

Hall Effect Experiment

by Dr. G. Bradley Armen
Department of Physics and Astronomy
401 Nielsen Physics Building
The University of Tennessee
Knoxville, Tennessee 37996-1200

*Copyright © April 2007 by George Bradley Armen**

**All rights are reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage or retrieval system, without permission in writing from the author.*

1. Introduction

The Hall effect is basic to solid-state physics and an important diagnostic tool for the characterization of materials – particularly semi-conductors. It provides a direct determination of both the sign of the charge carriers, *e.g.* electron or holes (appendix A), and their density in a given sample.

The basic setup is shown in Fig. 1: A thin strip (thickness δ) of the material to be studied is placed in a magnetic field \mathbf{B} oriented at right angles to the strip.

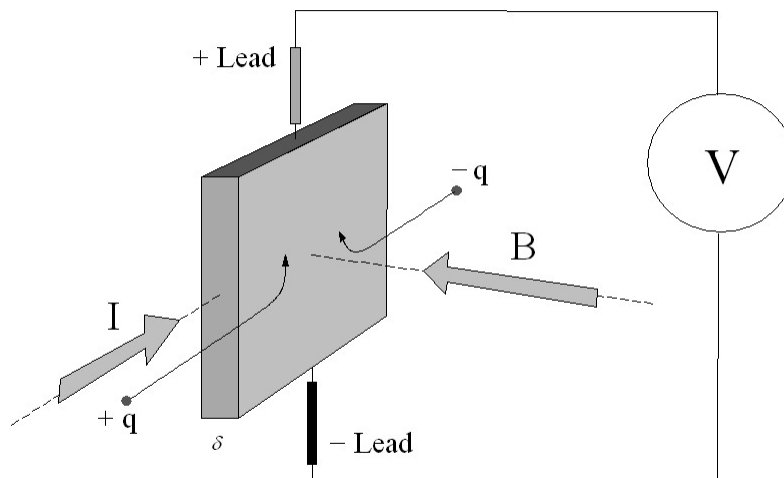


Fig. 1: Hall effect geometry. This arrangement corresponds to our laboratory setup.

A current I is arranged to flow through the strip from left to right, and the voltage difference between the top and bottom is measured. Assuming the voltmeter probes are vertically aligned, the voltage difference is zero when $B = 0$.

The current I flows in response to an applied electric field, with its direction established by convention. However, on the microscopic scale I is the result of either positive charges moving in the direction of I , or negative charges moving backwards. In either case, the magnetic Lorentz force $q\mathbf{v} \times \mathbf{B}$ causes the carriers to curve upwards. Since charge cannot leave the top or bottom of the strip, a *vertical* charge imbalance builds up in the strip. This charge imbalance produces a vertical electric field which counteracts the magnetic force, and a steady-state situation is reached. The vertical electric field can be measured as a transverse potential difference on the voltmeter.

Suppose now that the charge carriers were electrons ($q = -e$). In this case negative charge accumulates on the strip's top so the voltmeter would read a negative potential difference. Alternately, should the carriers be holes ($q = +e$) we measure a positive voltage.

The above argument provides a simple picture in which to think about the Hall effect — and in fact leads to the correct answer if pursued. However, it presupposes a *steady* current of charge carriers flowing in the conductor all in a single direction with constant speed. Why, for instance, don't the carriers accelerate? The true nature of macroscopic conduction is a bit more complicated, relying on a statistical average over individual carrier's motions. In the next section we will briefly look at this issue.

2. The Drude theory and the Hall effect

Before considering the effect of magnetic fields on conductors, we need some model to describe the flow of currents in response to electric fields. To do so we will use the Drude theory of conductors. This is a simple classical model, and many of its concepts extend to the quantum case. A more careful account of the Drude model can be found in the first chapter of “Solid State Physics” by Ashcroft and Mermin.

The Drude model envisions a conductor as a gas of free current-carrying charges. The freely moving carriers suffer randomizing collision events *on average* every τ seconds. The parameter τ is called the relaxation time¹, and is the only feature describing the (otherwise unspecified) collision events.

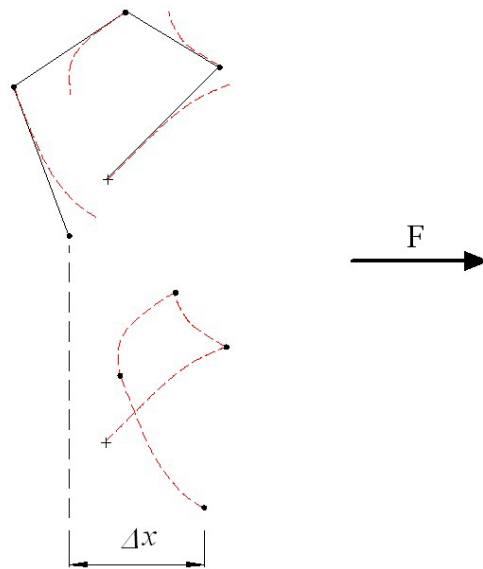


Fig. 2: Example of a charge carrier’s random trajectory after four collision events.

Figure 2 indicates the basic idea. The solid black lines show a specific trajectory of a charge carrier in the *absence* of any external force. It proceeds with some constant

¹ One might think that collisions occur more often the faster the particles are moving, so that τ could depend on the average particle velocity \mathbf{V} . However, we’ll see that the thermal velocities the carriers actually move with between collisions is very much greater than \mathbf{V} .

velocity in some direction until it experiences a thermal² collision. The red dashed lines indicate how each of these trajectories would be influenced by a constant external force \mathbf{F} : The particle now accelerates in the direction of \mathbf{F} consistent with its initial conditions dictated by the last collision. When these paths are assembled, we see that \mathbf{F} has caused a net displacement Δx after the four collisions; the particle has begun to ‘drift’ in the direction of \mathbf{F} .

For macroscopic quantities, only averages over all particles are of interest. With the above assumptions, one can obtain an equation of motion for the *average* momentum \mathbf{P} , given by

$$\frac{d\mathbf{P}}{dt} + \frac{1}{\tau}\mathbf{P} = \mathbf{F}.$$

Here, \mathbf{P} is interpreted as follows: a group of charge carriers (say $\approx 10^{23}$) all have some specific *average* momentum \mathbf{P}_0 at $t = 0$, the group’s average momentum $\mathbf{P}(t)$ at any future time is then determined by this initial value and the first order equation above. The average momentum \mathbf{P} of the group evolves in time *like* a single carrier subjected to the *average* external force \mathbf{F} per particle, but subjected additionally to a resistive damping. This damping is a consequence of all the random collisions; when τ is large collisions are uncommon.

For a constant force a terminal momentum is reached — *independent of the initial conditions*. This is in complete analogy with the motion of an object falling through air: the object either speeds up or slows down enough so that the air-resistance counteracts its weight. In our case \mathbf{P} becomes constant in time as $\mathbf{P} \rightarrow \tau \mathbf{F}$. Thus, if we apply a force \mathbf{F} and wait ‘long’ enough, a steady state develops where, *on average*, each particle drifts with a constant speed in the direction of \mathbf{F} .

In response to a constant electric field ($\mathbf{F} = q\mathbf{E}$) a constant current of charge carriers is established. The current density \mathbf{J} is the average charge crossing an (oriented) area per unit time. If there are n charge carriers per unit volume with average (drift)

² That is, the new velocity has a random direction and its magnitude, unrelated to the old velocity, is consistent with the temperature at the collision site (*i.e.* on average $\frac{1}{2}mv^2 = kT$)

velocity $\mathbf{V} = \mathbf{P}/m$, the associated current density is $\mathbf{J} = nq\mathbf{V} = nq \times \tau q \mathbf{E}/m$. Thus, we arrive at Ohm's law

$$\mathbf{J} = \frac{\tau q^2 n}{m} \mathbf{E} \equiv \sigma \mathbf{E} \quad \text{or} \quad \mathbf{E} = \rho \mathbf{J},$$

where σ is the conductivity, and $\rho = 1/\sigma$ the resistivity, of the material;

$$\rho = \frac{m}{\tau q^2 n}.$$

Note that ρ is inversely proportional to the relaxation time. Since we expect that τ decreases with increasing temperature (carrier collisions occur more often), we expect a metal's resistance to increase with temperature (which it does).

Now, consider what happens when we apply a magnetic field \mathbf{B} to a thin strip of material in which a current I is flowing, as shown in Fig. 3.

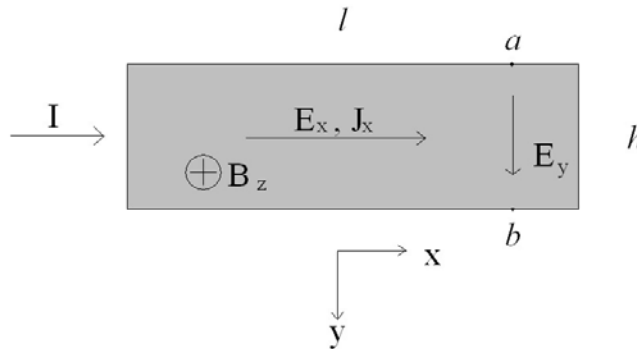


Fig. 3: Hall effect geometry again; the strip has a thickness δ , length l , and height h .

Here, the applied field \mathbf{B} is directed only in the z direction (into the paper). The x -component of \mathbf{E} drives the steady current I in the x direction, and a y -component of \mathbf{E} must appear to balance the Lorentz force on the charge carriers: The equation of motion (in SI units) is

$$\frac{d\mathbf{V}}{dt} + \frac{1}{\tau} \mathbf{V} = \frac{q}{m} [\mathbf{E} + \mathbf{V} \times \mathbf{B}].$$

Again, when a steady current flow is established the time derivative of \mathbf{V} at each point vanishes³. Multiplying the above by $\tau q n$, the time-independent result can be written in terms of \mathbf{J} as

$$\mathbf{J} = \sigma \mathbf{E} + \frac{\tau q}{m} (\mathbf{J} \times \mathbf{B}).$$

With the setup shown in Fig. 3 steady current can only flow in the x direction $\mathbf{J} = J \hat{\mathbf{x}}$, while $\mathbf{B} = B \hat{\mathbf{z}}$. Performing the cross product and resolving into components gives

$$\sigma E_x = J \quad \text{and} \quad \sigma E_y = \frac{\tau q}{m} JB.$$

Along with Ohm's law describing current flow along the strip, we now have a component to the electric field in the transverse direction — proportional to J and B . The Hall Coefficient (or Constant) R_H is officially defined as this proportionality constant:

$$E_y = R_H JB.$$

The Drude model thus predicts

$$R_H = \frac{1}{nq}.$$

The Hall constant thus gives a direct indication of the sign of the charge carriers; it is negative for electrons ($q = -e$) and positive for holes ($q = +e$). Note its independence of

³ Remember that \mathbf{V} is not the velocity of any given particle, but an average. Note that \mathbf{F} is a velocity-dependant force. When averaged, the single-particle velocity \mathbf{v} is replaced by the average \mathbf{V} .

the our model parameter τ : it turns out from a quantum mechanical analysis of the problem that this formula generalizes to a wide class of materials. Note also that the Hall field is large for low carrier concentrations: we might expect that the effect will be harder to detect in metals than in semi-conductors.

The Hall field E_y can be measured by the voltage difference between points a and b (Fig. 3) since $\Delta V_H = V_a - V_b = E_y h$, where h is the sample height. The total current flowing through the strip is $I = J \times (h\delta)$. Thus, in terms of laboratory quantities we have the equivalent definition of R_H in terms of the Hall voltage and the current:

$$\Delta V_H = \frac{R_H}{\delta} I B$$

Where, again, δ is the thickness of the strip. As we reasoned from the mechanics of Fig. 1, the Hall voltage-difference is negative for electron carriers and positive for hole conduction.

The Hall *constant* should not be confused with what researchers term the *Hall resistance* (often denoted by the same symbol used for the Hall constant). The normal resistance of the sample is just the voltage drop along the sample divided by I . In analogy, the Hall resistance (in ohms) is defined as $\Delta V_H / I$ — a kind of transverse resistance. Obviously, the Hall resistance is linearly dependent on B . However, in 2- D systems, at very low temperature and large fields, the Hall resistance show a step-like (rather than linear) dependence on B . These steps are completely independent of the type of sample and quantized to values $h/e^2 m$, where m is an integer. This is the famous Quantum Hall Effect⁴. The fundamental quantum of Hall resistance is $h/e^2 = 25,813 \Omega$. In this lab we will only be concerned with the Hall coefficient.

⁴ There is also a fractional quantum Hall effect.

In the SI system the Hall coefficient unit is (volt m)/(amp tesla) or $\Omega\text{m/T}$, however R_H is often reported in hybrid units such as (volt cm)/(amp gauss). One tesla [T — equal to one (newton sec)/(coulomb m)] is equivalent to ten kilogauss (kG).

Here is a list of values for some common elemental conductors

Element	R_H ($10^{-12} \Omega \text{ cm/G}$)	ρ ($\mu\Omega \text{ cm}$) @ 20C
Ag	- 0.8	1.63
Al	- 0.4	2.67
Au	- 0.7	2.20
Co	+ 2.5	6.34
Cu	- 0.5	1.69
Fe	+ 11.0	10.1
Mg	- 0.9	4.2
Mo	+ 1.3	5.7
Ni	- 5.1	6.9
Sb	+200.0	40.1
W	+ 1.2	5.4
Zn	+ 1.0	5.96

These values should be regarded as characteristic; real-world values of R_H are dependant on a number of parameters such as temperature, sample purity, and even magnetic field.

Determination of R_H and ρ provide some insight into the nature of conduction in solids. As an example, consider room-temperature silver with

$$\text{Ag: } \begin{cases} R_H = -0.8 \times 10^{-12} \text{ Vcm/AG} = -0.8 \times 10^{-10} \text{ Vm/AT} \\ \rho = 1.6 \mu\Omega\text{cm} \end{cases}$$

The negative Hall coefficient indicates that electrons are the charge carriers. From its magnitude, we can derive the carrier concentration

$$n = \frac{1}{|R_H|e} = 7.8 \times 10^{28} \text{ m}^{-3} = 7.8 \times 10^{22} \text{ cm}^{-3}.$$

We can compare this with the density of Ag atoms, which we find from its density of 10.5 g/cm^3 and mass number 107.87 g/mole . These values give an *atomic* density of $5.9 \times 10^{22} \text{ cm}^{-3}$. We see there is about one (1.3) conduction-electron/atom, more-or-less in keeping with what we might expect from Ag's $5s^1$ valence configuration.

This simple expectation that atomic valence electrons provide the conduction carriers breaks down for most metals, *e.g.* a similar calculation for Be (valence $2s^2$) yields 0.2 holes/atom!

The relaxation time can be found from the resistivity and the carrier concentration

$$\tau = \frac{m}{\rho e} |R_H| = 2.8 \times 10^{-14} \text{ s}.$$

Here, in keeping with the free-electron model, we've used the electron mass rather than an effective mass (see App. A). This is the *average* time between scattering events for the electrons conducting the current.

This gives us some insight into the scattering mechanism (at least within our Drude model). The mean free path, ℓ , is the average distance an electron travels between collisions. We can estimate this from the average thermal velocity of an electron ($mv_T^2 = 2kT$). For room temperature (300K) this is $v_T = 9.5 \times 10^4 \text{ m/s}$, and so

$$\ell \approx v_T \tau = 2.7 \times 10^{-9} \text{ m}.$$

Hence the electron travels about 27 \AA between collisions. This is larger than our expectation of the Ag atomic spacing ($\approx n^{-1/3} = 2.5 \text{ \AA}$). The quantum mechanical model shows that the principle scattering mechanism is not the atomic lattice sites, but the interaction with vibrations of the lattice (phonons).

What is the drift velocity of the carriers? This depends on the electric field applied. Suppose we apply a familiar field, say 100 V/m – this is the field we would have by placing a 1 V potential difference across a 1 cm length of material. The drift velocity in this case would be

$$V = \tau q E / m = 0.49 \text{ m/s} !$$

Notice how slow the average drift is to the thermal speed. This gives some justification in defining τ independently of \mathbf{V} . The applied forces barely perturb the actual trajectories, so that Fig. 2 is highly exaggerated.

3. Measuring the Hall Coefficient in Metals

On the practical side of things, consider actually measuring the Hall Voltage using the geometry of Fig. 2. To this end we have a number of metallic foils soldered to specially constructed circuit boards. Figure 4 shows our aluminum sample obtained from a piece of kitchen foil (burnished to thin it as much as possible):

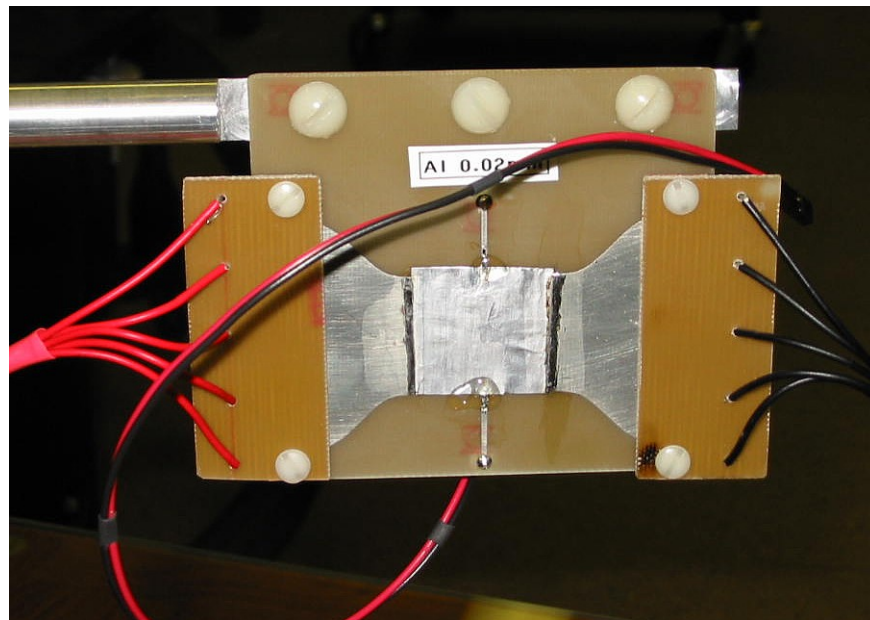


Fig. 4: Hall-effect sample mounted on its board. The sample is about 25mm square with a thickness of 0.02 mm.

A foil current is passed through the sample (left to right in the photo) *via* the 5-wire harnesses. This arrangement keeps the current flow in the foil spatially uniform. The transverse Hall voltage is measured on the twin red/black leads.

The measured voltage will be small. Our magnet attains a field strength of about 10 kG. For a foil thickness $\delta = 0.02\text{mm}$ then we expect a Hall voltage on the order of $20\ \mu\text{V}$. This is a very small potential to measure and we employ a very good DVM (Fluke model 45) on its most sensitive scale (resolution of $1\ \mu\text{V}$).

In measuring the potential difference between points a and b (Fig. 2) the voltage probes will not be precisely aligned vertically. The actual situation will be more like that of figure 4:

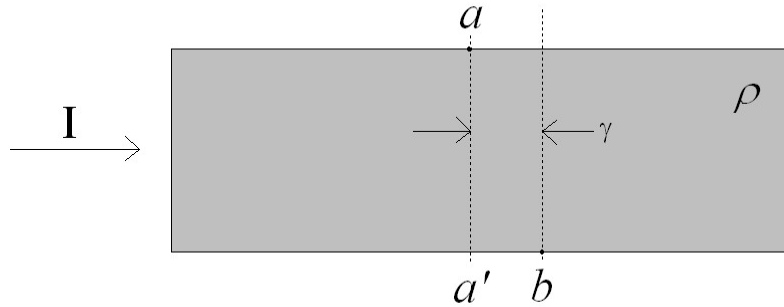


Fig. 5: Horizontal offset (γ) of Hall voltage measurement points.

The resistance of the area between a and b is $R = \rho\gamma / h\delta$, so that a potential difference IR exists *along* the strip. Thus, an offset voltage (proportional to the current) occurs in our Hall measurement:

$$\Delta V_{a,b} = \frac{R_H}{\delta} BI + RI.$$

Using again our Al example, if there was an effective mismatch of the probes by 0.5 mm, the resistance would be $27 \mu\Omega$ causing an offset $270 \mu V$ at 10 A foil current. This offset is large compared to the actual Hall voltage. Our scheme will be to fix the foil current I , and scan the magnetic field. A linear fit of $\Delta V_{a,b}$ vs. B will then have a slope $R_H I / \delta$.

Procedure

1. Read appendices **B** and **C**, which will give you an overview of how the magnet and scan are controlled.
2. Choose a metal-foil sample, and attach it to the aluminum sample holder by two or three nylon screws. Attach the five-wire current leads to the circuit board (copper side to copper side) as in Fig. 3 (also with the nylon screws provided). Be sure to maintain the proper red/black polarity. Each sample board comes with attached red/black leads for the Hall voltage measurement. Place the sample and holder between the pole faces of the magnet — make sure the sample doesn't interfere with the B-field probe located to the rear of the space. The holder rides on the optical mount running beneath the magnet. When you get the sample placed, tighten the base down securely.

The Hall voltage is measured by the Fluke 45 multimeter. A specially-made test lead is needed to connect to the sample board's Hall leads: it has red/black banana plugs on one end (connecting to the Fluke), and a female micro adapter on the other which mates with the male plug of the board leads. Using this lead, connect the Board to the meter as shown in Fig. 6:

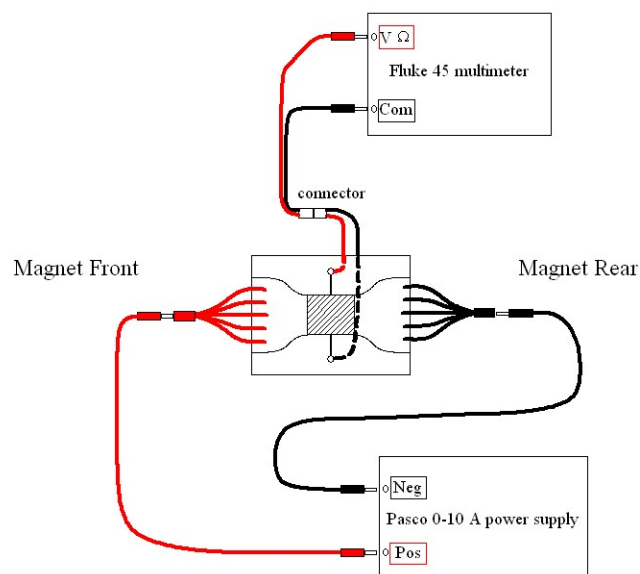


Fig. 6: Setup wiring for metal-foil samples.

Using banana-jack cables connect the Pasco power supply which will provide the foil current (Fig. 6 again). With the sample board orientation and electrical polarities indicated (and the magnet current in the 'Normal' setting), the experiment will have the same geometry as in Figs. 2 and 3. You should satisfy yourself that this is so.

3. Next, connect the USB cables to your PC: There are two (see Fig. B5), one from the Gaussmeter A/D box, and one shared by the magnet power supply and the Fluke meter. Turn on these instruments and deal with the usual Windows problem of finding 'new' USB devices.

The Sorensen magnet power supply is disabled until the magnet cooling water is turned on. This is accomplished by opening the ball valve on the copper plumbing on the wall behind the magnet.

4. Now, you must establish a constant current through the foil: On the Pasco power supply, zero the voltage- and current-limit knobs (counterclockwise all the way). Turn on the supply and turn the voltage adjust all the way up – you should still see zero current on the indicator. Next carefully turn up the current limit knob until the current meter reads your desired current. (You should take scans for at least three different currents: *e.g.* 6, 7.5 and 9 Amps or some similar values.) The supply should now maintain a constant current through the foil while the Hall scan is underway.

5. You are ready to take a scan (have you read appendix C yet?)! Run the HallScan program (which may or may not be on your desktop). After initiation the code sits idle until you run a scan. Mostly, the default scan parameters (2 to 10 kG) are reasonable however, it's best to increase the number of points to perhaps 10. For the nickel sample, even more points are preferable. Click the scan button and relax.

Currently, we have metal-foil samples mounted of Cu, Al, Fe, Ni, and Mo. You should do them all (with a variety of foil currents for each). Be careful to always maintain the standard setup geometry, and also make sure you have an adequate naming scheme for your output files.

Analysis

1. Each scan, when $\Delta V_{a,b}$ is plotted against B_{meas} , should result in a straight line with a slope of $R_H I / \delta$ and an offset RI . The nickel sample will be an exception, see part 4 below.

So, perform a linear fit to all your to data files. The easiest way to accomplish this is simply to import them into Excel, plot them, and add a linear trend line. This will give you a fitted slope and y-intercept. Using the foil current I and the thickness δ (listed on each sample) compute the Hall constant R_H and offset resistance R . For a given sample, each data file should give a similar result – do you find any trend in R_H as a function of I ?

2. Your final ‘measured’ value of R_H for a given metal can be taken as the average value of the R_H for each foil current. Which samples conduct via electron and which via hole currents? Compare your results of R_H to the values listed at the end of section 2 (be careful of your units!).

3. For each metal, calculate from R_H the measured carrier density n . Look up each sample’s density, mass number, and valence structure and estimate what you think n should be (as we did for Ag at the end of section 2). Compare.

4. Ni and Fe are ferromagnetic! This means that within the foil are many crystal domains with oriented magnetic moments. When an external field (what we’ve been calling \mathbf{B}) is applied, domains aligned with the external field grow and a net magnetization \mathbf{M} results. The actual \mathbf{B} inside the sample is thus a sum of the external field and that of the induced magnetization. Notice that Fe’s Hall voltage shows a fairly linear dependence on B , but that Ni begins to curve over. These differing behaviors are due to the fact that the magnetization saturates at a much lower field in Ni (6 kG) than in Fe (21 kG).

4. Measuring the Hall Coefficient in Semiconductors.

The metal samples we've been dealing with up until now are (for obvious reasons) called conductors. According to the band theory of solids, conductors are substances which have partially filled valance bands. In such bands the electrons are free to respond to external forces, and a macroscopic current is easily established. If the band is almost empty we have electron-like conduction, whereas if it is almost full we see hole-like conduction (appendix A). Recall that the resistance of metals is determined by thermal scattering of the carriers, and the resistance of a metal thus increases with temperature.

In contrast, insulators have completely filled valance bands. A filled band cannot support any macroscopic current since the electrons are 'frozen' in their individual orbitals by Pauli's exclusion principle. However, at temperatures greater than zero thermal agitation can promote electrons from a filled band into any empty bands, so that an insulator can have some free current carriers⁵. What determines the degree to which conduction is possible is the energy gap between the filled valance band and the empty (conduction) band. If this *band gap* is small compared with kT there will be a lot of promoted electrons available for conduction. For band gaps much larger than kT there will be very few conduction electrons, and we have a very good insulator.

Semiconductors are insulators with intermediate energy gaps. Silicon, for example, has a band gap of about 1.1 eV. At zero temperature *pure* Si would be a perfect insulator. At room temperature there is enough thermal excitation to promote about 10^{10} conduction electrons/cm³. This sounds like a lot, but recall there are on the order of 10^{22} atoms/cm³ in the solid, so there is a conduction electron for every 10^{12} atoms! In fact *pure* Si at room temperature is a pretty good insulator. At higher temperatures more and more valance-band electrons are promoted, *e.g.* at 200 °C the carrier concentration has risen to 10^{14} cm⁻³. This increase in carrier concentration with temperature generally overshadows the increased carrier scattering, so semi-conductors have the reverse behavior of metals: their resistance *decreases* with temperature.

⁵ Both holes and electrons.

Notice in the proceeding discussion that the word pure has been emphasized. In practice we can never attain a purity anywhere near the intrinsic carrier concentration of one part in 10^{12} . In fact, even a tiny impurity concentration completely alters the (room temperature) electrical characteristics of semiconductors from their intrinsic behavior. In the early days of semiconductor research, this fact caused a lot of conflicting results and confusion.

Impurity atoms (of the right type) act as sources of free charge carriers: When introduced into the semiconductor crystal, they form energy states only $\approx 10^{-2}$ eV removed from either the conduction or valance band. A donor impurity is located just below the conduction band and can be easily ionized thermally, donating electrons to the conduction band (*n*-type semiconductors). Likewise, acceptor impurities lie just above the valence band and provide a place for valence electrons to go. Acceptor impurities thus provide a way of introducing hole carriers (*p*-type semiconductors). By carefully controlling the concentration and type of impurities, engineers can fabricate crystals with desired electrical properties. As an example, Fig. 7 displays the resistivity of Si as a function of impurity concentration:

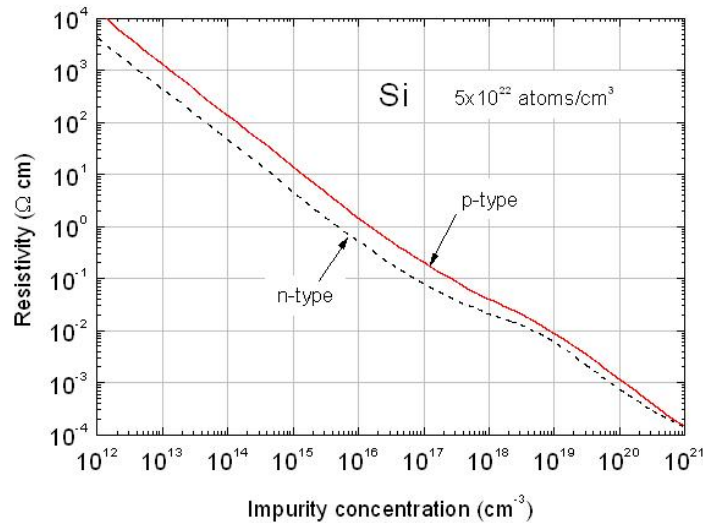


Fig. 7: Room temperature Si resistivity as a function of *n*- or *p*-type impurity concentration, over a range of 1 in 10^{10} up to 1 in 10. For comparison, recall the *intrinsic* carrier density is on the order of 10^{10} cm⁻³, and a metal's resistivity is $\approx 10^{-6}$ Ωcm.

Because the density of charge carriers (n) can change so readily with temperature, it's customary in discussing semiconductors to factorize the conductance as

$$\sigma = \mu n,$$

where μ is the *mobility*⁶. The mobility is an important property of the semiconductor, which we will not be concerned with — except to point out the following: By measuring both the hall constant (giving n) and the conductivity (or resistivity) the mobility can be determined. These pairs of measurements form an important and very common diagnostic procedure in both research and industry⁷.

Our measurement of the Hall effect in semiconductors is similar to our procedure for metals, with only a few changes. Firstly, the samples are more crudely mounted:

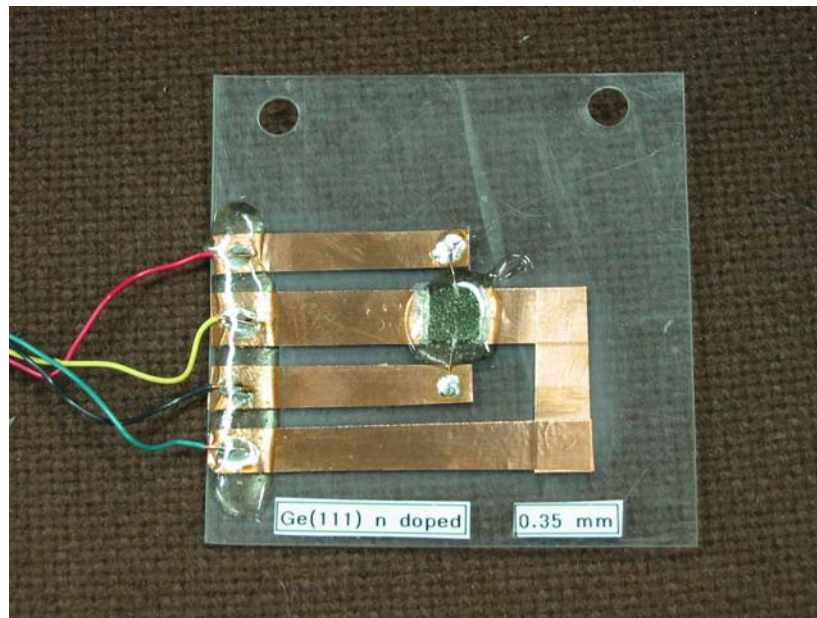


Fig. 8: Typical semiconductor sample.

⁶ Actually, since the current is can be comprised of both hole and electron movement, there is a hole and electron mobility to consider and the hall constant becomes a little more complicated to interpret. In this lab we will consider **only** materials which are doped with either donor or acceptor impurities—so the samples will only conduct *via* a single carrier type.

⁷ See for example the NIST website <http://www.eeel.nist.gov/812/hall.html> for a lot more detail.

Our samples are Si or Ge crystals soldered (using Indium solder) to copper tape strips. An epoxy coating is applied for protection. Since the semiconductor samples are much more resistive than the metal samples we can only pass a small current through them. For this purpose we use the 0-20 mA constant current source instead of the Pasco power supply to provide ‘foil’ current. For the same reason the resistance is high (*i.e.* low n), we will be measuring much larger Hall voltages, however the Fluke meter will handle this automatically.

Procedure and Analysis

The Fluke voltmeter and the current source hook to the samples using different leads than before. They are banana connectors on the instrument end and gator clips on the sample end. Figure 9 shows the hook-up scheme corresponding to our standard geometry.

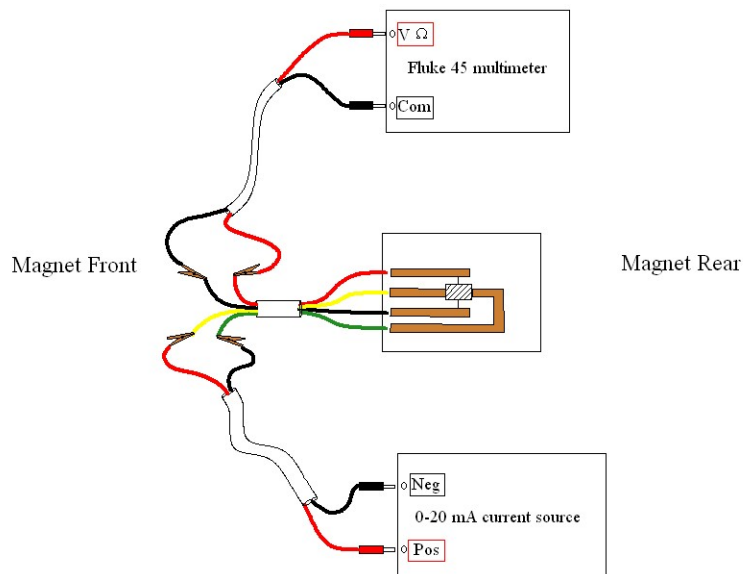


Fig. 9: Setup wiring for semiconductor samples.

While the wire colors are standardized on all the samples (so far) it's best to think about what you're doing with each setup.

Now the procedure and analysis goes exactly the same as for the metals. For each semiconductor sample, you should do a Hall scan for a relatively wider range of currents; perhaps 10, 5, 1, 0.5, and 0.1 mA. With semiconductors it's often important to make measurements at as low a current as is possible (because of thermal considerations). See how low you can go in current and still get a reasonable measurement.

APPENDIX A: Electron and hole conduction.

We are all well aware (nowadays) that electrical conduction in solids is due solely to the movement of electrons. A major catastrophe for the free-electron model of solids was the observation of positive Hall constants: how can there apparently be a motion of positive charge carriers in a metal? The mystery is resolved by application of quantum theory to the problem, leading to the idea of electron ‘holes’. Holes are a way of describing the *collective behavior of many electrons* by introducing fictitious particles.

The concept of an electron hole is, in fact, a subtle one. Here, we explore the idea with the aim of comforting ourselves with a little fuller understanding. To do so, we must first outline some of the most salient features of the quantum approach.

When we solve for electron wave functions in a periodic potential we find energy eigenstates $\varepsilon_n(\mathbf{k})$ which belong to various *bands*—each labeled by an index n . Each band has a maximum and minimum possible energy. For each band the wave vector \mathbf{k} varies continuously over a \mathbf{K} -space. However, the *spatial* structure of the problem’s periodic potential dictates that the energies $\varepsilon_n(\mathbf{k})$ are periodic in the ‘*reciprocal*’ lattice of \mathbf{K} -space. Hence, only wave vectors \mathbf{k} that are confined to a so-called primitive cell (the 1st Brillion zone) of this lattice are needed to identify a unique state of the n^{th} band.

Now, let’s examine how electrons occupying the various states of a given band respond when external electric and magnetic fields \mathbf{E} and \mathbf{B} are applied. To do so, we construct wave packets localized in space near $\mathbf{r}(t)$, and in wave vector near $\mathbf{k}(t)$. The velocity of such a localized wave function is given by the group velocity. The result is a semi-classical orbit $[\mathbf{r}(t), \mathbf{v}(t)]_n$ associated with energy $\varepsilon_n(\mathbf{k})$ and wave vector \mathbf{k} at time t . The band index n is a constant of the motion. The wave-packet’s equation of motion is

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}(t) = \mathbf{v}_n(\mathbf{k}) = \frac{1}{\hbar} \nabla_{\mathbf{k}} \varepsilon_n \quad \text{and}$$

$$\hbar \frac{d\mathbf{k}}{dt} = \mathbf{F}_{\text{ext}} = -e[\mathbf{E} + \mathbf{v}_n(\mathbf{k}) \times \mathbf{B}].$$

Note that $\hbar\mathbf{k}$, the so-called crystal momentum, is not the actual momentum of an electron following the orbit since its time-rate-of-change is dictated only by the *external* force \mathbf{F}_{ext} . Effects due to *internal* forces (i.e. the periodic potential) have already been accounted for in solving the band structure $\varepsilon_n(\mathbf{k})$.

Now we can ask “How do electrons near either the top or bottom of a band behave when an external electromagnetic field is applied?”

First, consider the response of electrons occupying orbitals with energies near to the maximum of the band, ε_{max} . Near such a maximum, the energy can be estimated by the leading term of a Taylor expansion

$$\varepsilon_n(\mathbf{k}) \approx \varepsilon_{max} - \frac{\hbar^2}{2m^*}(\mathbf{k} - \mathbf{k}_m)^2,$$

where \mathbf{k}_m is the wave vector at which the band energy is maximum. The quadratic expansion coefficient is written in terms of the (positive) parameter m^* , called the electron effective mass⁸. How does a semi-classical orbit with \mathbf{k} near \mathbf{k}_m evolve in the presence of an external electro-magnetic force? The velocity becomes

$$\mathbf{v}_n(\mathbf{k}) = \frac{1}{\hbar} \nabla_{\mathbf{k}} \varepsilon_n \approx -\frac{\hbar(\mathbf{k} - \mathbf{k}_m)}{m^*},$$

and so the acceleration is

$$\mathbf{a} = \frac{d\mathbf{v}_n(\mathbf{k})}{dt} \approx -\frac{\hbar}{m^*} \frac{d\mathbf{k}}{dt}.$$

For orbits near the top of the band, the acceleration is *opposite* in direction to the wave vector’s change. From the semi-classical equations of motion, we have

⁸ More generally, the energy surface need not be symmetric, in which case the Taylor expansion is written in terms of an inverse effective mass tensor \mathbf{M}^{-1} .

$$(-m^*)\mathbf{a} \approx (-e)[\mathbf{E} + \mathbf{v}_n(\mathbf{k}) \times \mathbf{B}].$$

Thus, the time evolution of the orbit can be viewed as the motion of an electron with a negative mass or better yet: as a positive particle of positive mass m^* .

Now consider the behavior of orbits with energies near the band minimum. Now the dispersion relation is approximated by

$$\varepsilon_n(\mathbf{k}) \approx \varepsilon_{\min} + \frac{\hbar^2}{2m^\diamond}(\mathbf{k} - \mathbf{k}_m)^2$$

with some possibly different (but still positive) effective mass m^\diamond . Orbits at the bottom of the band thus behave in an electron-like way:

$$(+m^\diamond)\mathbf{a} \approx (-e)[\mathbf{E} + \mathbf{v}_n(\mathbf{k}) \times \mathbf{B}].$$

To summarize:

Fact I: *Electrons occupying orbitals near the band minimum respond to external fields like particles of charge $-e$ and mass m^\diamond . Electrons occupying orbitals near the band maximum respond to external fields like particles of charge $+e$ and mass m^* .*

We can now construct the electronic structure of the solid by filling orbitals with the available number of electrons using Fermi-Dirac statistics; an orbital (quantum numbers n , \mathbf{k} , and spin $s = \uparrow$ or \downarrow) is either occupied or unoccupied at time t by an electron. The total number of occupied orbitals must be constant.

The contribution to the current density associated with an electron occupying such a state is simply $(-e)\mathbf{v}_n(\mathbf{k})$. The finite contribution to \mathbf{J} due to a totally occupied region Ω of \mathbf{K} -space in the n^{th} band is

$$\mathbf{J} = (-e) \int_{\Omega} \frac{d^3\mathbf{k}}{8\pi^3} \mathbf{v}_n(\mathbf{k}).$$

The factor $1/8\pi^3$ is a density of \mathbf{K} -space states which appears when making a continuum approximation, and need not worry us. The integral implicitly includes a sum over the spin quantum states.

Now, it is not too difficult to show (using the time-reversal symmetry of the wave equation) that for every wave vector \mathbf{k} in the 1st Brillion zone, $\mathbf{v}(-\mathbf{k}, \uparrow) = -\mathbf{v}(\mathbf{k}, \downarrow)$. The immediate corollary is that the current density due to an *entirely occupied* band is zero:

$$(-e) \int_{\text{band}} \frac{d^3\mathbf{k}}{8\pi^3} \mathbf{v}_n(\mathbf{k}) = 0.$$

For a band which is only partially occupied, we can use this fact to write **Fact II**:

$$\mathbf{J} = (-e) \int_{\text{occupied}} \frac{d^3\mathbf{k}}{8\pi^3} \mathbf{v}_n(\mathbf{k}) = (+e) \int_{\text{unoccupied}} \frac{d^3\mathbf{k}}{8\pi^3} \mathbf{v}_n(\mathbf{k}).$$

Hence, the current density associated with a group of electrons occupying certain orbitals of a band is the same *as if* there were positively charged electrons (holes) occupying the unoccupied *electron* orbitals. **Furthermore**, in normal circumstances where the holes are near the band maximum, these hole orbitals respond to external fields like positively-charged particles (Fact I). It's in this way that nearly-filled bands exhibit hole-like current densities.

APPENDIX B: Our lab magnet

Specifications for the large Bell-Labs electromagnet in our modern physics laboratory have been lost long ago, however something *is* known about it in its present configuration.

The magnet coils are separated into four sub-coils on each side. To minimize voltage requirements these have been connected in parallel as shown below:

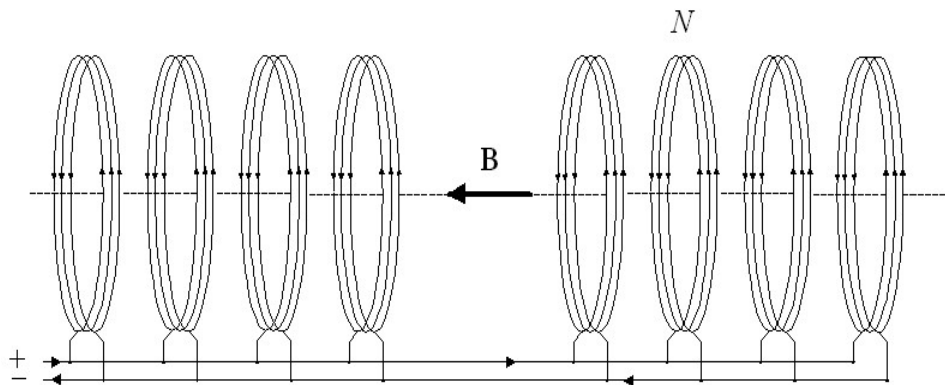


Fig. B1: Magnet coil configuration

There are N (unknown) turns per sub-coil with total DC resistance of $129\ \Omega$. For a *supply* current I – the current supplied by the power supply — there is a current of $i = I/8$ flowing in each sub-coil.

Our new Sorensen power supply can provide a current of up to 10 A at 200 VDC. Note that on the magnet's rear there's a reversing switch if you want the current (and thus the field) to go the other way. *Never switch the current direction unless the current is zero.* There is also an interlock which disables the power supply unless the magnet cooling water valve is on.

Now, consider the magnet in cross section as sketched in Fig. B2:

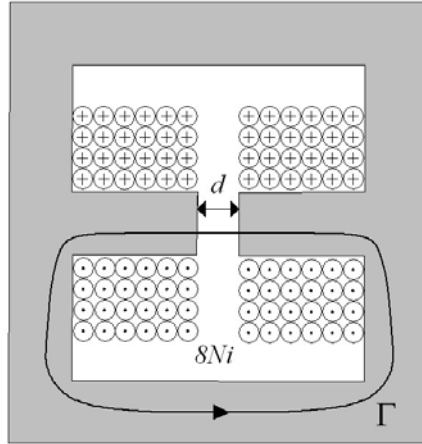


Fig. B2: Magnet cross section showing integration path Γ .

The sense of current flow through the coils is indicated by dots or arrow tails in the usual way. The spacing between the pole faces is $d \approx 4.3$ cm. We can gain some insight into the magnet's characteristics by applying Ampere's law around the closed loop Γ .

$$\oint_{\Gamma} \mathbf{H} \cdot d\mathbf{s} = I_{\text{enclosed}}$$

Recall that the magnetic field strength \mathbf{H} and the magnetic induction \mathbf{B} are related at a point in a material by the local magnetization \mathbf{M} ,

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}).$$

The physical significance of the field strength is that $\mu_0\mathbf{H}$ would be the induction produced using the same external arrangement of currents, but with the magnetic material absent. Thus, \mathbf{B} is a sum of two terms: $\mu_0\mathbf{H}$, the induction due to external current sources and $\mu_0\mathbf{M}$, the induction due to the local concentration of atomic dipoles. It's the \mathbf{H} -field that we have direct control over in the laboratory via the free currents I_{enclosed} of Ampere's law.

For simplicity, let's assume that all fields are constant around the integration path. In the air gap $\mathbf{H} = \mathbf{B} / \mu_0$, so that we can write

$$d \frac{B}{\mu_0} + L \left[\frac{B}{\mu_0} - M \right] \approx I_{\text{enclosed}},$$

where $L + d$ is the total length of the path Γ . The enclosed current is $8Ni = NI$ so that

$$B \approx \frac{\mu_0 NI}{L} + \frac{(L-d)}{L} \mu_0 M.$$

We see that the B -field within the gap is increased by the magnetization of the iron magnet core — in fact the magnetization is the dominant contribution (which is why the heavy core is used):

$$B \approx \mu_0 M.$$

Recall that \mathbf{M} is dependent on \mathbf{H} and to the past history (hysteresis) of the sample. For 'hard' ferromagnetic materials like our magnet iron, the hysteresis effects are small but noticeable. In this case we can roughly consider \mathbf{M} to depend only on the present value of \mathbf{H} . Since \mathbf{H} is proportional to I , we can also regard \mathbf{M} as a function of I , however $\mathbf{M}(I)$ is not a linear function:

For 'small' \mathbf{H} (and I) the susceptibility $\chi = M / H$ is large and constant so that the induction $B \propto I$. However as \mathbf{H} (and I) is increased further, all the permanent atomic dipoles eventually align and \mathbf{M} approaches a constant 'saturation magnetization' \mathbf{M}_s . For iron the $M_s \approx 1.7 \times 10^6$ A/m and the maximum contribution to B from the core magnetization is $\mu_0 M_s \approx 2.1$ T or 21 kG.

So, we expect that the induction B in the magnet's air gap is linear for small coil currents I , but then begins to flatten out as the core begins to saturate—somewhere on the order of a Tesla. In fact this is what happens:

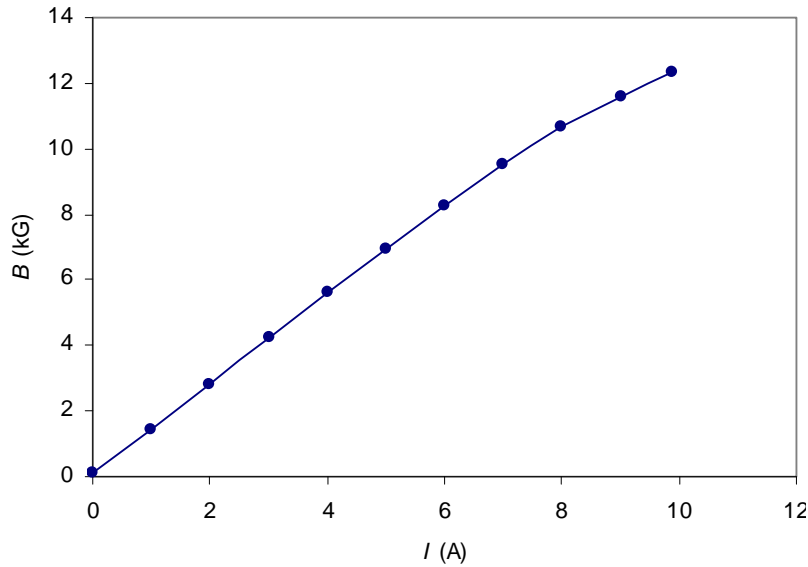


Fig. B3: Average magnet field verses coil current.

The figure A1 shows the plot of the induction (measured with a Hall probe) as a function of I . Note the turn over at larger fields.

At currents $0 \leq I \leq 6$ A the field is quite linear with

$$B \approx (1.381 \text{ kG/A})I + .107 \text{ kG} .$$

Note that there is a residual magnetization (and thus B) when the current is zero. As the magnet is used more and more with only positive currents we can expect this to grow.

For $I > 6$ A a good fit to the data is

$$B \approx (-0.0982 \text{ kG/A}^2)I^2 + (2.587 \text{ kG/A})I - 3.68 \text{ kG} .$$

These data points are the average of measurements taken while ramping the current up and then down. For any given coil current I , B is a little stronger on the way back down.

This hysteresis can best be described by considering the percent difference in the field between ramping down and up:

$$Diff = 200 \times \frac{B_{down} - B_{up}}{B_{down} + B_{up}}.$$

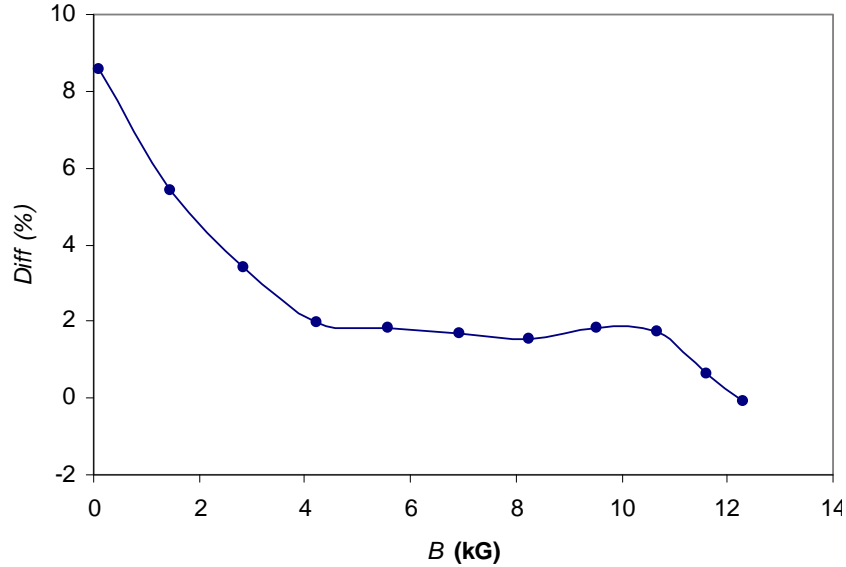


Fig. B4: Typical magnet hysteresis upon ramping current up and down.

The point of all this is that there is no exact relationship between the coil current we have control over, and the actual field produced between the pole faces. However, in controlling an experiment it would be preferable to work in terms of B . To this end we define a *nominal* magnetic field, dependent on I by the above fitted equations.

In setting the magnet to a nominal B field, the necessary coil current is calculated by inverting the fits. The controlling computer then sets the coil current via interface with the Sorensen power supply. In this way we have a *rough* control over the desired field, which is accurate to perhaps 5% depending on the past history of the magnet.

The actual field is measured using a solid-state hall probe located between the pole faces at the rear. The probe we use is a little cheap and non-linear, so that the Gauss meter's LCD reading (based on linear electronics) is only accurate to about 1%. However, this analog reading is in turn interfaced to the computer via a USB A/D unit, and the computer can deal with the non-linearity problem. Using our Bell Gauss meter as calibration, the computer outputs the measured B to about 0.1% accuracy.

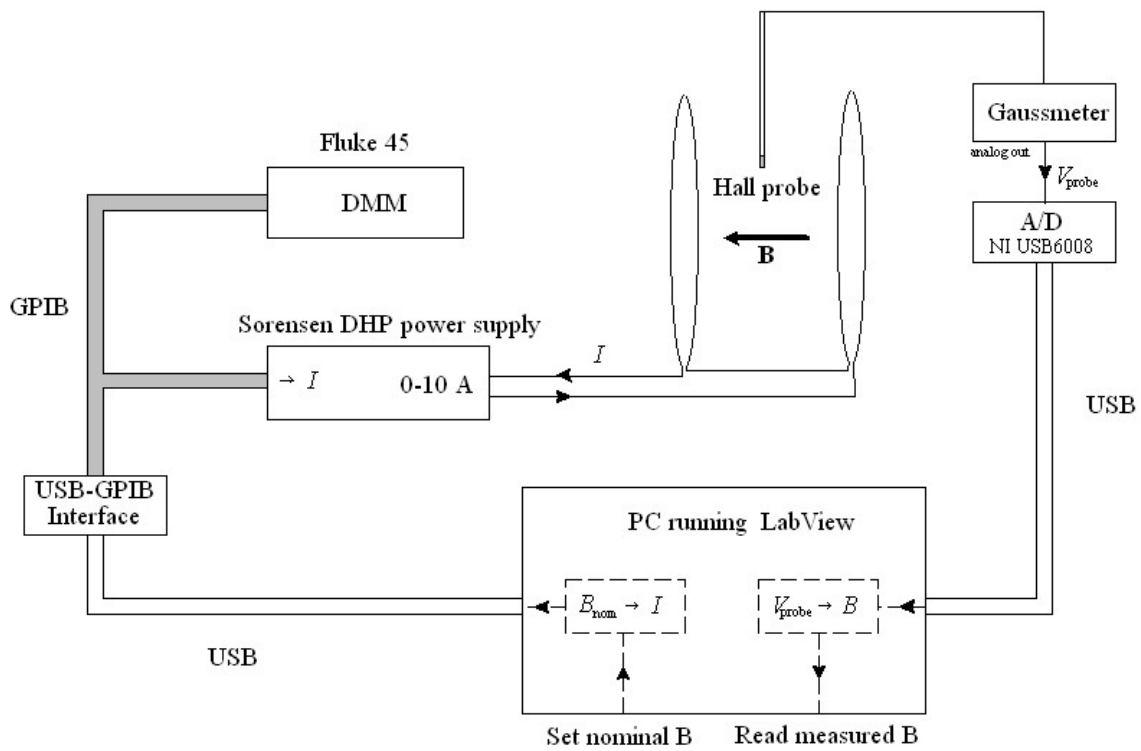
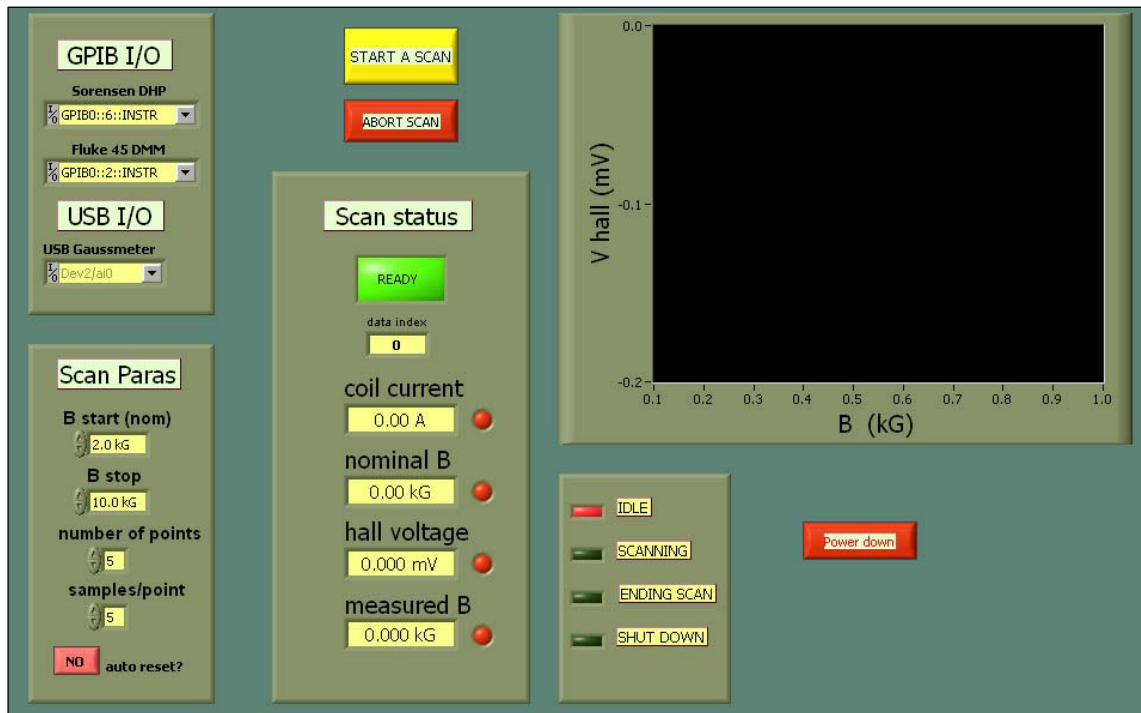


Fig. B5: Schematic plan for setting a nominal B and reading actual B .

APPENDIX C: The HallScan Program

This section briefly describes our LabView program for performing a Hall scan, *i.e.* measuring a sample's Hall voltage as a function of applied magnetic field. After startup and initialization you'll have the following window to control the experiment:



While initializing, an overlay window will briefly appear. This window informs you that the program is reading some calibration data about the magnet. (The file and path name is *C:\Magnetdata\B1calibration.txt*, and contains parameters enabling the code to convert between nominal B-field and the magnet's coil current.)

The lower-left panel is all there is to adjust: Simply set the start and stop values of the magnetic field, and the number of points. For example, a scan from 2.0 to 10.0 kG with 9 points will take data at $B = 2, 3, 4, 5, 6, 7, 8, 9,$ and 10 kG (nominal). At each of these B settings, the program samples the hall voltage a number of times determined by the *samples/point* parameter. The *auto-reset* toggle tells the computer whether or not to reset B to the starting value after the current scan is complete.

Once you have selected your desired scan parameters, simply click the start-scan button. A window will appear asking you to enter the *sample's foil current*; all this is

used for is to appended to the scan's output file. The scan status window gives you information about the progress of the scan, *i.e.* which point it's on and the measurements that have been completed so far (red indicators lighted). At the end of a scan, a window appears prompting you for an output filename; once entered, the data is written to disk and the program goes back to idle.

When you're finished taking scans, be sure to click the *power down* button and *wait* until the program ends before exiting. This ensures the coil current is ramped slowly down to zero.

The upper left panel deals with serial communication ID's for the various instruments used in the scan. In theory these will never need adjustment.

Here's an example of a typical output file:

Icoil	Bnom	VHall	stddev	Bmeas
1.3700E+0	1.9990E+0	2.9314E+0	5.4772E-4	1.8060E+0
2.0900E+0	2.9932E+0	2.9400E+0	0.0000E+0	2.7974E+0
2.8200E+0	4.0012E+0	2.9492E+0	4.4721E-4	3.8041E+0
3.5400E+0	4.9953E+0	2.9576E+0	5.4772E-4	4.7874E+0
4.2700E+0	6.0033E+0	2.9660E+0	0.0000E+0	5.7834E+0
4.9900E+0	6.9975E+0	2.9738E+0	4.4721E-4	6.7579E+0
5.7200E+0	8.0055E+0	2.9820E+0	0.0000E+0	7.7343E+0
6.5100E+0	9.0013E+0	2.9900E+0	0.0000E+0	8.7568E+0
7.3200E+0	9.9966E+0	2.9976E+0	5.4772E-4	9.7651E+0
8.2700E+0	1.1000E+1	3.0054E+0	5.4772E-4	1.0851E+1

foil current was 7.870000

This can be imported to an Excel worksheet for analysis. The 1st and 2nd columns give the magnet coil-current (A) and nominal B-field (kG). The Hall voltage's *average* (of the samples taken) is given in column 3 (mV), and its standard deviation in column 4 (mV). Finally, the last column gives the 'actual' magnetic field (kG) measured using the hall probe.