Hall Handbook

April 2016

Theory and practice of the usage of the apparatus to measure Hall effect, Energy gap and Resistance of a germanium sample over a wide range of temperatures.

## Introduction

Edwin Herbert Hall discovered the “Hall effect” in 1879 while working on his doctoral thesis in Physics investigating the influence of magnets on the resistance of a coil excited by a current. Hall discovered that a magnetic field would skew equipotential lines in a current-carrying conductor. This effect is observed as a voltage (Hall voltage) perpendicular to the direction of the current in the conductor.

The magnitude of this discovery is even more impressive considering how little was known about electricity in his time. The electron, for instance, was not identified until more than 10 years later.

The “Hall effect” remained a laboratory curiosity until the latter half of the XX century because the materials available, such as metals, would only produce small Hall voltages. With the advent of semiconductor technology and the development of various III-V compounds, it became possible to produce Hall voltages many orders of magnitude greater, allowing the production of Hall sensors, mostly made of indium antimonide (InSb), indium arsenide (InAs) and gallium arsenide (GaAs).

### A macroscopic approach to Ohm’s laws

The usual *macroscopic* approach to electrical conduction is based on the following experimental observations on metallic conductors:

1. The application of a steady voltage difference to a metallic wire produces a steady electric current proportional to . This holds true at least for small values of when the temperature of the wire does not increase appreciably.

* This allows the definition of the *electric resistance* as follows, according to the first Ohm's law:

1. For high currents the wire temperature increases and the power supplied by the generator to the moving charges, instead of accelerating more and more the circulating charge, is converted into heat (Joule effect).

* It could appear that the moving charges are subjected to some kind of force, like a body falling in a viscous medium, so that they reach a steady motion and give up part of their kinetic energy to the “body” of the wire (i.e. to the crystal lattice).

1. The resistance increases with increasing temperature.
2. Using wires of different length and sections the second Ohm's law determines the resistance:

* where the constant (the *electrical resistivity*) has a characteristic value for any material and increases with temperature.

The inverse quantity, the *electrical conductivity* , can be expressed using the two Ohm's laws as:

where is the current density and the electric field intensity. This is the starting relation needed to pass to a *microscopic* picture, that will allow a better understanding of the phenomena.

## A semiclassical microscopic model

The simplest microscopic model one can use is the “free electron gas" model of metals, in which the valence electrons are supposed to be practically free from their original atoms, and thus to move in the crystal lattice formed by the metal ions. In the absence of an applied electric field, the electron velocities are randomly distributed, with zero mean value and a *root mean square* value that may be evaluated from the equation:

where is the Boltzmann constant, the electron mass and the absolute temperature: at room temperature turns out to be of the order of .

Only when an electric field is externally applied the electron motion acquires an ordered component with a *mean value* (the *drift velocity*) which turns out to be very small with respect to as we will show later.

The drift velocity, i.e. this ordered component of the motion due to the electric field and to the the scattering of the electrons with the lattice, is simply proportional to the electric field intensity. The constant ratio between and (both in modulus) is called the *drift mobility* .

During a time of free motion between two collisions, the electrons increase their speed of the quantity:

where is electron charge. The kinetic energy of the electrons also increases, but it can be assumed that with each collision they loose additional energy. The transfer of such energy to the lattice ions explains the Joule effect.

In figure 1 it can be noticed[[1]](#footnote-24) that , after the application of the electric field, the average speed of the electrons is not zero but instead:

where is the mean free time between collisions[[2]](#footnote-25), so that the drift mobility has the microscopic expression :

Using these concepts of drift speed and mobility the current density can be written as:

where is the free electron concentration and relation (3) and (7) allow us to give a *microscopic definition* of the electrical conductivity:

Relation (9) tells us that all the physics of electrical conduction is described by the two parameters and .

In the electron gas model should be, for a monovalent metal:

where is Avogadro's number, is the atomic mass and the density. As an example for copper. Of course should not depend on the temperature.

A rough order of magnitude for the electron mobility may be derived using (7). A reasonable value for is: , where is the electron mean free path, of the order of the interatomic distance in the metal (i.e. a few ) so that, at room temperature, is of the order of .

Using the values of the elementary charge and of the electron mass , the electron mobility should be of the order of some .

Even for very large electric fields (up to ) the drift velocity is thus much smaller than .

Drift mobility should decrease with increasing temperature[[3]](#footnote-26) because of the increased thermal vibrations of the lattice ions. This effect can be studied by measuring the dependence of the electrical conductivity on temperature, but only if the free carrier concentration is the same at all temperatures.

In order to check experimentally the microscopic model we must measure not only the electrical resistance (which gives the product of and ) but also the free charge density : this can be obtained by performing a measurement of the Hall effect.

## The Hall effect

The Hall effect is essentially due to the Lorentz force acting on each electric charge moving with velocity in a magnetic field .

Let us consider a conducting bar (figure 1) immersed into a uniform magnetic field directed along the axis, with an electric current flowing along the axis. The Lorentz force on moving charges, both positive and negative, acts in the direction shown by the arrow (figure 1) (independently from the charge sign).

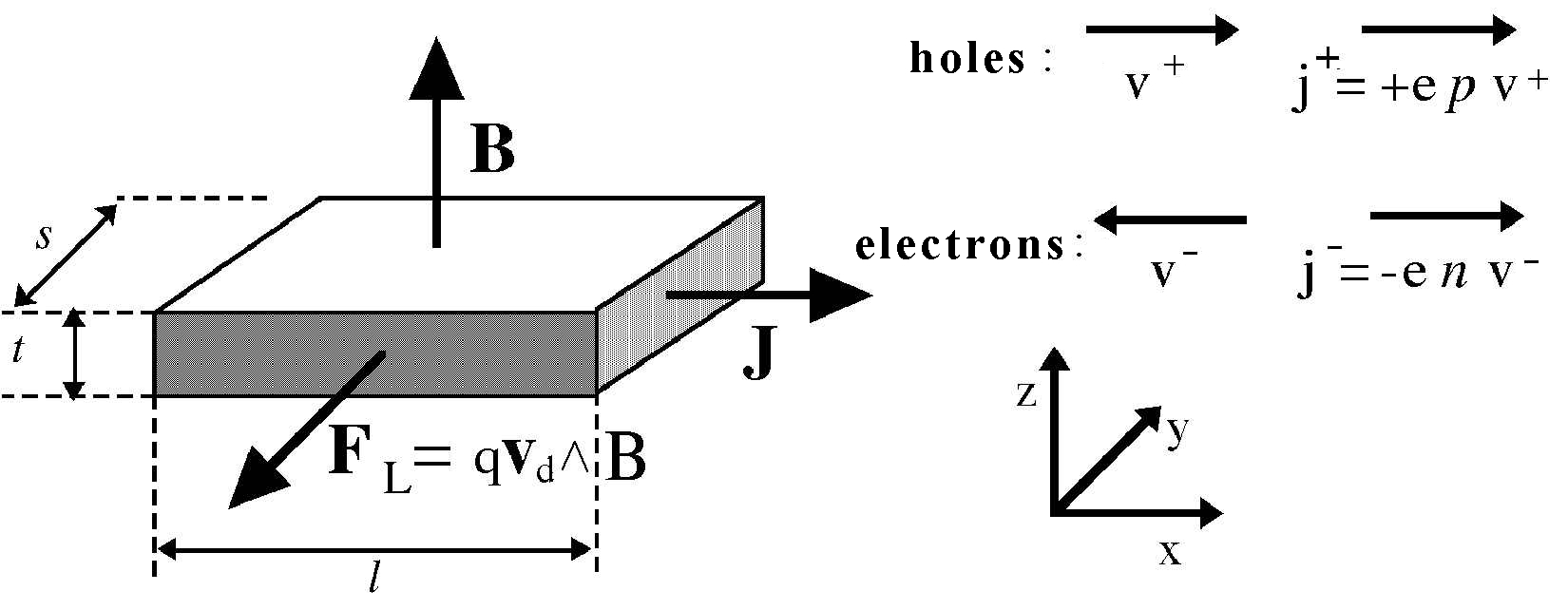


Figure 1. Hall effect geometry

In metals the electric current is only due to electrons. In semiconductors the charge carriers may be either electrons or holes.

In a pure semiconductor the electron density and the hole density is identical, in doped semiconductor we have (in N-doped material) or (in P-doped material). In doped semiconductors only one type of charge carriers is therefore important.

Let us consider first a metal or a N-doped semiconductor sample, where the relevant charge carriers are electrons.

In the electric field the electrons gain a drift velocity and they are subject to the Lorentz force , pointing towards the negative . While drifting in the direction they tend to crowd at the sample surface orthogonal to the axis and placed towards the reader in figure 1.

This charge density increase at the sample lateral surface produces a difference of potential along the axis and therefore an electric field . The value of the *Hall field* at equilibrium will correspond to an electric force equal and opposite to the Lorentz force, i.e. . This relation tells us that the Hall field is proportional both to the current density (through ) and to the magnetic field. It is therefore convenient to define the Hall coefficient as:

Recalling the relations (or ) we get :

or otherwise, for P-doped conductors:

Depending on the type of conductor, either metal (13) , N-doped (13) or P-doped (14).

Measuring we can determine the concentration of majority carriers and their sign (if we know the direction of the vectors ).

We can obtain relation (13) by assuming identical drift velocity for all charge carriers. This is an approximate relation, found in the literature:

where is a parameter that accounts for the statistical velocity distribution of the charge carriers, as well as the different scattering mechanisms: for mainly phonon scattering (lattice vibrations) and for mainly impurity scattering.

The Hall coefficient in semiconductors is many order of magnitude larger than the one in metals, due to the smaller charge density. This makes easier to measure Hall voltages in semiconductors, where a bias current of a few may conveniently generate a Hall voltage of in the order of a few .

To measure we must know , , and the sample thickness :

It is worth noting that the the Lorentz force direction does not depend on the charge sign.

The general expression for , valid (see Appendix 1) when *both electrons and holes* are present with densities and and mobility and is:

Which corresponds to the relations and (13) and (14) for or

When two types of charge carriers are present the electrical conductivity becomes:

The product , is named Hall mobility (note the capital index "" that distinguish it from hole mobility ).

For a doped semiconductor the Hall mobility approximates the *majority carriers* drift mobility :

From relation (17) we see that by increasing the temperature in a *P-doped sample*, generating many intrinsic[[4]](#footnote-29) carriers (i.e. electron-hole pairs), the Hall coefficient (which is positive at room temperature in the extrinsic region) tends to decrease, and it may even change sign. This is explained by the mobility ratio . *Note that this does not happen with a* N*-doped sample*.

From relation (17) at the temperature where (“*inversion point*”) we get , with and (where is the dopant density and is the thermally-generated charge density in the intrinsic zone). Therefore , or .

The intrinsic conductivity *measured* at the inversion point is:

In the extrinsic region of a P-doped sample, where the charge carrier density is constant , the conductivity is proportional to the carrier mobility : i.e.

The experimentally measured temperature dependence of the mobility is a power-law , where the exponent (in the range ) depends on the type of prevailing interaction of the charge carriers with phonons, lattice defects or impurities).

Therefore we may *extrapolate the extrinsic conductivity at the inversion point* , and from the ratio we get the value:

which can as well be written as:

where is the measured sample resistance at the inversion point and is the resistance extrapolated from the extrinsic region (low temperature) to the value it would have at the inversion temperature.

The dopant concentration is related to the value of the *Hall constant at the inversion point* (in the extrinsic region only the hole concentration is significant) by the equations (13) and (14), i.e. :

## The experimental setup

The apparatus uses a Ge sample, cut from a standard P-doped wafer, placed inside a isothermal aluminum case. It is placed in the gap between two poles of a permanent magnet, realized from two neodimium magnets and a U shaped soft-steel core, acting like a torus.

The sample has 7 wires tin soldered in the positions shown in figure 2 as follows:

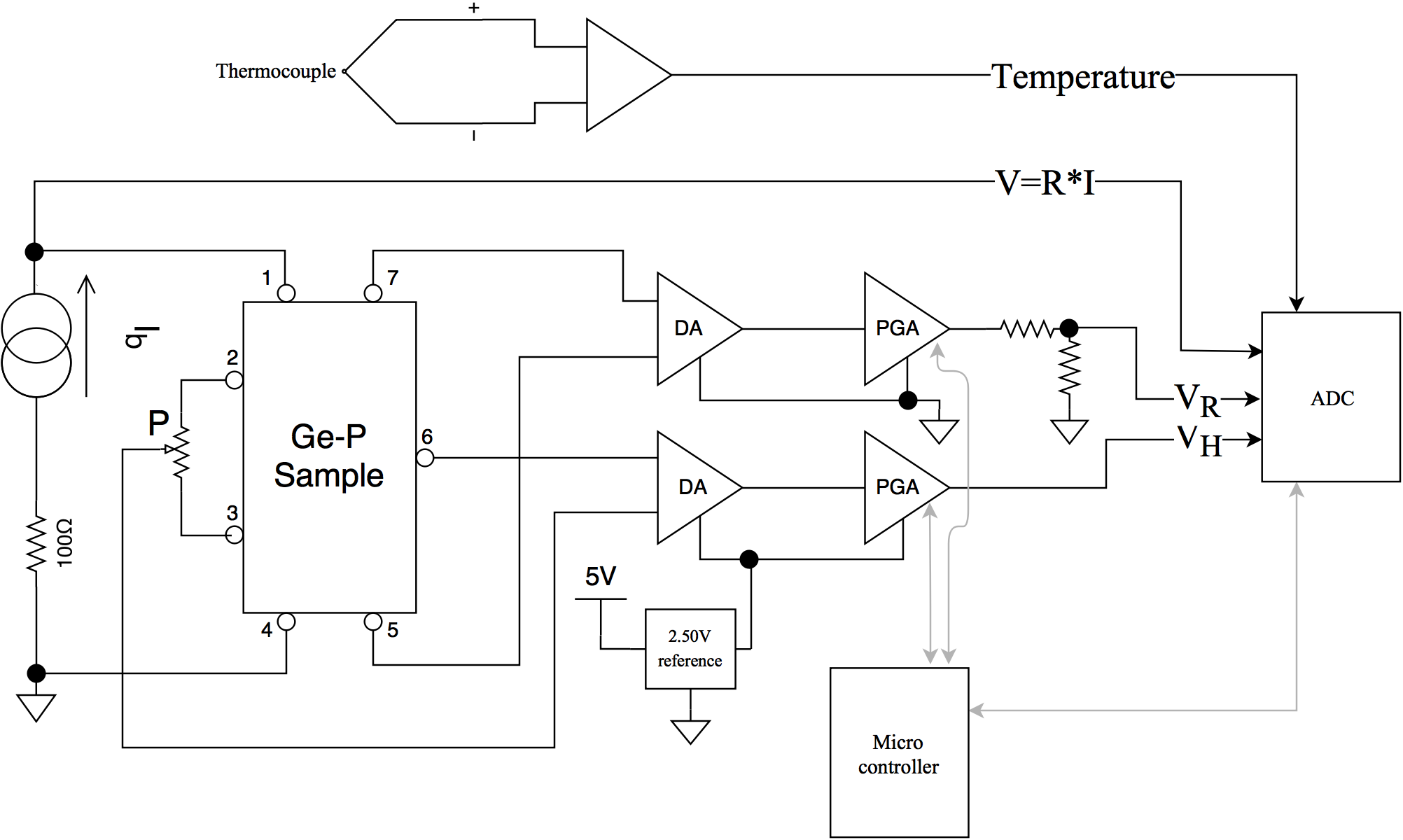


Figure 2. simplified schematic of the sample circuitry

* Contacts 1 and 4 (with Test Points named TP1 and TP4) are used to feed the bias current Ib produced by a constant current generator [fix, see figure x]
* Contacts 7 and 5 (TP7 and TP5) are used to measure (through a differential amplifier, DA for short) the voltage across the sample, in a 4-wire (often called kelvin) resistance measurement.
* Contacts 2-3 and 6 are the output of the Hall voltage and fed to the a second DA.
* Contact 6 (TP6) is the reference point for the Hall voltage and contacts 2 and 3 are used to set the balancing potentiometer P after having removed the sample from the magnetic field (the Hall voltage should be zero in absence of applied magnetic field). The potentiometer P is needed since the contacts 2 and 3 cannot be precisely positioned.

The DA outputs are amplified by Programmable Gain Amplifiers (PGA for short) whose outputs are referred to ground voltage in order to feed the signals to a data-logger.

The numbering of the contact on the sample corresponds to the number of the pins in the rj45 connector of the sample assembly.

The two DAs have fixed gains , set to for and to for [[5]](#footnote-32), and they're powered from a power supply.

The PGA gains are selectable among the following values through the front panel.

The output voltages on the front panel are restrained in a number of cases:

* If the input voltage is the DA saturates.
* If the output of the DA is not it is clamped down by a Schottky diode to prevent damage to the circuitry.
* If the output voltage of a PGA is not the PGA saturates

Saturation gives a warning (ERROR) message on the front panel.

The bias current is measured from the voltage drop across a resistor , giving thus an output of .

A digital-to-analog converter (DAC) acquires the analog values, sampling them from a to a integer value from to

The best value for the bias current is a compromise between the need to obtain a large to make measurements precise reducing the SNR (Signal to Noise Ratio) and a low self-heating of the element due to the Joule effect: and signal are proportional to while the Joule self-heating is

### Hall voltage and resistance measurements at room temperature

With a finite value of magnetic field B orthogonal to the large face of the sample, we must measure identical values for (but with opposite sign) when rotating of the sample. This behavior must be tested before proceeding to further measurements: if reversing the direction (i.e. rotating the sample of degrees) different values are measured, the offsets should be better adjusted using potetiometer P in figure 2.

The absolute value of may be varied by changing the width of the gap between the magnetic poles (see figure 3). One of the two permanent-magnet mounted on the soft-steel structure may be moved horizontally by turning the brass-screw: increasing the gap, the value of decreases.

[fix, update image]

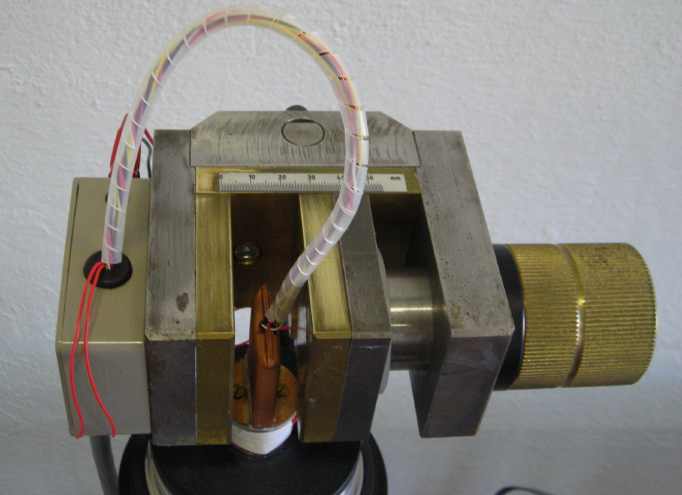


Figure 3. The screw device for changing the effective magnetic field

A calibration of the magnetic field as a function of the gap may be made using a gauss-meter probe placed between the poles (see figure [fix]).

The magnetic field may be calibrated using a gauss-meter probe placed at the center between the poles.

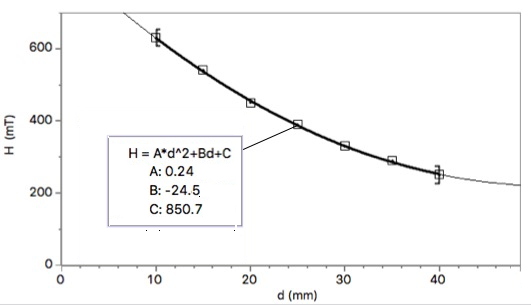


Figure 4. Measured values vs gap length

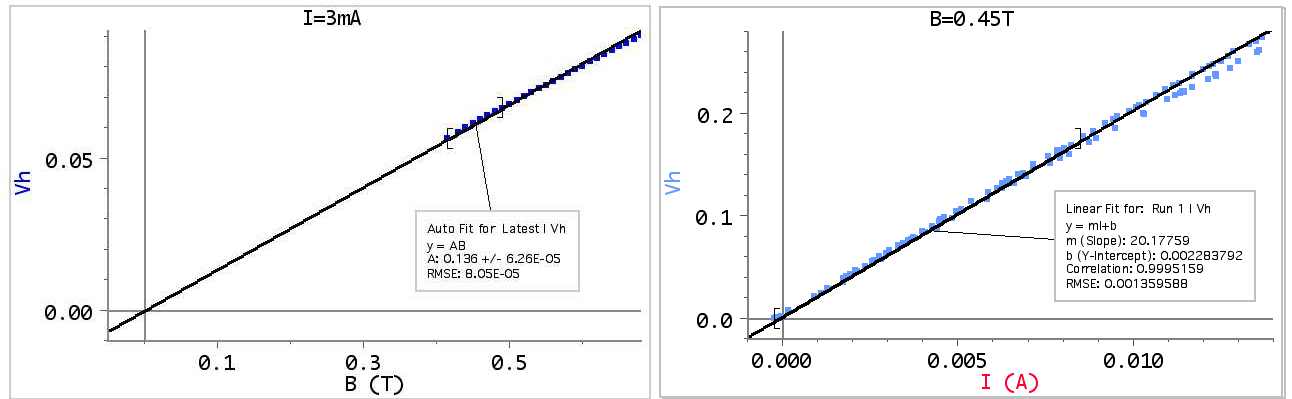


Figure 5. Hall voltage versus bias current (), and versus ( ).

Figure [fix] shows the linear dependence of on and . Deviations from linearity at high values of the magnetic field may be explained by calculating a second order approximation of (Appendix 1)

### Measurements at constant and while varying the temperature

The stainless-steel dewar can be filled of liquid nitrogen or a mixture of acetone and dry-ice (solid carbon dioxyde). The cold finger (the aluminum bar screwed into the base of the sample) is surrounded by the liquid nitrogen, allowing the sample to be brought lo low temperatures.

The temperature is measured by a tipe K (Chromel-Alumel) thermocouple thermally coupled to the sample. The small voltage generated by the thermocouple is amplified by the AD8495[[6]](#footnote-38) integrated circuit. The output is vaguely proportional to the temperature, with a sensitivity of . It is successively amplified by a non-inverting amplifier (not shown in picture [fix] ) to get with . While the K type thermocouple is fairly linear a room temperature, it is not in the wide temperature range covered by the apparatus, as can be seen in figure 6.

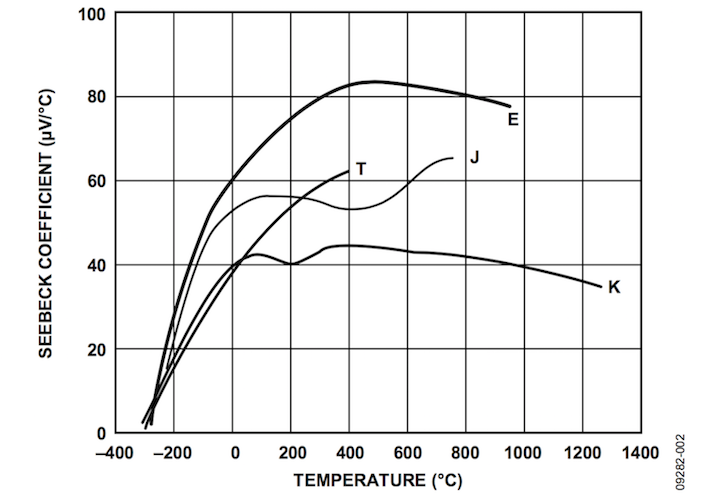


Figure 6. In every type of thermocouple, the Seebeck coefficiently is heavly non-linear in relation to the temperature.

In order to get a correct measurement it is necessary to compensate for the non-linearity (see figure 6) of the thermocouple linearly extrapolating the following polynomial:w

where is the output voltage of the thermocouple.

In the following table[[7]](#footnote-41) two different groups of coefficients are given, depending on the temperature range:

|  |  |  |
| --- | --- | --- |
| t range | -200°C to 0°C | 0°C to 500°C |
|  | 0.0000000E+00 | 0.000000E+00 |
|  | 2.5173462E+01 | 2.508355E+01 |
|  | -1.1662878E+00 | 7.860106E-02 |
|  | -1.0833638E+00 | -2.503131E-01 |
|  | -8.9773540E-01 | 8.315270E-02 |
| **E range** | **-5.891mV to 0mV** | **0mV to 20.644mV** |

Linearly extrapolating a polynomial (24) of the fifth order is more that sufficient given the precision of our equipment.

The voltage at the termocouple junction can be obtained[[8]](#footnote-42) from the following equation:

where is the output of the instrument (on the front panel), the voltage that indicates a temperature , is the error voltage at 0°C to achieve 125 mV at 25°C and is the internal gain of the AD8495 amplifier.

Consequently, linearly extrapolating with the compensation polynomial allows us to finally obtain the correct temperature[[9]](#footnote-44):

### Cooling-Heating procedure

To obtain accurate measurement it's suggested to first cool off the sample by placing liquid nitrogen in the dewar until a stable temperature is reached, then manually empty the dewar letting it slowly drift up to room temperature (due to the unavoidable thermal coupling sample-to-ambient).

During the cooling process measurements may be taken, but the rapid cooling generates a large thermal gradient across the apparatus that makes such measurements less reliable.

A resistive element is wound around the base of the sample, allowing to heath it up after reaching room temperature. The resistive element it's controlled in PWM: it's rapidly powered-on and powered-off with a duty cycle that can be selected (in percetange) on the front panel of the instrument.

[fix, image pwn]

This technique allows for a linear temperature sweep. Starting with a low power allows to better approximate an equilibrium situation, considering it can subsequently be increased if necessary to reach higher temperatures. The instruments automatically shuts down if .

Measurements can be obtained by feeding the analog outputs [fix, quali?] to a data-logger and setting a suitable acquisition rate (e.g. 1 sample per minute)

To obtain precise measurements, at least one complete hour is required for the whole temperature sweep.

### Typical results

The sample shown in this example has thickness , width and lenght . The measured Hall voltage and resistance vs temperature (with a constant bias current of and in a magnetic field) is shown in figure [fix].

In figure [fix] the resistance is calculated from data of figure [fix].

The logarithmic plot of the resistance versus (where is the Boltzmann constant) is shown in figure [fix].

From the slope in the intrinsic region we get the value of the energy gap , extrapolating linearly from , that can be compared to the known value for germanium (, cfr. Appendix 2)

[fix, insert image]

## Appendix 1: calculation of for small and high magnetic field

The motion equation for charge carriers can as well be written as:

where the charge is the for holes and electrons and we account for the mean time between collisions and for the Lorentz force. In stationary conditions the acceleration is zero. Therefore the velocities along ( is directed along ) for electrons and holes are respectively:

And, for velocities along y:

The current density along the axis can as well be written as:

where we made the approximation , neglecting here the Lorentz force. Recalling that , for small magnetic fields (33) may be approximated by:

For negligible current density along y we have:

or using and definitions:

If again we assume (neglecting, for small B, the correction for the Lorentz force we can write:

In this way the Hall coefficient becomes:

The formula (36) holds true only for *small values* of . For large values we must use (36) for the definition (33) instead of (34), obtaining for the Hall coefficient :

which tends to saturate at high B values.

### Appendix 2: Temperature dependence of

Experimental results consistently shows that the energy gap depends on temperature and for Germanium we can find in the literature the following empirical law:

This may be approximated, in the high temperature region, by a linear law as follows:

where the constants is the value of *linearly extrapolated* to :

Since in the intrinsic region (high temperature) the resistance depends on the absolute temperature as , a plot of vs using a linear approximation for results in a straight line with slope

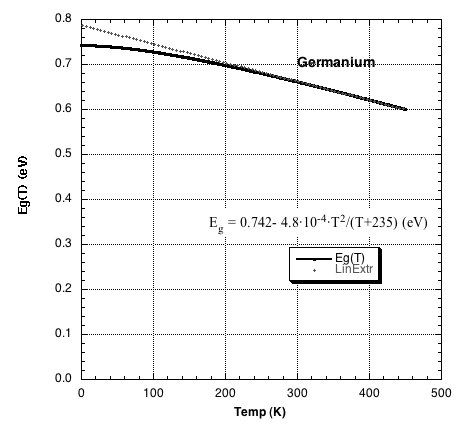


Figure 7. Temperature dependance of the energy gap

##Warnings

**Using high magnetic field require some caution:**

* You must avoid approaching any magnetizable object (clocks, electronic devices, screwdrivers...), which when brought too close may be permanently magnetized.
* A pinch hazard subsists if steel or other ferromagnetic material is placed near the magnets.
* Do not attempt to unscrew the magnets.
* The Magnets are brittle. A rapid shock with another magnet or ferromagnetic material may release shards dangerous for the eye.
* The apparatus **MUST NOT** be used by people with pacemakers.

##References

* J.C. Slater *Quantum Theory of matter*, mcGraw-Hill 1951.
* C.L.Chin e C.R.Westgate, *The Hall Effect and Its Applications*, Plenum Press, NY, 1979
* J.R.Hook , H.E.Hall *Solid State Physics*, John Wiley &Sons 1991.
* A. C melissinos *Experiments in modern Physics*, Academic Press, 1993.
* *New Semiconductor materials. Characteristics and Properties*, http://www.ioffe.rssi.ru/SVA/NSm/introduction.html (Electronic archive)
* *The Semiconductor informations WebSite* (properties of Germanium), http://www.semiconductors.co.uk/propiviv5431.htm

## Authorship

This Handbook was originally written by Giacomo Torzo of [Labtrek](http://labtrek.it)

Integrations and corrections by Davide Bortolami of [Fermium LABS](http://fermiumlabs.com)

1. See for instance *The Feynman lectures on Physics* vol.I 43-1,3 Addison-Wesley 1963. [↑](#footnote-ref-24)
2. This time does not depend on the electric field because the average speed increment due to the applied electric field is very small with respect to the r.m.s. speed due to thermal motion . [↑](#footnote-ref-25)
3. Drift mobility in semiconductors decreases with the absolute temperature as , where depending on the prevailing type of interactions of the free carriers (with phonons, lattice defects, or impurities). [↑](#footnote-ref-26)
4. *Intrinsic* term labels properties related to pure semiconductors or to doped semiconductors at hight temperature, where the thermally generated carriers density is much larger than the (*extrinsic*) carrier density due to the dopant. [↑](#footnote-ref-29)
5. The gain can change due to specifications and calibration. Please refer to the values calculated on the screen of the device. [↑](#footnote-ref-32)
6. [AD8495 datasheet, Analog Semiconductors](http://www.analog.com/en/products/amplifiers/specialty-amplifiers/thermocouple-interface-amplifiers/AD8495.html) [↑](#footnote-ref-38)
7. NIST t-90 tables for K type thermocouples, http://srdata.nist.gov/its90/download/type\_k.tab [↑](#footnote-ref-41)
8. [AN-1087, Analog Semiconductors](http://www.analog.com/media/en/technical-documentation/application-notes/AN-1087.PDF)   [↑](#footnote-ref-42)
9. [AN-1087, Analog Semiconductors](http://www.analog.com/media/en/technical-documentation/application-notes/AN-1087.PDF)   [↑](#footnote-ref-44)