The Hall Effect

Alexander Seaton

April 28, 2014

Abstract

The Hall constant of four different materials was measured: silver, tungsten, p-type and n-type germanium. This was achieved via the elimination of systematic errors caused by various effects. The constant for these materials was found to have values of $-7.4 \pm 3.7 \times 10^{-11} \mathrm{m}^3\,\mathrm{C}^{-1}$, $1.1 \pm 1.7 \times 10^{-10} \mathrm{m}^3\,\mathrm{C}^{-1}$, $6.9 \pm 0.3 \times 10^{-3} \mathrm{m}^3\,\mathrm{C}^{-1}$ and $-11.3 \pm 0.3 \times 10^{-3} \mathrm{m}^3\,\mathrm{C}^{-1}$ respectively, demonstrating the presence of positive charge carriers in n-type Ge and tungsten and negative charge carriers in p-type Ge and silver.

1 Introduction

The Hall effect is the deviation in path of an electrical current due to an externally imposed magnetic field, which produces a voltage across the material. It was first discovered by Edwin Hall in 1879[1] - notably before the discovery of the electron in 1897 or indeed the advent of quantum mechanical models of the behaviour of electrons in solids. Since then it has proved extremely useful and is used to measure the strength of magnetic fields in addition to the type and density of electrical charge carriers in materials (especially semiconductors).

In this experiment we sought to measure the constant of proportionality relating the Hall voltage and the magnetic field strength and current (the Hall constant) for four materials: silver, tungsten, p-type germanium and n-type germanium. To perform this accurately, various systematic errors needed to be removed, caused by effects described below.

2 Theoretical Considerations

The Hall effect can be quantified by considering a rectangular block of material, thickness t and width w which is positioned such that its length and width are in the $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ directions respectively. If a current is passed through it in the $\hat{\mathbf{x}}$ direction and a magnetic field is imposed in the $\hat{\mathbf{z}}$ direction, the charge carriers will experience a force causing them to move to one side of the material. This produces a build up of charge and thus an opposing electrical potential across the width of the material in the $\hat{\mathbf{y}}$ direction that balances the force from the magnetic field. With the assumption that charge carriers conduct in the material according to the Drude model, the following equation is obtained:

$$V_y = \frac{1}{nq} \frac{B_z I_x}{t} \tag{1}$$

The constant of proportionality relating V_y and $B_z I_x/t$ is known as the Hall coefficient, R_H, and in this case is:

$$R_H = \frac{1}{nq} \tag{2}$$

There are two important points when considering this expression:

- The dominant charge carrier's polarity is indicated by the sign of R_H.
- R_H varies as the reciprocal of the dominant charge carrier's number density - for small carrier densities a larger effect is observed. This is useful for measuring the dopant concentration of semiconductors.

In addition to the Hall effect there are several other effects present in this situation. These are the misalignment voltage and the Ettingshausen, Nernst and Righi-Leduc effects:

2.1 Misalignment Voltage

The Misalignment voltage is caused by the voltage probes not being placed on opposite sides of the material. Due to electrical resistance an additional voltage is measured between the two probes when a current is passed along the sample:

$$V_M \propto I_x \tag{3}$$

2.2 The Ettingshausen Effect

This produces a temperature gradient along the $\hat{\mathbf{y}}$ direction, quantified as[2]:

$$\Delta T \propto \frac{I_x B_z}{t} \tag{4}$$

The temperature difference across the sample's width results in a thermocouple voltage between the probes. This is dependent on exactly the same quantities as the Hall effect, so it is difficult to distinguish the two.

2.3 The Nernst Effect

This produces a potential difference across the width of the sample when a thermal current flows in the $\hat{\mathbf{x}}$ direction[2]:

$$V_N \propto w_x B_z \tag{5}$$

Where w_x is thermal current density in the $\hat{\mathbf{x}}$ direction

2.4 The Righi-Leduc Effect

This produces a temperature gradient in the $\hat{\mathbf{y}}$ direction caused by a thermal current in the $\hat{\mathbf{x}}$ direction and a magnetic field in the $\hat{\mathbf{z}}$ direction[2]. As with the Ettingshausen effect, a thermocouple voltage is produced across the sample:

$$\Delta T \propto w_x B_z w \tag{6}$$

Where as before w_x is the thermal current density in the $\hat{\mathbf{x}}$ direction and w is the width of the sample.

With the exception of the Ettingshausen effect, it is possible to eliminate these effects by taking measurements in which either the current or magnetic field is reversed due to their parity in I_x and B_z. The combination of measurements required is shown below[2]:

$$V_H(B,I) = \frac{V(+B,+I) - V(+B,-I) + V(-B,-I) - V(-B,+I)}{4} \tag{7}$$

3 Experimental Method

The samples were used of dimensions listed in table 1 in appendix B, measured using a micrometer. The thickness of the sample is purposefully chosen small to increase the measured voltages.

The samples were positioned between two electromagnets allowing control of the magnetic field. The magnetic field in the region to be occupied by the sample was measured with a Hall probe and found to be uniform to the resolution of the probe (1mT), thus effects due to non-uniform magnetic field could be discounted. It was not possible to measure the strength of the field perpendicular to the $\hat{\mathbf{z}}$ direction defined previously, however it is expected that these fields would be small and ought to have a significantly smaller effect on the measured voltage.

To control thermal effects, a fan and thermocouple attached to each sample was used to monitor and control the temperature of the samples during measurements. This could only prevent heating greater than 5 K. In addition when current was being varied, random values were chosen to prevent the current systematically being kept high and causing cumulative heating. This was particularly important for the metal samples which were subjected to much higher currents and experienced greater heating. A further consideration was the heating of the coils, which became a significant factor for large magnetic fields due to the large currents required. The coils were thus cooled by a further two fans.

At each point sampled from current-magnetic field parameter space, measurements were made of the current, magnetic field, voltage across the width of the sample and temperature. For the two semiconductor samples, the current through the sample was held constant and the magnetic field varied, whereas the opposite was done for the metal samples. This was decided in particular for the metals to reduce sustained heating.

4 Data Analysis

The data recorded is sampled from (I_x, B_z) parameter space and consists of runs in which one of the variables is held constant whilst the other is varied. For the following discussion, the variable held constant is referred to as Y and the varied quantity as X. For each run with Y a run was performed for $-Y$ so as to allow combining these data as described in equation 7. However the X values selected were not necessarily symmetric about zero and were different for different runs i, necessitating some method of combining these assymetric values of X.

To do this each run was separated into two series containing X of fixed sign - e.g. all $X > 0$ or all $X > 0$. These different series were then transformed so as to make equation 7 the arithmetic mean. This is done using these transformations:

$$Y' = |Y| \qquad X' = |X| \qquad V' = V\frac{XY}{|XY|}$$

A quadratic was then fitted to each series i to allow each to be summarised by a vector of coefficients, $\boldsymbol{\beta}_i$. This has the benefit of smoothing out statistical noise, however it is possible that information of importance is also eliminated or indeed that the quadratics introduce false trends into the data. To mitigate these concerns, the quality of the fit with all series measured was carefully monitored. It was clear from the results that the quadratic fits appeared to model the data extremely closely justifying their use. To retain information about the statistical variation of the data, the maximum error in the voltage measured in each series was retained to indicate the confidence limits of its quadratic fit (σ_i). In other words the error at each point on the curve is σ_i:

$$V(X', Y_i') = \begin{pmatrix} X'^2 & X' & 1 \end{pmatrix} \cdot \boldsymbol{\beta}_i \pm \sigma_i \tag{8}$$

Following this, quadratic fits with the same value of Y_i' were grouped and combined to eliminate their individual systematic errors by taking the mean of their coefficient vectors giving a reduced set of corrected quadratic coefficients, $\boldsymbol{\beta}_j^{\mu}$. This is possible as the quadratic equation is linear in the coefficients, whilst the transformation above ensures that equation 7 is simply the mean. In addition, their respective errors were combined using the standard error propagation formula for a linear weighted sum to give σ_j^{μ}.

At this stage we have hopefully eliminated the systematic errors from the data and may compare it against the theoretical Hall voltage - equation 1. This is accomplished by integrating the squared difference between the theoretical and observed curves weighted according to the error of the quadratic under consideration. This is essentially based on converting the the reduced Chi-squared statistic to act on continuous rather than discrete data points:

$$\chi^2_{\text{discrete}} = \frac{1}{N}\sum_{i=0}^{i=N}\frac{(O_i - E_i)^2}{\sigma_i^2} \rightarrow \chi^2_{\text{cont.}} = \frac{1}{x_1 - x_0}\int_{x_0}^{x_1}\frac{(O(x) - E(x))^2}{\sigma(x)^2} \tag{9}$$

Thus our goodness of fit measure G is:

$$G\left(\frac{R_H}{t}\right) = \frac{1}{N}\sum_j \frac{1}{X_j'^{\max}}\int_0^{X_j'^{\max}} \frac{1}{(\sigma_j^{\mu})^2}\left[\begin{pmatrix} X'^2 \\ X' \\ 1 \end{pmatrix}^{\mathrm{T}} \cdot \boldsymbol{\beta}_j^{\mu} - \frac{R_H X Y_j'}{t}\right]^2 dX \tag{10}$$

Where N is the number of corrected quadratic coefficient vectors we have and $X_j'^{max}$ is the maximum value of X' recorded in the original data that went into corrected quadratic j.

We can now determine the best fitting value of R_H by minimizing the function $G(R_H)$. This was performed numerically using the Nelder-Mead method

in log space via the functions $\mathcal{G}_+(R'_H/t') = \log(G(\log(R_H/t))$ and $\mathcal{G}_-(R'_H/t') = \log(G(-\log(R_H/t))$ since R_H/t spans many orders of magnitude depending on the material and the minimisation algorithm often missed minima when R_H/t was particularly small.

5 Results and Discussion

5.1 Raw Data

As described in the previous section, measurements were taken in a series of 'runs'. These are plotted for the different materials and are found in appendix A with their quadratic fits.

The semiconductor samples exhibit a strong Hall effect, with voltages recorded in the mV range. The curves for the two materials have opposite gradients for the the same regions of parameter space, indicating that as expected the dominant charge carriers have different polarity. Furthermore, for these materials the data follows a virtually perfect linear relationship with the current. However, in the case of the p-type sample a slight amount of curvature is visible for larger currents. Notably, this perturbation to the linear relationship changes sign when either of the the magnetic field or the current are reversed meaning it will not be removed by our elimination procedure and could therefore be due to the Ettingshausen effect which also has this property.

The metal samples on the other hand produced extremely small voltages of 10-100µV. In addition, the large currents passed through them (up to 10A) resulted in significant amounts of heating, particularly in the case of the tungsten sample. For the silver samples there appears to be a very good linear fit to the data, however unlike the semiconductor samples, the sign of the voltage measured does not reverse when the sign of the magnetic field changes. This would appear to indicate that a large portion of the signal measured here is not in fact due to the Hall effect and will be removed by the elimination procedure. This was also seen for the tungsten sample in addition to large amounts of curvature when the magnetic field was negative. Interestingly this amount of curvature was not observed for positive magnetic fields. Since this is not consistent with the systematic errors expected, it must be due to further systematic error such as inadequately controlled temperature during particular series of measurements. Indeed the temperatures recorded for negative fields were in general lower than those measured for positive fields (approx in the ranges 23°C-25°C vs. 24°C-27°C respectively).

5.2 Combined Data

After the mean of the matching quadratics was taken, the results were plotted along with the best-fitting Hall effect curve, calculated with the fitting procedure described above. These are shown in figures 1, 2 and 3. For all plots the combined quadratic is shown in blue, and confidence limits indicated by the dashed grey lines enclosing the shaded yellow area. The best fitting Hall effect curve is plotted in green.

For both semiconductors, it is clear that the combined quadratics make a good fit with the hall effect model. Both combined quadratics display a small

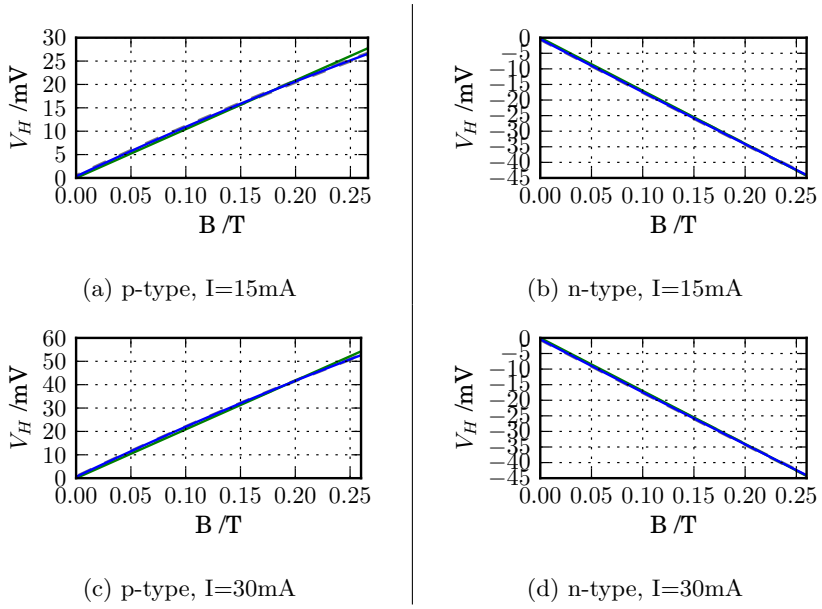
offset from zero and also a slight amount of curvature - with the curvature more promiment in the case of the p-type conductor as predicted.

The corrected quadratics of the metals display large amounts of curvature and also have relatively large zero offsets. As predicted the Hall voltage extracted from the silver data is very small, of comparable size with that produced from the tungsten data. Furthermore, with some of the systematic errors removed the silver sample appears to have a greater curvature than that shown by the corrected quadratics for tungsten. This would suggest that the silver samples are exposed to a stronger systematic error. A rough agreement is generally observed between the quadratics representing the data and the theoretical Hall voltage.

5.3 Goodness of Fit of Theoretical Hall Voltage

The value of the goodness of fit function $G(R_H/t)$ was plotted in figure 4. A red point is plotted at the position calculated by the minimisation routine of the best fitting value of R_H/t.

The value of $G(R_H/t)$ for the best fitting values of R_H/t is generally found to be on the order of 1, indicating a reasonably good fit, however despite the p-type and n-type data producing generally better fits the statistic is in fact worse for them than for the metals. This is largely a result of the very small relative errors which were listed for the semiconductor measured voltages, by comparison to the large relative errors for the metal samples. Since the errors



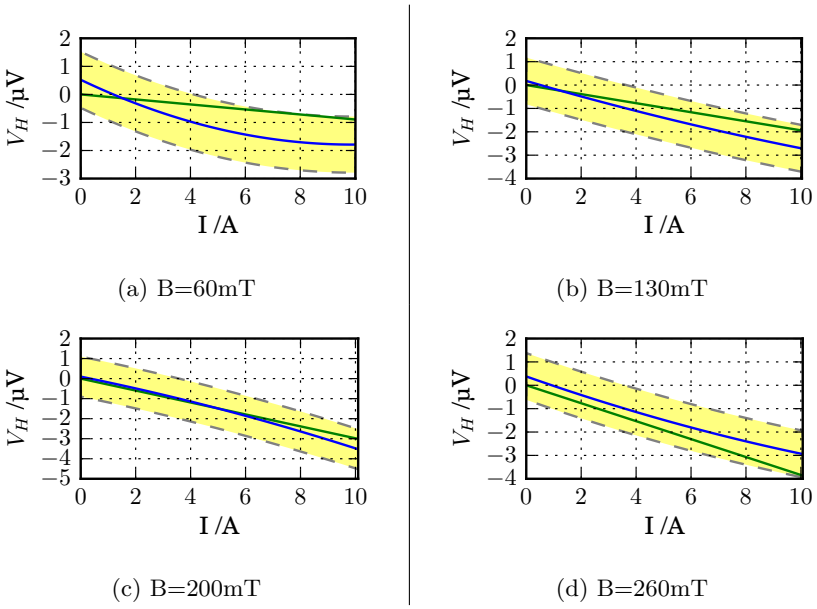
(a) p-type, I=15mA

(b) n-type, I=15mA

(c) p-type, I=30mA

(d) n-type, I=30mA

Figure 1: Corrected quadratics (blue) representing data for germanium samples with confidence limits denoted by shaded yellow regions (not visible due to small magnitude of errors). Best fitting theoretical Hall effect curve plotted in green, showing an excellent fit. p-type Ge curves both exhibit noticeable curvature - not as apparent in n-type plots.



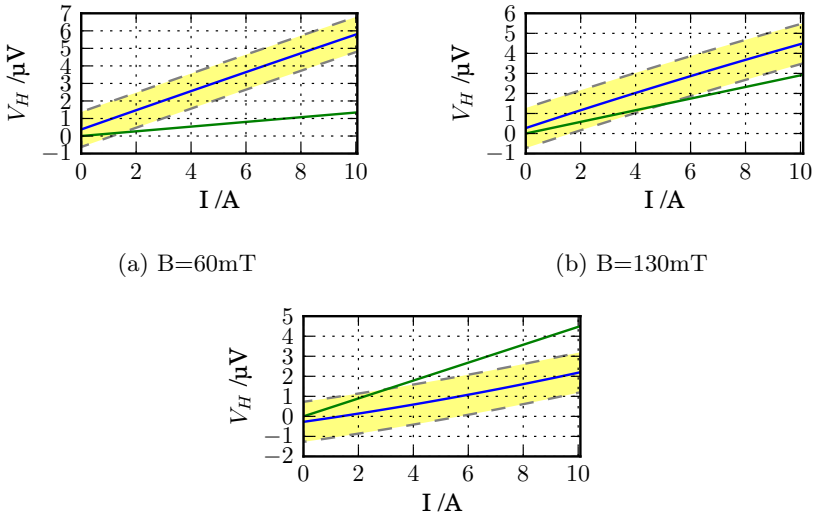
(a) B=60mT

(b) B=130mT

(c) B=200mT

(d) B=260mT

Figure 2: Corrected quadratics (blue) representing data for silver samples with confidence limits denoted by shaded yellow regions. Best fitting theoretical Hall effect curve plotted in green indicates a generally reasonable fit. All curves exhibit strong curvature.

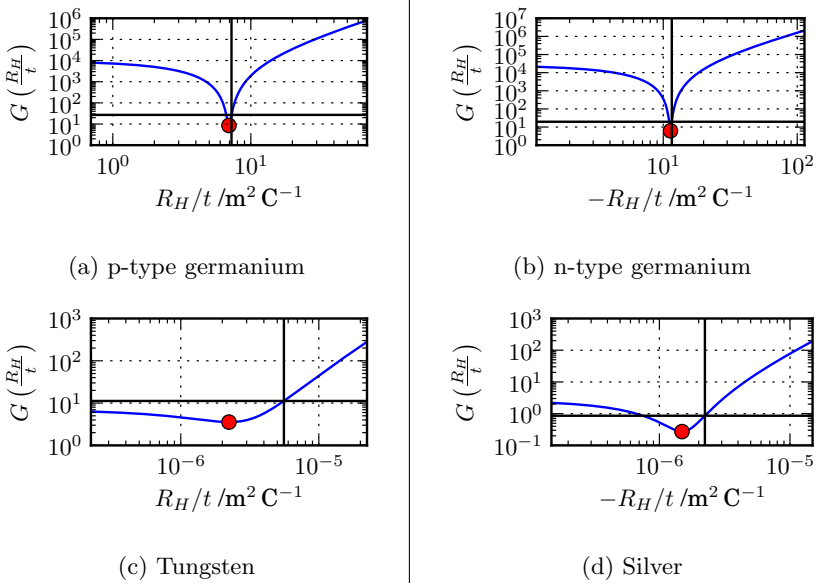


(a) B=60mT

(b) B=130mT

(c) B=200mT

Figure 3: Corrected quadratics (blue) representing data for tungsten samples with confidence limits denoted by shaded yellow regions. Best fitting theoretical Hall effect curve plotted in green indicates relatively poor fit. Curves are generally linear but indicate that $\frac{\partial V}{\partial B}$ is significantly different to theoretical fit.



(a) p-type germanium

(b) n-type germanium

(c) Tungsten

(d) Silver

Figure 4: Goodness of fit, G (see eq. 10) vs. R_H of quadratics representing data with theoretical Hall voltage for different materials (blue curves). Best fitting value marked by red dot, point at which curve offset from this by half an order of magnitude marked by black lines - this is used to define error in best fitting R_H.

assigned are not likely to be good estimators of the standard deviation (other than as an order of magnitude estimate) this result is probably not meaningful.

What is perhaps more meaningful is the shape of the graphs in the vicinity of the errors. This can be considered analogous to measuring how close points x on a straight line come to a point, where $x = 0$ is the position on the line where it makes its closest approach to the point and r_0 is their separation at $x = 0$. The separation at position x is thus $r = \sqrt{r_0^2 + x^2}$, with r_0 controlling how 'sharp' the curve $r(x)$ is. Similarly the sharpness of the minima on the error curves give a further indication of how close the theoretical hall voltage comes to the data. Thus the semiconductors appear to fit the data significantly more closely than the metals. We can estimate the error in the measurement by finding the value of R_H/t for which the integrated error curve is offset by a certain amount from the minimum. This offset was set at 0.5 in log space and its position is marked by the vertical and horizontal black lines on each plot. The values of R_H, n (from eqn. 2) and associated errors are thus:

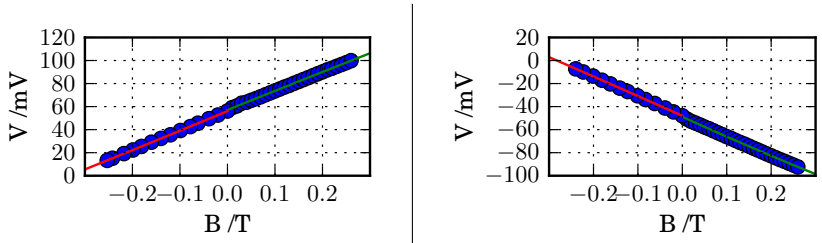
Material	R_H /m^3 C^{-1}	σ_{R_H} /m^3 C^{-1}	$n = \lvert eR_H\rvert^{-1}$ /m^{-3}	σ_n /m^{-3}
Silver	-7.4×10^{-11}	3.7×10^{-11}	8×10^{28}	4×10^{19}
Tungsten	1.1×10^{-10}	1.7×10^{-10}	6×10^{28}	8×10^{28}
p-type Ge	6.9×10^{-3}	0.3×10^{-3}	90×10^{19}	4×10^{19}
n-type Ge	-11.3×10^{-3}	0.3×10^{-3}	55×10^{19}	1×10^{19}

From these results it is clear that firstly the silver and the n-type germanium illustrate that their dominant charge carriers are negative (electrons), whilst the p-type sample and tungsten samples' dominant charge carriers are positive (holes). In addition, the metals have a significantly smaller Hall coefficient and thus much greater charge carrier density than the germanium samples - given that metals are conductors and germanium is a semiconductor this is no surprise. Finally, the charge carrier number density for the silver sample was found to be greater than that for the tunsten sample. This is also expected as silver is known to be a very good conductor whereas tungsten is a poor one and conductivity is directly proportional to the carrier density.

6 Conclusion

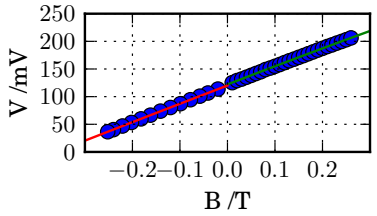
To conclude, we have measured the Hall coefficients of silver, tungsten, p-type and n-type germanium. From these it was shown that electric currents are carried by electons in n-doped germanium and silver, whilst in p-doped germanium and tungsten they are carried by holes. To accomplish this measurement, systematic errors from various effects were eliminated using a method of combining data symmetric in one independent variable and randomly chosen in the other. This was based on fitting of quadratic curves which were found to fit very well to the data obtained.

A Additional Figures

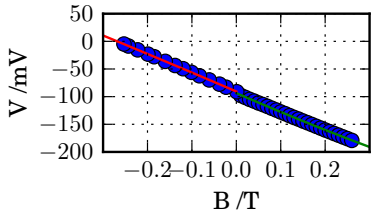


(a) I=-15mA

(b) I=15mA

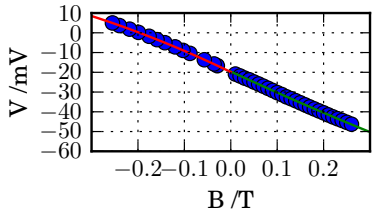


(c) I=-30mA

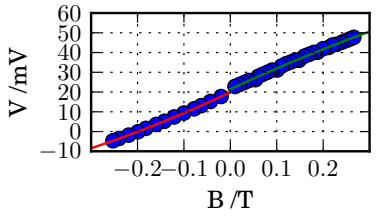


(d) I=30mA

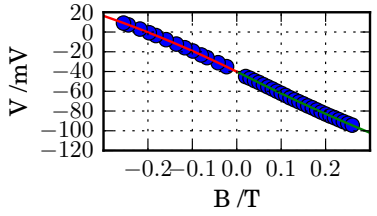
Figure 5: n-type Ge raw data (blue dots) with quadratic fits (red and green lines). Data almost perfectly linear, fits are thus extremely good.



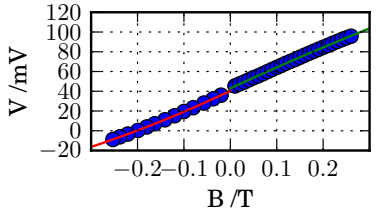
(a) I=-15mA



(b) I=15mA

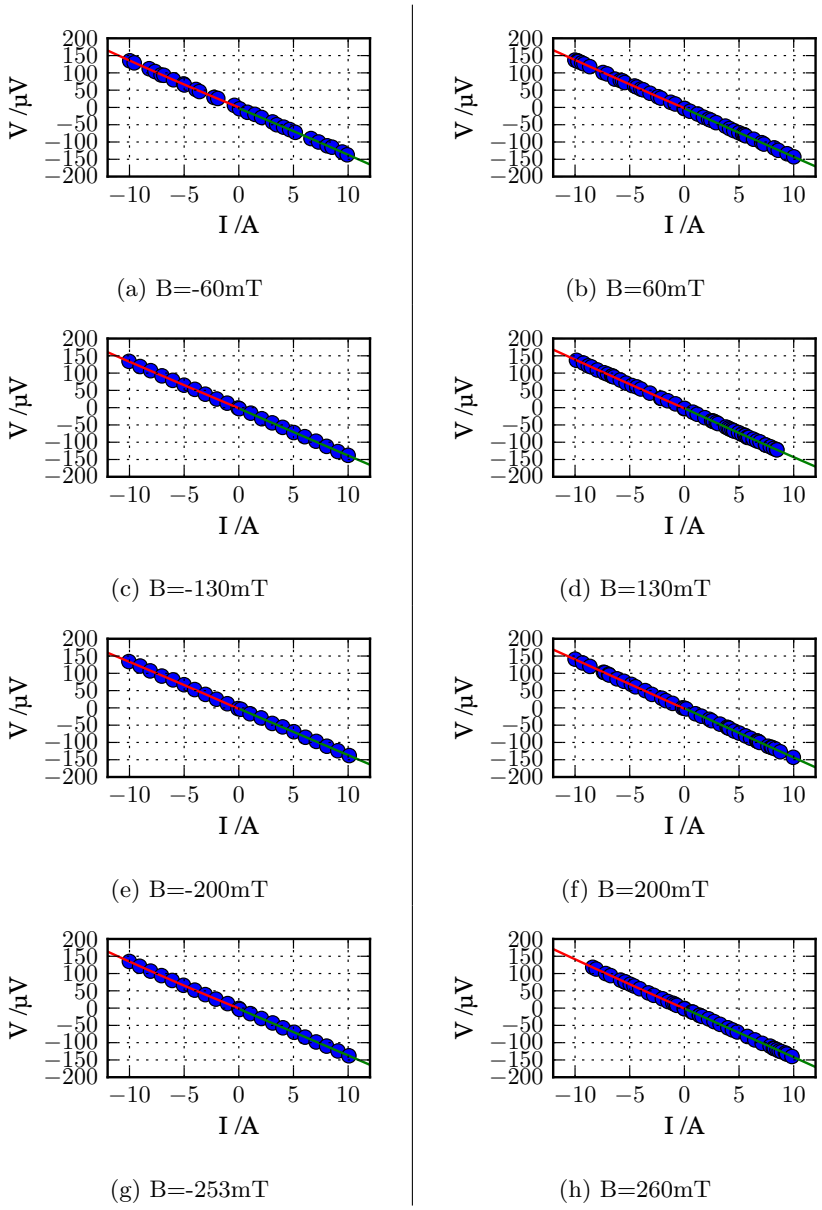


(c) I=-30mA



(d) I=30mA

Figure 6: p-type Ge raw data (blue dots) with quadratic fits (red and green lines). Data almost perfectly linear with slight curvature which is extremely well modelled by quadratic fits.



(a) B=-60mT

(b) B=60mT

(c) B=-130mT

(d) B=130mT

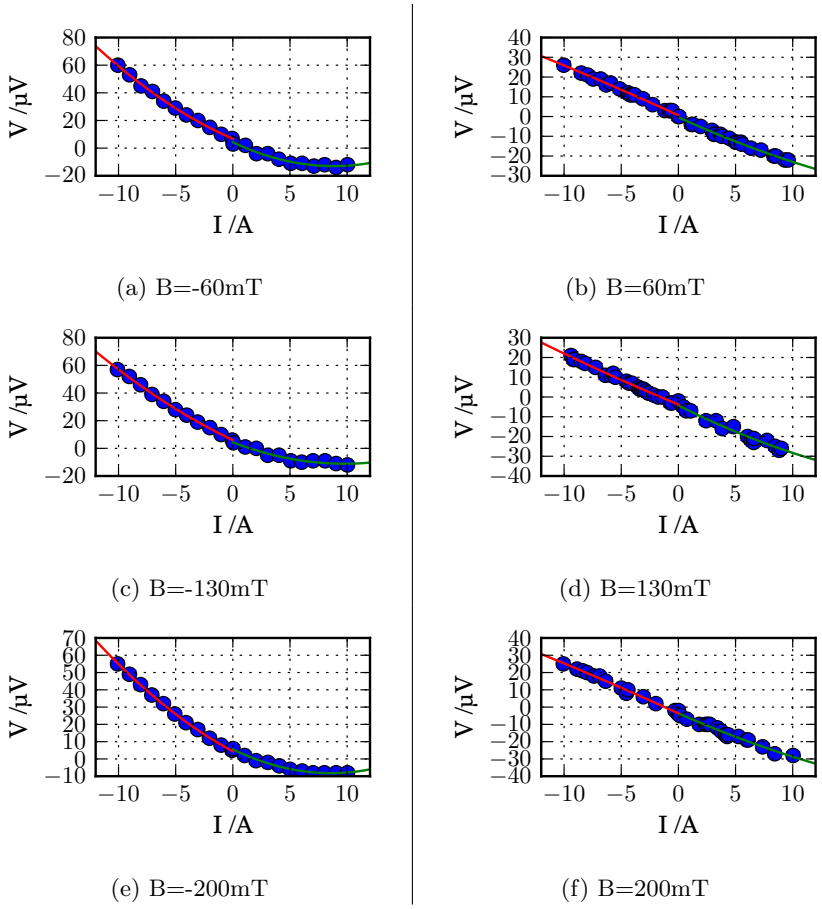
(e) B=-200mT

(f) B=200mT

(g) B=-253mT

(h) B=260mT

Figure 7: Silver raw data (blue dots) with quadratic fits (red and green lines). Data almost perfectly linear, almost perfectly modelled by the quadratics



(a) B=-60mT

(b) B=60mT

(c) B=-130mT

(d) B=130mT

(e) B=-200mT

(f) B=200mT

Figure 8: Tungsten raw data (blue dots) with quadratic fits (red and green lines). Data for negative B displays high degree of curvature which is very well modelled by the fits. In contrast data for positive B displays almost no curvature and is again modelled well by the fits.

B Tables

Material	Thickness /mm	Width /mm	Length /mm
Silver	0.05	20.0 ± 0.5	67.0 ± 0.5
Tungsten	0.05	20.0 ± 0.5	67.0 ± 0.5
p-type Ge	1.0	5.0 ± 0.5	10.0 ± 0.5
n-type Ge	1.0	5.0 ± 0.5	10.0 ± 0.5

Table 1: Dimensions of the materials examined

References

[1] E. H. Hall, *On a New Action of the Magnet on Electric Currents*, American Journal of Mathematics, Vol 2, p.287-292 1879

[2] O. Lindberg, *Hall Effect*, Proceedings of the IRE, 1952