

Silicon detectors for particle physics laboratory

R. Bates¹, L. Eklund², N. Pacifico² and M. Milovanovic³

This laboratory would not have been possible without the financial help from the particle physics groups of the Universities of Geneva and Glasgow and the sponsorship from the Scottish Universities Physics Alliance (SUPA) and the Science and Technology Facilities Council, UK.



¹ University of Glasgow, UK

² CERN, Geneva, Switzerland

³ University of Liverpool, UK

General Introduction

Semiconductor theory

A semiconductor is a solid whose electrical conductivity can be controlled by doping; adding impurities. It has an electrical resistivity between that of a conductor and insulator. The material is characterised by an energy band diagram with an energy gap between the conduction and valance bands of a few electron volts (more than 4 eV the material is said to be an insulator). The current in a semiconductor is carried by a flow of electrons and or by positively charged particles called 'holes' in the electron structure of the material. The electrons are free to conduct (know as free electrons) when they inhabit the conduction band. This takes place due to thermal excitation or doping. When an electron is excited from the valance band to the conduction band it leaves behind a hole in the valance band; which can be treated as a charge carrier with positive charge.

The conductivity of the material is controlled by the impurity concentration. Silicon is a group IV element with four valance electrons. Doping with a group V material (with 5 valence electrons) introduces an extra electron into the crystal structure which inhabits the conduction band. This is known as n-type doped material. Similarly doping with a group III element (three valence electrons) introduces a lack of an electron, or a hole, into the material. This is known as p-type material. A simple diode structure, called a p-n junction, is formed when a piece of p-type and a piece of n-type doped material are brought together. In the p-n junction electrons diffuse, due to the carrier concentration difference, from the ntype material to the p-material leaving a net positive fixed charge in the n-type region. The electrons recombine with the holes in the p-type material. Likewise free holes diffuse from the p-type material to the n-type, recombining with electrons, leaving a region of negative fixed charge in the p-type material. The fixed positive charge causes a force on the charge carriers and creates a drift current that is opposite in direction to the diffusion current. Eventually a steady state is reached. The region of fixed space charge is known as the depletion region as it is depleted (free of) mobile charge carriers. Due to the fixed charge a potential is present over the depletion region, and therefore an electric field exists. The application of a reverse bias over the p-n junction causes the depletion region to grow and eventually the material will become fully depleted, i.e. leading to a region free from mobile charge carriers.

For a silicon detector a p^+ -i-n⁺ structure is typically used. The intrinsic material (denoted by an i) is in fact either very slightly n or p type (denoted as n⁻ and p⁻). The positive superscript signifies that the doping concentration is high. The doping concentrations are typically 10¹⁸ atoms cm⁻³ for the n⁺ and p⁺ type doped sections and 10¹¹ atoms cm⁻³ for the near intrinsic region; while intrinsic silicon has a doping level of 1.5 x 10¹⁰ atoms cm⁻³ at room temperature. The detector's doped areas can be segmented into strips or pixels to enable position sensitivity. The heavily doped regions are only a few microns in thickness while the near intrinsic region is typically hundreds of microns in thickness. For operation as a detector the diode is reverse biased, where biasing refers to the application of an external voltage across the p-n junction of the detector, to set up a depletion region across the full thickness of the intrinsic material and therefore an electric field across the entire device. When a charged particle or photon enters the material the silicon is ionized and free electron hole pairs (charge carriers) are produced. The electric field causes the charge carriers to separate, before they can recombine, and to drift towards the heavily doped regions which form the external electrodes of the diode. The introduction of the charge carriers and their subsequent drifting induces a signal on the external electrodes which is measured by external electronics.

The p-n junction is illustrated in Figure 1. The first two sub-figures show the material with the fixed space charge illustrated as a sign inside a circle and the free charge carries as "-" and "+" signs. The other sub-figures illustrate physical attributes of the junction.



Figure 1 : The p-n junction. 1) p and n-type material 2) p and n junction in thermal equilibrium 3) absolute doping concentration 4) space charge density 5) free carrier density 6) electric field 7) electric potential The electric field and the electrostatic potential inside an abrupt p⁺-n junction can be calculated with the use of Poisson's equation, given in Equation 1, where Neff is the effective doping density of the semiconductor.

$$\frac{dE}{dx} = \frac{d^2 V}{dx^2} = -\frac{\rho}{\varepsilon} = \frac{q}{\varepsilon_{Si}\varepsilon_0} N_{efj}$$

Equation 1 : Poisson's equation

The boundary conditions are that the electric field and the potential are both equal to zero at the edge of the space charge region, that is:

$$\frac{dV(x = w)}{dx} = 0$$

and
$$V(x = w) = 0$$

Equation 2 : boundary conditions for the abrupt p⁺-n junction

Solving Poisson's equation gives the potential as a function of distance inside the junction as given by Equation 3. Inserting the device thickness gives the full depletion voltage of the diode. The depletion width can also be expressed as a function of the applied potential, as given in Equation 4.

$$V(x) = \frac{q}{2\varepsilon_{Si}\varepsilon_0} \left| N_{eff} \right| x^2$$

Equation 3 : Potential inside the p⁺-n abrupt junction as a function of distance

$$w(V) = \sqrt{\frac{2\varepsilon_{Si}\varepsilon_0}{q|N_{eff}|}}V$$

Equation 4 : The width of the depletion region as a function of the applied bias

Using the fact that the capacitance is given as the rate of change of the charge with potential, the capacitance of the junction can be expressed as a function of the doping density of depletion width as shown in Equation 5 to Equation 7.

$$C = \frac{dQ}{dV} = \frac{dQ}{dw} \cdot \frac{dw}{dV}$$

Equation 5 : The definition of capacitance as a function of depth inside the junction

$$C(V) = A \cdot \sqrt{\frac{\varepsilon_0 \varepsilon_{Si} q \left| N_{eff} \right|}{2V}}$$

Equation 6 : The junction capacitance as a function of applied voltage

$$C(w) = \frac{\varepsilon_{Si}\varepsilon_0 A}{w}$$

Equation 7: The capacitance as a function of depletion width

Device Types

Pad detector

The pad detector is a very basic silicon detector. It consists of a p^+-i-n^+ junction with one large heavily doped pad on either face of the intrinsic bulk material. One face is p^+ doped and the other n^+ doped. This means there is no way of deducing a particles interaction point except to say that is occurred within the detector. The signal induced will depend on the type of ionizing particle interacting in the detector and the depletion width (controlled by the applied bias voltage).

Strip and Pixel Detectors

The detector's highly doped regions (electrodes) can be segmented into strips or pixels. When an ionising particle enters the detector ionising the material free electron-hole pairs are created. These drift under the influence of the electric field in the depletion region and a signal is induced on the electrodes in the vicinity of the original ionisation. This enables the original position of the ionising particle to be reconstructed from the signal collected on the segmented electrodes. Both types of detectors work in this way, the difference between them being that the pixel detector has individual contacts that are arranged in a 2D array while the strip device has only a 1D array of contacts. Therefore the pixel detector allows the 2D determination of the original ionisation point rather than just the 1D. This laboratory will only use strip devices.

Energy deposition from a high energy particle

When a silicon detector is traversed by a high energy particle and only a small amount of energy is lost by that particle per unit thickness of the detector, such a particle is known as a Minimum Ionising Particle (a MIP). When the ionising particle passes a silicon atom in the lattice it will tend to liberate a loosely bound valance electron. The electron will be ionised at an excited energy state and via collision will enter thermal equilibrium with the lattice as a free electron in the conduction band. Occasionally the ionising particle will interact with a more tightly bound electron than the silicon valence electrons and as a result create an electron with a significant amount of kinetic energy. This liberated electron can cause secondary ionisation of the silicon. The energy deposited for this interaction is therefore higher than that for the interaction with the valence electrons. As a result of these rare high energy events the energy spectrum of the deposited energy in the silicon detector is a non-symmetric distribution as shown in Figure 2. The distribution was first described by Landau and therefore takes his name.



Figure 2 : The Landau distribution.

Due to the non-symmetric nature of the distribution the mean value is higher than the most probable value. For a MIP the mean energy loss per unit length is 3.88 MeV cm⁻¹ or 116 keV for 300 μ m of silicon; while the most probable energy loss is approximately 0.7 times this or 81 keV for a 300 μ m thick detector. In silicon it takes 3.6 eV to create an electron hole pair under ionisation by a MIP. Therefore the mean

number of charge carriers created is 108 per micrometer and the most probable is 72 per micrometer. This results in a mean signal from a 300 μ m thick detector of 32000 electrons or a most probable value of 22500 electrons; which is equal to 3.6 fC. It is the most probably energy that is used for calculations of collected charge from a silicon detector.

Signal Generation and Ramo's Theorem

The signal is induced on the external electrodes by moving charges and is observed as pulses on an oscilloscope. Ramo's theorem provides a way of calculating signals induced on the electrodes of a detector by the movement of charge carriers. The drift field determines the trajectory and velocity of the charge carriers. The induced signal on the electrode was first formulated by Ramo (Published in Proc.IRE.27:584-585,1939) for a vacuum tube system, and is valid for a semiconductor detector, using the Gauss identity. The theorem states that the induced current is given by the dot product of the velocity vector \underline{v}_{α} and the weighting field Ew, as given in Equation 8. The weighting field is calculated for the condition that all the electrodes in the system are held at zero potential except the collecting electrode which is held at 1V.

 $i = q \vec{E}_{\scriptscriptstyle W} \cdot \vec{v}_q$ Equation 8 : Ramo's theorem for the induced current

The induced charge is simply given as the integral of the induced current for the movement of the charge carrier from point r_0 to point r_1 , as given in Equation 9.

$$Q_a = \int i_a dt = -q \int_{r_0}^{r_1} \vec{E}_w \cdot dr$$

Equation 9 : Ramo's theorem for the collected charge from the movement of a charge q from point r₀ to r₁.

The velocity of the charge carriers is given by Equation 10, where μ is the carrier mobility and E the electric field.

$v = \mu E$ Equation 10 : Carrier drift velocity

The mobility is carrier type dependent and given below for silicon (subscript n for electrons, p for holes).

$$\mu_n = 1350 \text{ cm}^2 \text{V}^{-1} \text{s}^{-1}$$

 $\mu_p = 480 \text{ cm}^2 \text{V}^{-1} \text{s}^{-1}$

For a simple pad detector the electric field is given as V/d and the weighting field is given as 1/d where d is the separation of the electrodes (equal to the thickness of the device). As a consequence the signal is generated equally for a given increment of travel for a charge carrier throughout the thickness of the detector.

In a strip or pixel detector the drift electric field is close to that of a pad device. However, the weighting field is more complex as the 1V potential is applied only to the sense strip or pixel. This gives a very non-uniform weighting field, as shown in Figure 3. As a consequence the signal on the collecting electrode is dominated by the movement of the charge carriers close to the collection electrode rather than equally throughout the bulk of the detector.



Figure 3 : Weighting potential for a strip (pixel) detector.

Strip Detector

<u>Aim</u>

This section of the laboratory has several aims which are:

- To introduce the silicon detector module: consisting of the silicon detector, readout chip and thermal and mechanical aspects.
- To observe the noise of a silicon strip detector as a function of bias voltage.
- To observe the signal spectra due to a minimum ionising particle in a silicon detector and demonstrate the Landau distribution shape of collected charge.
- To observe the physical size of a charge cluster from a minimum ionising particle and relate this to the position resolution of the detector.

<u>Apparatus</u>

The apparatus for this experiment consists of a silicon strip detector, a readout amplifier chip, the readout system to control the readout chip, a Sr-90 beta electron source and a Scintillator coupled to PMT for the trigger. The whole system is driven over the USB by a laptop running Linux. There is a dedicated data acquisition code with a graphical user interface, gui, and a root based data analysis code. Each item is described below and shown in pictures at the end of the lab script.

The apparatus: The source and trigger

The experiment will attempt to characterise the silicon detector under the illumination of minimum ionising particles, as found in a particle physics experiment. The experiment uses a Strontium-90 source as the source of ionisation.

Strontium-90 source

Strontium-90 beta decays to Yttrium-90 which beta decays to Zirconium-90 which is stable. End point energy of Sr-90 decay is 0.546MeV

End point energy of V 00 decay is 0.340We

End point energy of Y-90 decay is 2.28MeV

Sr-90 can be considered to be a pure electron emitter.



Figure 4 : Y-90 beta spectrum

The absorption of the beta electron is described by via the Bethe-Bloch process:







The electrons pass through the silicon and deposit all of their remaining energy in the scintillator. The size of the signal from the scintillator/PMT is proportional to the electron energy.

The signal from the PMT will trigger the readout of the silicon detector if the signal is over a user set threshold value.

Look on the scope at the trigger signal from the PMT.

Adjust the threshold of the scope to decide on a good value for the threshold

- What happens if the threshold is too low?
- What happens if the threshold is too high?
- If the trigger threshold is set to trigger on low energy electrons, what problem in the collected charge would happen (look at the Sr-90 beta spectra and the Bethe-Bloch graph)?

The apparatus: The strip detector module

We have a 1cm long ATLAS prototype silicon strip detector with 100 strips. We have two readout chips in our module with 128 channels each. Each channel is an independent preamplifier with shaping to get full charge collection within 25ns.

Below is a picture of the module that will be used in the experiment.



- Identify the different parts of the module (silicon strips, amplifier chip, etc)
- Why are the chips and silicon detectors not next to each other?
- What is between the chip and detector?
- What does this represent electrically to the amplifier?

The strip detector is a p-n diode. The strips can be either n^+ or p^+ , see the diagram



Figure 5 : A diagram of the silicon strip detector

Typical values of doping concentrations are: n & p ~ 10^{-12} cm⁻³ n⁺ & p⁺ ~ 10^{-15} to 10^{-18} cm⁻³

- For the two diagrams what is the sign of the bias voltage applied to the back side of the detector required to reverse bias the detector?
- What happens to the silicon material after heavy irradiation?
- What effect does this have on the position of the p-n junction and how it grows with bias voltage?

Noise

The total noise of the system measured as an equivalent noise charge at the input of the amplifier, ENC, is given by:

$$ENC_{tot}^{2} = ENC_{pa}^{2} + ENC_{i}^{2} + ENC_{RP}^{2} + ENC_{RS}^{2}$$

The individual noise sources are:

$$ENC_{pa} = A + B \times C_{load}$$

Equation 11 : ENC for the preamplifier, where A and B are pre-amplifier constants and C_{load} is the capacitive load at the input of the pre-amplifier

$$ENC_i = \frac{e}{q} \sqrt{\frac{qI\tau}{4}}$$

Equation 12 : ENC due to the detector leakage current for CR-RC shaping. τ is the shaping time of the shaper.

$$ENC_{RP} = \frac{e}{q} \sqrt{\frac{\tau kT}{2R_b}}$$

Equation 13 : ENC for the thermal noise from the parallel resistors. R_b is the bias resistor of the detector.

$$ENC_{RS} = 0.395 \times C_{load} \sqrt{\frac{R_s}{\tau}}$$

Equation 14 : ENC for the series resistance in the circuit. Rs is the resistance of the readout strip

Run the DAQ software and run a pedestal run.

Look at the pedestal and noise histograms.

- What is being plotted in each?
- What is the pedestal?
- Is the value constant
 - o across the two chips?
 - o with time?
- Can you identify the two chips from the pedestal plot?
- What is the noise in the noise histogram?
- Can you identify where the detector is?
- Comment on the high and low points on the noise histogram
- Estimate what the average noise of the detector is?

Increase the bias voltage slowly from 0 to 100V

- What happens to the noise in the detector?
- What is changing with increasing bias and how does this change the noise, look at equations above?
- Think about what could happen if the detector current was higher and increasing with bias voltage.

Signal

The signal is induced on the external electrodes by moving charges (electrons and holes)

To find the signal in a given event, the signal collected on each given amplifier channel is compared in turn to a cut, know as the seed cut. If the value is higher than the cut the channel is considered to have signal.

If the signal is higher than the seed cut, the output the neighbouring channel is looked at. A second cut, the neighbour inclusion cut, is applied to this channel to include the signal on the neighbour into the event (cluster). If the neighbouring channel is included the neighbour inclusion cut is applied to its neighbouring channel. This process is repeated until the signal on the neighbouring strip is below the inclusion cut.

The signal size is the sum of the signals on all the strips in the cluster.

The cut applied must reduce the likelihood of including noise, but should keep the signal. Look at the following table of likelihood of an event being under a Gaussian function as a function of the cut applied.

Table 1. Probabbling of being inside a Gaussian distribution	
Number of standard deviations	Probability of being under the curve
1	0.682689492137
2	0.954499736104
3	0.997300203937
4	0.999936657516
5	0.999999426697
6	0.99999998027

Table 1 : Probabbility of being inside a Gaussian distribution

- What do you think is a good value of the seed cut in units of standard deviations of the noise Gaussian?
- Explain why you have chosen this value
- What do you think is a reasonable value for the neighbour inclusion cut?

Data collection

Bias the detector to 100 V

Align the Sr-90 source, Detector module and Scintillator

Run the DAQ software

(In the gui there should be three lines of text. If they are in red it means there has been a problem somewhere and you should exit the program, switch the motherboard off at the wall, unplug the USB cables and start again. If they are black then everything has been set up correctly and is ready to run).

Check that the trigger is set as you require (OR and Trigger in) and the threshold value as you determined from the oscilloscope.

Log the Data to a file for the data from the detector (you need to decide upon a name) Click on RS (radioactive source) Click on Run

You should have a data rate of 15-20 Hz.

For the on-line analysis a seed cut of 5 noise sigmas is used.

Look at the different pull down menus

- What is being plotted in each window?

In the signal page look at the two chips individually

- What can you say about the signal in the two chips?
- Is the S/N cut good enough?
- What addition feature adds to the noise that you can remove in an off-line analysis?
- In the signal page look at the time projection
 - What is this showing you?
 - Would this detector be any good for the LHC? (bunch crossing 25ns)

You should collect data for about 10-15 minutes (10000 events will suffice). If time is short then don't worry there is an example data file for the off-line analysis.

Data Analysis

Run the analysis code. Several windows (know as canvases) will appear. Look at the analysis canvas

What is plotted in the top left of the canvas?

- Why are chip 1 and chip 2 different?
- What does this say about your hit occupancy?
- If you were designing a detector system would you read out all the channels, if not why not?

Look at the bottom left plot, data is only included in this plot if it passes the seed inclusion and neighbour inclusion cuts.

- What is plotted?
- Comment on the shape of the plot

A timing cut is applied to the data to only include the data that is in the peak section (+/-10ns around the peak). This data is called the peak data. Data is also plotted for a timing cut of 80-100ns and called the baseline.

- What do you expect to be in the base line data plots?

Loot at the top right hand plot

- Do you see any noise events?
- How does the noise affect the signal?
- What can you say about the shape of the signal plot?
- Can you explain the physics of the shape?
- Is the reported peak at the same place as the apparent peak of the fit?
- If not do you think the fit is correct?

Look at the Henrycanvas (not Henrycanvas2)

- What is plotted?
- Comment on the shape?
- Why is the cluster width 2 highest? Do you think that this is reasonable?
- Why are there clusters more than 2 strips wide?

Look at Peak Distribution canvas

- What are you looking at?
- What can you say about the difference in the plots on the top left and bottom left?
- Do you think that this has a physical reason?

Thinking about the plots in the Henrycanvas and Peak distribution canvas

- Do you think the position resolution will be affected by larger clusters?
- If so will the affect be always good, always bad, or sometimes good and sometimes bad? Explain why.
- Read the slides "delta electrons and resolution in a silicon detector"

Conclusions

Look at both sets of results from each part of the experiment. Consider the different attributes from each detector and the different ways in which they could be used.



Pictures of the Strip laboratory set-up

Figure 6 : A sr-90 source aligned above the silicon detector with an aligned Scintillator trigger below.



Figure 7 : Motherboard with flat cable to daughter board, trigger input from PMT and USB cable connected.



Figure 8 : Example of the trigger signal on the oscilloscope.

<u>Appendix</u>

Strip detector

The strip detector is a more complex detector than the pad device. The strip layout means the signal can be picked up on several electrodes and depending on the relative strength of the signals the interaction position of the particle can to be determined. The silicon sensor in this experiment is 1 cm x 1 cm and 300 μ m thick. There are 100 strips at a pitch of 80 μ m. The strip detector has a silicon oxide layer between the implant and the aluminium readout electrode. This oxide layer acts as a capacitor between the diode and the amplifier. Why do you think this might be? To enable a D.C. connection across the diode a set of bias resistors are used to connect each strip implant to a common implant (know as the bias rail or bias ring). This common connection is connected to the H.V. return 9 at the ground potential of the daughter board). The H.V. bias is supplied via a back side contact to the strip detector. The detector is mounted on the daughter board and connects to a Beetle amplifier chips. The two experiments use different types of strip detector. One has a p⁺-n-n⁺ diode and the other an n⁺-p-p⁺ diode, where the first implant is the segmented strip structure.

Beetle Chip

The Beetle chip is an analogue readout chip. The daughter board contains two Beetle readout chips. They each have 128 independent input channels of analogue amplifier and shaper with a 25ns peaking time. The analogue signal from each channel is readout from each chip as a multiplexed analogue signal. The input dynamic range of the amplifier is around ±110000 electrons, that is to say about 5 MIPs of either positive or negative polarity.

Trigger, Pipeline

Triggered systems are used to identify what events should be stored for later analysis. As only a limited amount of events can be stored the trigger is used to rapidly decide which ones are interesting enough to keep. In the case of this lab the trigger will select events where a particle passed through the sensor and deposits enough energy in the scintillator for the output of the PMT to exceed a predefined value. Therefore in this set-up a scintillator trigger is used.

The pipeline stores the events on the Beetle chip waiting for the trigger to decide whether or not they are 'interesting'.

Readout System

The readout system consists of the daughter board which supports the Beetle chips and the mother board. The daughter board has an additional amplifier to amplify the signal for transmission to the mother board. The mother board digitizes the data and transfers it via the USB to the PC. The system contains two front-end readout chips. There are two analogue outputs on the motherboard in order to probe the analogue output signal of each Beetle chip before they are digitised. The mother board accepts an analogue signal from a PMT to use as a trigger for data acquisition.

The Software

There are two software codes; the data acquisition code to collect the data and perform online data analysis and the data analysis code to perform detailed off-line data analysis. To start the DAQ system.

- Start by opening a terminal window then type 'ls'. This brings up the list of directories. Make sure you see the directory Analysis_Compiled then type 'cd Analysis_Compiled' to take you into that directory.
- 2. Plug the USB cable into the laptop and the motherboard, and plug in the motherboard's power supply. A green and red LED on the motherboard should illuminate to show you it's switched on.
- 3. In the terminal type 'dmesg | grep tty'. The *dmesg* code checks amongst other things the connections to the USB hub, and *grep* looks for the structure *tty* in the codes output. The pipe (i.e. "|") can be found by holding the "Alt Gr" button and the button above tab. The last line of the code that appears should read "FTDI USB Serial device convertor now attached to ttyUSB0". The 0 could be replaced by any other number and this is the number you use in the next line of code. We'll just assume it's 0 but make sure to change it if it's any different.
- 4. Type 'change_priv 0' and press the blue reset button on the motherboard.
- 5. Type 'alibava-gui dev=/dev/ttyUSB0 lars.ini'. This will open the Alibava program gui window. The input file lars.ini contains the set-up details for the Beetle chips.
- On-line data analysis:

The graphs produced by this program are:

- **Signal** used for the calibration and radioactive source runs. Only of interest for the radioactive source run.
- Pedestals and noise used for the pedestal and radioactive source runs. The chip output is about 500 with no input signal. This varies from channel to channel and event to event. The zero value is known as the pedestal. The noise on each channel is calculated by calculating the standard deviation of the signal (pedestal value) on that channel. The value is constantly updated. You will see on the graph the distinction between the two chips on the daughterboard as only one is connected to the detector so we will get a larger noise on one of them. Which one is it? A large spike indicates a noisy channel and a low spike indicates an unbonded channel. Both need to be masked out later. Also need to mask off 28 channels on the chip connected to the detector as there is no signal on them as the detector only has 100 strips. The noise will change with boas voltage which way will it change and why?
- **Hitmap** shows you the signal on the channel that is hit.
- **Temperature** of the daughterboard
- **Time structure** shows the time a trigger occurs, this should be uniform
- **Event display** this doesn't give the average, it displays the data on each channel for just 1 event updated every 100 events.
- Noise/Common mode the noise graph gives the average noise for each chip as a function of event number, the common noise is how the chip channels change in a common fashion (for example they all jump up together). This is displayed as a function of the event number, not channel number.