of thermal gradients across different conductors comprising a circuit (such as sample itself, contacts to a sample, wires connected to contacts, resistors used in the circuit). It is the same as the emf (or voltage) appearing in a *thermocouple*, except that in a thermocouple it is desired, while in a measurement such as the one we are describing -- it is not.

How can we eliminate this systematic error?

- a). Having to do with temperature differences, thermal emf tends to decrease or completely disappear as the sample temperature is equalized with its surroundings. Sometimes it is enough to wait until thermal emf disappears.
- b). If it does not go away, the way to remove systematic error due to thermal emf is to look at the voltage (measured between Hall or resistive leads) at zero current. The remaining voltage will be most likely due to thermal emf, and it should be subtracted from V which will be measured with the current (in doing so one should be careful about its sign!). If there is appreciable thermal emf, one should measure it before and after each V measurement, as it tends to change with time (probably due to thermal gradients changing with time).
- c). Thermal emf can be also eliminated by taking measurements with reversed current and then averaging the absolute values. If V_+ is taken with current in one direction and V_- with current in the other, reverse direction, then $V = (|V_+| + |V_-|)/2$ will eliminate thermal emf offset.
- d). Measuring the slope of V_H vs. B , which, as was said above, is removing geometrical offset, also tends to largely remove systematic error due to thermal emf. Indeed, thermal emf is added to the true value of V_H ; this addition, if it is the same at different points, will not influence the slope of V_H vs. B ; however, if it drifts with time, this will not help.

Sample position in the magnet:

The apparatus in our Lab does not provide for precise positioning of sample's xy-plane perpendicular to the magnetic field. This can in principle lead to a systematic error in R_H determination. To a large degree this can be improved by turning the probe and watching the Hall voltage: the maximum Hall voltage at a given field will correspond to the best sample position with respect to turning around the probe axis. But, clearly, this orients the sample only partially; it may still be tilted in the other, orthogonal direction. Students should try and estimate this systematic error; indeed, if it turns out to be larger than the other errors we are worrying about, there will be no excuse in missing it.

What can we do here? Decide, from examining a sample probe inside the magnet, what is the realistic maximum tilt θ_{\max} that you may have: is it 5° ?, 10° ?, more? You may produce such tilting in the Gaussmeter probe (you will need to think of some way for the rough calibration of probe angles) and see what happens to your field measurement. You can also estimate this error by calculating a $\cos(\theta_{\max})$: indeed, such will be the mathematical effect of tilting.

11. AC vs. DC measurements

In DC measurements there is an advantage of simplicity, and in Hall measurement you have to use DC at least at some point in order to find the sign of the carriers in your sample. However, in a DC measurement, there are offsets due to thermal emf as discussed above in Section 10, and, additionally, DC voltage is subject to various drifts and noise. Therefore, often one uses an AC measurement, which consists of

passing known rms AC current and measuring rms AC voltage. This is equivalent of continually doing the reversed current measurement of true $V = (|V_+| + |V_-|)/2$ as described above. The most refined way of doing AC measurement is with the use of a frequency and phase sensitive detection called *lock-in* method. The instrument is called *lock-in amplifier*. This technique is part of a general principle of bandwidth-narrowing. You can read about bandwidth-narrowing techniques and specifically about lock-in method in Horowitz and Hill [13], and it is well explained in lock-in amplifier manuals. Sometimes, when there are significant parasitic inductances in the circuit, the lock-in measurement may suffer from the unwanted frequency dependence, and one has to then measure at the lowest frequencies of only a few Hz, or even fraction of a Hz. As a rule of thumb, when signals are below a mV, it is beneficial to use *lock-in amplifier*; above a mV one can typically do a DC measurement.

12. Instruments

Magnet and Gaussmeter

The magnet will produce a maximum field of 0.15 T (at its maximum setting the power supply gives a current of about 1 A; it is intentionally limited at this current, to avoid magnet wiring heating). The field can be measured with the probe (Gaussmeter), which is provided. Incidentally, it is also a Hall probe! Be sure to measure the field when the probe is positioned to give a maximum reading, which corresponds to its surface being perpendicular to the magnetic field. When it is turned at an angle, the reading will decrease.

You will also need to find the North and South magnet poles in order to find carrier sign in your sample. This can be done with the help of a simple compass (which should be kept away from the strongest field, or it may magnetize incorrectly), or with the same Gaussmeter. One test is to look at the magnetic field of the Earth (your Gaussmeter is sensitive enough for that – and it is a fun thing to do; look also at the Earth magnetic field *inclination*), and, knowing the geographical North direction, determine the polarity of the Gaussmeter signal, and then use it in the magnet to find the direction of the field. Remember that the end of the compass needle marked “N” is designed to point to the geographical North Pole, which is actually the South Pole of the Earth magnetic field. Thus the “N” side of the needle will point to the “S” pole of the magnet too. Reverse the current in the magnet, and observe the behavior of the compass needle.

Multimeters

Multimeters are described in their reference manuals. Initially, when you check sample leads, it is convenient to use ohmmeter with just two probes. However, you should always remember that you cannot pass more than 10 μA through the sample. A multimeter on an ohmmeter setting will send a certain current through the sample, and measure voltage required to provide for this current, the ratio being resistance. Thus watch for the multimeter specs given in the reference manual: the current it will push depends on a resistance range, and it may be much greater than 10 μA ! It is sufficient to make sure that you are measuring in a *manual* mode (*man* written on the display) on a 100 $\text{k}\Omega$ or higher ($\text{M}\Omega$) scale. The instrument will not switch to a different range if it is set to a manual mode, and on 100 $\text{k}\Omega$ or higher scale the current is small enough. **Make sure of that before doing a two-probe measurement on a sensitive sample.**

You will use multimeters to measure voltages and currents. Note that at 10 μA , in our multimeters, the current can be measured only to 3 digits (reliably, probably to 2 digits only). This will limit your precision. You may want to determine current by an indirect but more precise measurement of a voltage across a well-known or well-measured resistor. See however the discussion of the positioning error above: is it much larger than the expected error from current uncertainty? Of course, the positioning error only relates to Hall voltage measurement, and not to resistivity. But all of our microscopic parameters will be calculated based on both measurements, so in this sense a large error in R_H will imply large errors in all calculated quantities.

Liquid Nitrogen dewar

The *dewar* is to be filled with liquid nitrogen. Ask Lab instructor to help you with that task. **Be very careful with the sample probe and the glass dewar. Do not apply force taking probe in and out of the dewar. Observe safety rules with liquid nitrogen.**

When a sample probe is taken out of liquid nitrogen, being very cold, it will get wet from the moisture condensing on it. It is best to use a fan to quickly dry it off. Even though the sample is encapsulated in a holder, the moisture may damage the delicate heterostructure and its contacts.

13. Sample layered structure, effective masses, contact numbering in a sample holder

Our two samples are GaAs/AlGaAs heterostructures grown by MBE in Prof. Emilio Mendez's Lab, by Bent Nielsen.

The vertical structure:

Our samples consist of the following layers, numbered above the GaAs substrate from the bottom up; not drawn to scale here in terms of layer thicknesses:

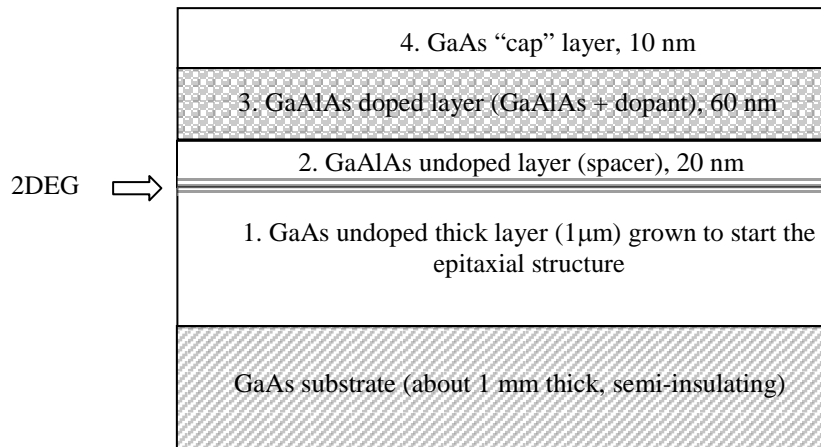


Fig. 12. Sample's layered structure (not to scale)

The layers are grown on a single-crystalline GaAs substrate about 1 mm thick. From the substrate up:

1. The first layer, 1 μm thick, is grown in part to start the single-crystalline structure, and in part to provide the interface where *band-bending* creates an inversion layer and a narrow potential well responsible for the carrier quantization (see Fig. 13 below). The interface containing 2DEG is indicated in Fig. 12 and Fig. 13. The approximate extend of the wavefunction of the conducting carriers in the direction perpendicular to this layer is 10 nm.
2. The second layer: undoped 20 nm GaAlAs *spacer* which separates the conducting inversion layer on top of layer 1 and the source of *dopants* (i.e. atoms contributing charge carriers) distributed in the next layer 3. This is the principle of *modulation doping* [7], when *dopants* are spatially separated from the mobile carriers in order to reduce carrier scattering and enhance mobility.
3. The third layer is a 60 nm GaAlAs layer containing dopant atoms incorporated into a GaAs lattice. The dopant may be Silicon for n-type (as shown in Fig. 13 below) and Boron for p-type.
4. The forth layer on top is a capping layer which is required for the proper utilization of the doping layer 3 and for protection of the structure.

As you can see, the total thickness of the part of the structure involved with electrical 2-D conduction and quantization resides in a thin upper part of layer 1 and in layers 2 – 4; the total thickness of this active part

of the structure is only about 100 nm, i.e. approximately 400 atomic layers (taking lattice spacing at about 0.25 nm). The 2DEG exists in a ≈ 10 nm inversion layer, i.e. in about 40 atomic layers. A schematic *energy diagram* in Fig. 13 shows energy profile in layers 1, 2 and 3.

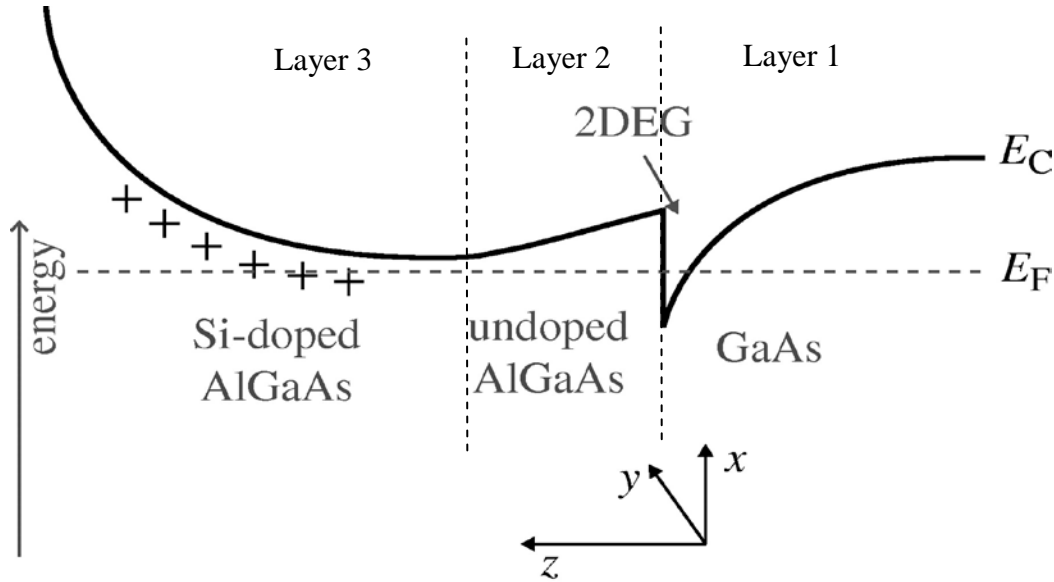


Fig. 13. Schematic energy diagram for the conduction band in n-type modulation-doped GaAs/GaAlAs heterostructure. The solid line labelled E_C represents energy profile at the bottom of the conduction band. The dotted line labelled E_F represents Fermi energy lying below the bottom of the conduction band, this band therefore being empty except for the state labelled E_F in the triangular-shaped potential well (also called *inversion layer*) at the interface between layers 1 and 2 -- the 2DEG.

The effective masses for electrons and holes in GaAs:

Electrons in GaAs have $m_e^* = 0.067 m$; holes in GaAs have $m_h^* = 0.35 m$. Once you identified the carriers, you will need these values for your calculations of microscopic parameters.

The planar pattern and the leads arrangement:

A single-crystalline layer of GaAs, typically about 1 mm in thickness, which serves as a substrate for MBE growth, is called a *wafer*. A multilayered structure grown in an MBE apparatus uniformly covers the whole wafer. The wafer with the grown structure on it is then being cut into *chips*, and the planar Hall and resistivity geometry is delineated on a chip with subsequent use of *photolithography* [14] and etching. Metal contacts are deposited (typically, evaporated in a separate vacuum system) over the pads and subsequently annealed in a furnace to provide good ohmic contacts (see Section 8). The resulting sample is a small chip 3-5 mm in size. This chip is mounted on a small standard round connector, with all the “fingers” of the planar pattern wire-bonded to ultra-thin gold wires (see photograph in Fig. 15 below) connected to the connector’s pins. The connector is plugged into the probe, which has a box attached to it, which provides labeled contacts to connector pins. The numbering arrangement corresponding to holder box numbers is shown in Fig. 14. Note the position of the “tab” which is visible on the outside of the holder; it is close to #5 lead.

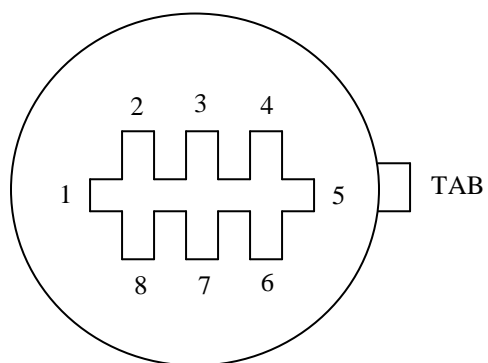


Fig. 14. Schematic sample planar geometry and numbering in the holder (see photo in Fig. 15 for the real geometry)

In Fig. 15 there is a photograph of the central part of one of our samples with the 100 micron scale shown in the lower right corner. The light-colored films seen on the square pads are the metal contacts (Ti-Au alloy) which we were depicting as black circles in the text above.

One can measure the actual sizes and the number of squares between voltage leads using this photo. A standard lithographic mask of these dimensions is used on both samples; we will have to assume that the dimensions are the same in all sets of leads. Students should use this measurement to estimate geometrical errors (uncertainties), which will in turn influence the uncertainties in each of the measured and, eventually, calculated quantities.

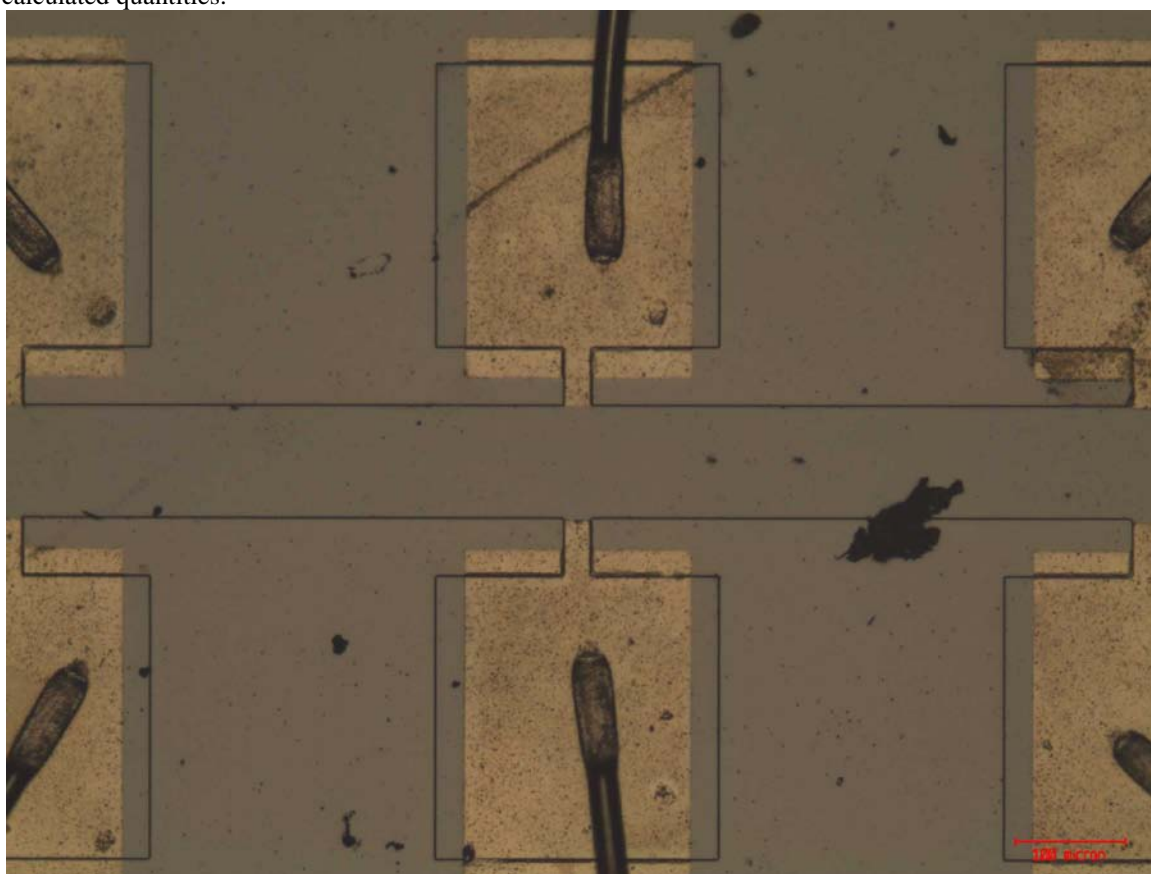


Fig. 15. Photo of one of our samples. Notice light metal pads (contacts) and thin gold wires attached to the pads.

14. Errors, uncertainties

When doing measurements, you will encounter sources of *systematic errors* (some of which we have discussed above), as well as *random errors*, or *uncertainties*. Systematic errors should be identified and eliminated to the best of our ability. There should be “zero tolerance” for systematic errors. Yet there are some things which cannot be easily improved, short of changing the experimental setup. We may have to include uncertainties resulting from them when we report the final results. As to the random errors, here is what we can advise: Try to estimate uncertainty in each of your measurements that will contribute to the final experimental result -- in this experiment, to R_{\square} and to R_H . After all is set and done, the goal is to find and report correct values of $R_{\square} \pm \Delta R_{\square}$ and $R_H \pm \Delta R_H$, with *realistic* uncertainties ΔR_{\square} and ΔR_H justified by your actual apparatus and the measurements that were performed in your experiment. You will encounter *instrumental* uncertainties determined by how accurately you can measure a given quantity; for example, this refers to measuring a number of squares γ using a sample photograph, or to measuring a voltage V on your multimeter, etc. You will also encounter *random errors* which can be gauged by doing repeated measurements of the same quantity. If you do a least squares fit to a set of data, you will also need to find the uncertainty associated with the result (for example, the slope of a linear dependence).

Keep a perspective, however, on the relative importance of these various uncertainties. Indeed, some uncertainties may be much more important than the others. If you identify, for example, some unavoidable uncertainty in your experiment which will lead to, say, 5% error in V_H (we use 5% just to take some specific number here), then it does not make much sense trying to improve your measurement in terms of another 0.2% uncertainty. If, however, different uncertainties are of comparable weight, then it makes sense to keep track of all of them.

Once you determined experimental uncertainties ΔR_{\square} and ΔV_H , they will serve to find uncertainties in your calculated parameters which follow from R_{\square} and to V_H (such as for example n , μ , mfp, etc.) To do that you will need to use *error propagation formulas*.

15. Suggested steps in doing this experiment and in analyzing the data

General: Some of the following steps are only a suggestion; one can and should improvise and do or calculate other things, or the same things in a different order, provided that the sample will not be damaged, and the experiment will be eventually done in a meaningful and complete fashion.

In any case, **start by reading this write-up**, by understanding the theory and the experiment. You will need to write the essential theoretical ideas in your final report. We do not want you to just copy and paste from this write up; we would like you to explain the essential Physics of electrical conduction and of Hall Effect in your own words.

Keep good records of everything you do, with circuit diagrams, notes, numbers, graphs, estimates of uncertainties, questions that may occur to you during the experiment, etc.

1. Study multimeter manuals; particularly, find out where, to which inputs, one should connect samples for different measurements. Before measuring the actual Hall sample, test your multimeter Ohms (Ω) function on the “**manual**” setting by making resistance measurements on a regular 1 to 10 k Ω resistor (plug it into a breadboard) and, using a second multimeter as an ammeter, simultaneously monitor the measuring current to make sure that it is not higher than 10 μ A.
2. Preparing for the 4-probe measurement, make sure that your sample current power supply has a hard-wired large (1-2 M Ω) series resistor (actually, two resistors) to limit the current. Explore the limits of current that you can get from your supply. Measure current using a multimeter.

Question: How can you measure small μ A-range currents with precision exceeding the precision you can get out of a multimeter on the most sensitive available current setting (I)?

3. Measure a regular 5 - 10 k Ω resistor (placed in a “bread board”) using the “pseudo” 4-probe measurement: pass 10 μ A current and measure voltage across this resistor; compare the result with your previous multimeter (ohmmeter) measurement of the same resistor.
4. Take one of the real samples, check all the leads in a two-probe measurement, first making sure that your ohmmeter does not pass a current larger than 10 μ A through the sample (see sample contacts arrangement in Sec. 13 above). Make a diagram marking on it working and damaged leads. Note values of resistances you are getting between different leads.

Question: when you measure between different leads, involving different parts of a sample, do these resistances make sense given the sample geometry?

5. Using the 4-probe method, measure sample resistance at room temperature and, using the sample photo to measure the number of squares, find R_{\square} . Decrease the current to zero and see if you have any thermal emf present. If you do, change current direction and average the voltage as described above to eliminate thermal emf. Do this for all available nominally identical sets of leads along the sample (such as 2-3; 3-4, etc).
6. You may at this point want to estimate the uncertainty in your determination of R_{\square} . Estimate the uncertainty in your measurement of the number of squares between sets of leads visible in a photo. Next, get the experimental uncertainty in this measurement from the measured values for different sets of leads. Also take into account instrumental uncertainty in measuring I and V. Based on all this, and, when applicable, using error propagation formulas, find final uncertainty in your R_{\square} . Decide what contributed most to this uncertainty.
7. Explore the magnet, magnet power supply, and magnetic field measurements with the Gaussmeter; decide on the direction of the magnetic field in the magnet.
8. Place the probe in the magnet, inside the glass *dewar* (which is now empty), positioning it in the center of the pole-pieces of the magnet. Switch on the magnet current and go to the maximum current allowed by the power supply (i.e. to the maximum field). At the maximum field, make sure your sample is aligned perpendicular to the field,

Question: how will you do that?

and then measure resistance for any pair of resistive leads along the sample, and compare it to sample resistance in zero field. Do you find any magnetoresistance? How do you interpret what you see?

9. Estimate the uncertainty in a sample orientation in the magnet which will contribute to the uncertainty in the value of V_H and eventually to uncertainty in n , as was discussed in Section 10. Is this uncertainty serious, or can it be ignored in your experiment?
10. Taking all relevant directions and polarities into account, determine the sign of the carriers (see Section 4). Make careful notes and sketches in your lab book so that you will be able to explain this in the report.
11. Measure Hall voltage at room temperature at maximum B; check offsets at zero B field and at zero sample current. Subtract offsets from your V_H ; estimate the number of carriers.

NOTE: it is always advisable to quickly, even if approximately, estimate the value of interest during the experiment, to see if your measurement makes sense. If the measurement involves a graph, plot graphs in “real time”, as you take the measurements. If you do this only at the end, when the measurements are all finished, you may find too late that you want to go back and re-do some parts.

12. Try to improve on your Hall voltage measurement by measuring V_H as a function of B keeping the sample current fixed at about 10 μ A. Switch the current direction, as in resistivity measurement, and find the average slope (Hall coefficient) for the two current directions. Do this again for all available sets of Hall contacts (across the sample), and take the average value, excluding *outlying points*, if any (read error analysis texts to find what we mean by *outlying points*). You may encounter an offset in V_H vs. B, i.e. V_H may have an intercept at zero B field. Think of the possible

reasons (see Section 10). Note, however, that you can still find the slope, even if there is an intercept. To find that slope do a least squares fit to the data. However, reserve this to the time when you are sure that you obtained the best reliable data that you can get. Get V_H uncertainty from knowing error bars in your plot, and varying the slope within the limits of these error bars. If you do a least squares fit, find uncertainty from your fitting routine, and verify that it is consistent to the one you got by varying the slope within error bars.

13. You can alternatively measure Hall voltage with fixed B (maximum available in your magnet) while changing the sample current (up to $10\mu\text{A}$). Try this method too and compare results for the Hall coefficient.
14. Repeat the same measurements: resistance, magnetoresistance (if any), Hall coefficient (estimate n) in liquid Nitrogen (77K). Ask the instructor to fill the dewar with liquid Nitrogen – do not do it yourself.

Question: Do you expect R_{xx} to be higher or lower at 77 K compared to 300 K? Explain why. Write about this in your report. You may want to read in the literature (for example, in [8]) about carrier scattering mechanisms. Do you expect mobility to be higher or lower at 77 K? Relaxation time? Mean free path? Carrier density? (the answer to the last question may be hard to guess in modulation-doped 2DEG samples, though).

15. Once you determined the carrier signs for different samples, you can use the appropriate effective masses (Section 13) in further calculations. Calculate all you can from what you have measured both at 300 K and at 77 K (n , μ , τ , v_f , Fermi energy, mfp). To find mfp you will need to estimate Fermi velocity in a free 2DEG from your value of n and from the effective mass. Derive the 2-D free-electron formula for v_f (present your derivation in the report). Calculate also free-electron Fermi energy in 2-D. Comment on whether your semiconductor is *degenerate* or *non-degenerate* at 300 K and at 77 K (the criterion of degeneracy is Fermi energy vs. kT , see [8]).
16. Whenever possible, compare microscopic parameters you found with simple-minded theoretical estimates, with known relevant sample dimensions, etc.
 - Compare your n to the value n_{metal} that would be found in a 2-D (single atomic layer) mono-valent metal. To get n_{metal} , assume a lattice constant of $a_0 = 0.25 \text{ nm}$ in a simple quadratic (in 2-D) crystalline structure, so that one carrier is contributed per a_0^2 . Can you estimate the number of donors per square in your sample? What is the average distance between donor atoms in your sample, assuming they form a monolayer?
 - What is the average distance between free carriers in your sample?
 - Compare mfp at 300 K and at 77 K with some meaningful length scales relating to your sample: you can compare mfp to the distance between atoms (taking it to be 0.25 nm), to the distance between carriers (see above), to deBroglie's wavelength $\lambda_{\text{dB}} = h/p_f = h/(m^*v_f)$, which you can also calculate, to the cyclotron radius at the maximum magnetic field in your experiment. Will carriers in your sample be able to complete one cyclotron orbit at 300 K? At 77 K? If not, what should be the value of magnetic field which will allow them to complete one such orbit (one circle)?
 - Calculate v_d in your sample at the maximum $10 \mu\text{A}$ current and compare it to v_f .
17. Summarize; comment on all your findings. Point out puzzles and unclear aspects of the experiment, if any.

References

There are some web sites listed here as references; many more can be found on the Web, in Wikipedia, and in the extensive literature on these subjects

1. Edwin Hall, "[On a New Action of the Magnet on Electric Currents](#)". *American Journal of Mathematics* (American Journal of Mathematics, Vol. 2, No. 3) **2** (3): 287–92 (1879).
2. *25 years of Quantum Hall Effect*, K. von Klitzing, Poincaré Seminar (Paris-2004), <http://www.bourbaphy.fr/klitzing.pdf> ; Yennie, D. R. (1987). "Integral quantum Hall effect for nonspecialists". *Rev. Mod. Phys.* **59** (3): 781–824. doi:[10.1103/RevModPhys.59.781](https://doi.org/10.1103/RevModPhys.59.781) ; Quantum Hall Effect, www.warwick.ac.uk/~phsbn/qhe.htm
3. The fractional Quantum Hall Effect, www.bell-labs.com/news/1998/october/13/2.html
4. D. F. Holcomb, "Quantum electrical transport in samples of limited dimensions", *Am. J. Phys.* **67** (4), p. 278, April 1999.
5. A. Y. Cho, J. R. Arthur (1975). "Molecular beam epitaxy", *Prog. Solid State Chem.* **10**: 157–192. doi:[10.1016/0079-6786\(75\)90005-9](https://doi.org/10.1016/0079-6786(75)90005-9); also see A. Cho, "Film Deposition by Molecular Beam Techniques", *J. Vac. Sci. Tech.*, **8**, p. S31-S38, (1971).
6. L. Eski, *History and Perspective on Semiconductor Superlattices*, in "Synthetic Modulated Structures", Edited by L. L. Chang and B. C. Giessen, p. 3, Academic press, (1985).
7. R. Dingle, H. Stormer, A.C. Gossard, W. Wiegmann, *Appl. Phys. Lett* **33**, pp. 665-667 (1978); A. C. Gossard and A. Pinzuk, *Modulation-Doped Semiconductors*, in "Synthetic Modulated Structures", Edited by L. L. Chang and B. C. Giessen, p. 215, Academic Press, (1985).
8. Neil W. Ashcroft; N. David Mermin, *Solid State Physics*, Saunders College, (1976); M. A. Omar, "Elementary Solid State Physics", Addison-Wesley (1993).
9. M. Gurvitch, *Phys. Rev. B* **24**, 7404-7407 (1981).
10. Harris, J. J.; Foxon, C. T.; Barnham, K. W. J.; Lacklison, D. E.; Hewett, J.; White, C. "Two-dimensional electron gas structures with mobilities in excess of $3 \times 10^6 \text{ cm}^2 \text{V}^{-1} \text{s}^{-1}$ ". *Journal of Applied Physics* **61**: 1219 (1987).
11. Geim, A. K. and Novoselov, K. S. (2007). "[The rise of graphene](#)". *Nature Materials* **6** (3): 183–191.
12. E. J. Zimney, G. H. B. Dommett, R.S. Ruoff, and D. A. Dikin, *Meas. Sci. Technol* **18**, pp. 2067-2073 (2007).
13. P. Horowitz and W. Hill, *The Art of Electronics*, p. 1026 - 1033, 2-nd Edition, Cambridge University Press, (1989).
14. Jaeger, Richard C. "Lithography". *Introduction to Microelectronic Fabrication*. Upper Saddle River: Prentice Hall (2002).