Seminar pri predmetu Eksperimentalne metode in detektorji II

Semiconductor detectors for detection of elementary particles

Gašper Kukec Mezek

Fizika II. stopnje
Študijsko leto 2012/2013

Contents

1 Sources of ionizing radiation

Particle detectors and their geometry depend on the type of ionizing radiation we want to measure. In principle, ionizing radiation can be divided into three categories, depending on its production type, energy[1] and experimental use.

- **Radioactive and nuclear sources**
  Ionizing radiation resulting from radioactive decay of particles. These include α, β and γ decays and neutrons from fission reactions. Usually not widely used in measurements and direct detection, other than dosimetry and detector calibration. Their energies are the lowest of the three, with γ–rays typically being under 10 MeV and fast neutrons with energies up to 20 MeV.
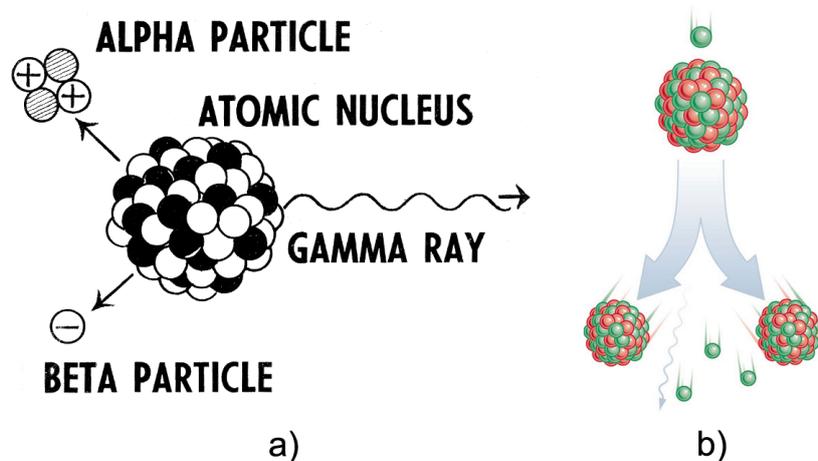
Figure 1: Representation of a) radioactive decay (α, β and γ) [1] and b) nuclear fission [2]. In both figures protons and nucleons are coloured differently.

- **Accelerated particles**
  This group includes ionizing particles from any type of man–made acceleration source, from cyclotrons and synchrotrons to particle colliders and linear accelerators. Particles accelerated in laboratories are usually measured indirectly to study the structure of an object (spectrometry) or particles arising from collisions (particle physics). In that sense, they are tightly connected to radioactive and nuclear sources, since the particles we eventually detect come from nuclear reactions or decay products during an interaction. We only track the primary particle (or opposing particles for a collider) when we want to figure out what happened to it after interacting (transmission, exit energy, scattering direction,...). The energies of accelerated particles are generally on the order of GeV, with highest energies up to TeV range (Tevatron, LHC).

---

[1]In particle and nuclear physics, energy is typically given in units of electron–volt eV (TeV = 10^3 GeV = 10^6 MeV = 10^9 keV = 10^{12} eV)
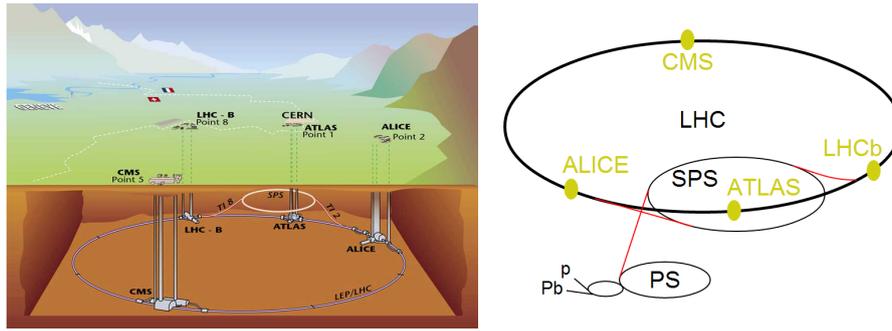
Figure 2: Two schematic views of the Large hadron collider (LHC), currently the largest and most energetic particle accelerator. p and Pb are linear accelerator injectors (for protons and lead), PS and SPS are storage rings (synchrotrons), while CMS, ATLAS, ALICE and LHC–B are experimental stations [3, 4].

- **Cosmic rays**

  The third group of sources include any kind of particles coming to Earth from space. All such particles are called cosmic rays. Their identity can vary from protons, neutrons, electrons, photons and neutrinos to heavier elements like iron. They cover a very wide range of energies, from lowest coming from the Sun at energies on the order of GeV to high energy particles from unknown sources with energies exceeding 10^{20} eV $= 10^{11}$ GeV. Unfortunately, the usefulness of cosmic rays is limited by the low flux at extreme energies and the fact that they are not directed like particles from accelerators, but arrive from a random direction. Therefore, measurements in cosmic ray physics are mostly used for identification and information on extra–terrestrial sources that produce them.

Figure 3: A side view of a cosmic ray induced extensive air shower. The tank grid on the ground are Cherenkov detectors, the telescope detector in the bottom right measures the fluorescence of nitrogen upon interaction [5].

3

2   Detection of ionizing radiation in semiconducting materials

In the recent decades semiconductor–based detectors have become well established for measurements of ionizing radiation and are, in many applications, quickly replacing other less efficient materials. Their special energy level structure gives them unique conducting pr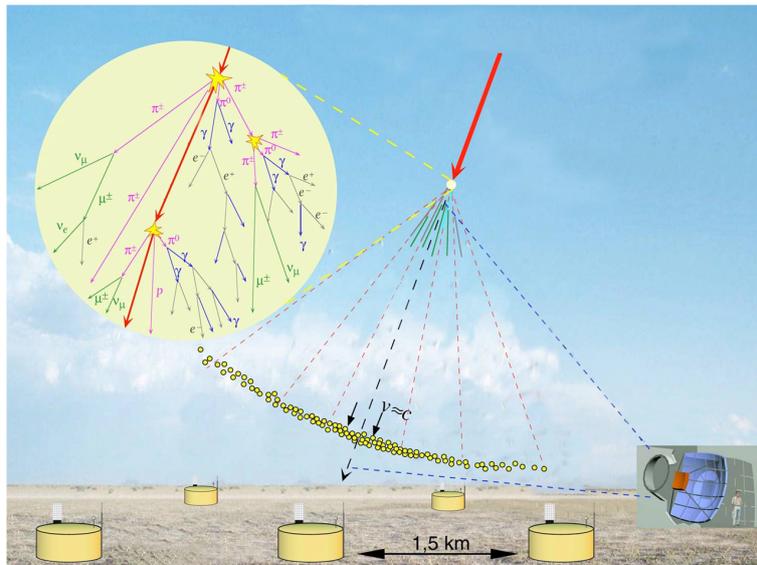operties we don't find in insulators or metals, while smaller size needed to detect ionizing radiation means they take up less space, where we can't afford it, like in colliders.

2.1   Energy band structure

When distinguishing materials by their ability to conduct electricity, the easiest explanation is the one using their energy level structure. If we gather energy levels with free electrons in a conduction band and those with bound electrons in a valence band, we get an energy band structure as illustrated in figure 4. The only current
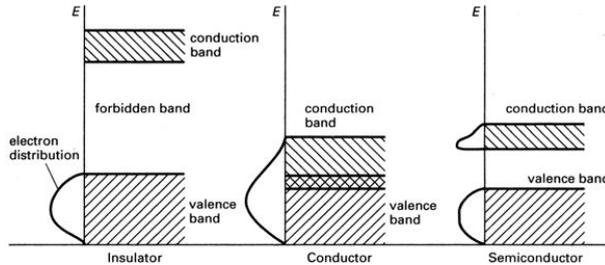
Figure 4: Energy band structure of insulators, conductors and semiconductors. The forbidden band or band–gap has no available states for electrons, the valence band holds only bound electrons, while the free electrons in the conduction band contribute to conduction of electric current [6].

carrying electrons are the ones in the conduction band, thus the band name. Whenever electrons from valence band are given enough energy to cross the band–gap, either from thermal excitations or ionization, they will settle in the conduction band and contribute to the conduction of electric current. Metals (conductors) have no band–gap, with their valence and conduction bands overlapping, making them extremely suitable for conducting electric current. On the other side, insulators have a large enough energy gap that most or all electrons can't pass from the valence to the conduction band. The special property of semiconductors arise from their smaller band–gap, giving them the ability to have a portion of bound electrons crossing to the conduction band, leaving behind positively charged holes. The semiconductor as a whole remains neutral. Typical band–gap energies are $E_{\mathrm{ins}} \sim 6\,\mathrm{eV}$ for insulators and $E_{\mathrm{semi}} \sim 1\,\mathrm{eV}$ for semiconductors. Naturally, the exact values depend on the material type and temperature.

2.2   Charge carrier mobility and doping

In semiconductors we have two types of charge carriers, the more mobile electrons and less mobile holes. The movement of holes is visible because neighbouring elec-

trons cross up to the conduction band, while the original hole is filled by another electron. In the presence of an applied electric field to the semiconductor, the charge carriers start to move through a semiconductor with drift velocities

$$v_e = \mu_e E \qquad \text{and} \qquad v_h = \mu_h E, \qquad (2.1)$$

where indices e and h correspond to electrons and holes, respectively. The mobility μ gives a direct correlation between speed of the charge carrier and electric field E and is also dependent on temperature. The higher the temperature, less mobile the charge carriers are. Typical values for these are $\mu_e = 1350 \, \mathrm{cm^2(Vs)^{-1}}$ and $\mu_h = 480 \, \mathrm{cm^2(Vs)^{-1}}$ for silicon and $\mu_e = 3900 \, \mathrm{cm^2(Vs)^{-1}}$ and $\mu_h = 1900 \, \mathrm{cm^2(Vs)^{-1}}$ for germanium. The mobility in turn gives us the resistivity of a semiconductor, with values around $\rho_{\mathrm{Si}} = 230 \, \mathrm{k\Omega cm}$ for silicon and $\rho_{\mathrm{Ge}} = 45 \, \mathrm{\Omega cm}$ for germanium. In general, semiconductors are divided into two groups: intrinsic semiconductors with equal free electrons and holes, and extrinsic or doped semiconductors with an excess of one of the two charge carriers. Figure 5 represents the two doping types, n–type (a) and p–type (b). The semiconductor becomes of n–type, when a normal
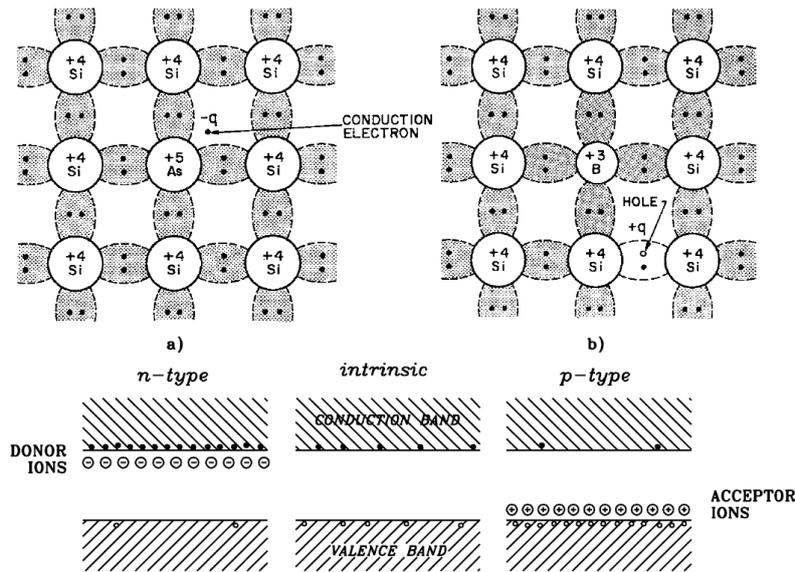
Figure 5: Representation of the crystal and energy band structure for a) n–type doping and b) p–type doping. The energy band structure of an intrinsic semiconductor is added for comparison [7].

+4 valence atom is replaced with a higher valence atom (donor), bringing an extra unbound electron. The usual doping elements are +5 valence elements such as arsenic (As), phosphorus (P) and antimony (Sb). The n–type dopant atoms are in essence impurities and thus create energy levels inside the energy gap close to the conduction band, making it easier for electrons to excite to the conduction band. Similarly, semiconductors are of p–type, when we replace the normal atom with a lower valence atom (acceptor). The dopant gives extra holes to the semiconductor and is usually a +3 valence element like gallium (Ga), boron (B) and indium (In). Just as for the n–type doping, we get an energy level inside the energy gap, but this time it's close to the valence band.

In spite of all the excess charge carriers in doped semiconductors, they are still neutral, with equal positive and negative charge densities

$$N_D + p = N_A + n. \tag{2.2}$$

N_D and N_A are donor and acceptor concentrations, p and n are the concentrations of holes and electrons in a semiconductor. For example, an n–type semiconductor ($N_A = 0$) gets the majority of electrons from the dopant ($n \simeq N_D$). A doped semiconductor can in principle still have both dopant types, but its type is still determined by the difference between the two doping concentrations. If $N_A > N_D$, the semiconductor is a p–type, otherwise $N_D > N_A$ gives us an n–type. In another possibility, the two concentrations can be equal $N_A = N_D$ and produce an i–type semiconductor, designated as a compensated semiconductor.

As expected, the properties of a doped semiconductor are different from those of the intrinsic one. The dependence of mobility and resistivity on doping level is shown in figure 6, with both of them decreasing with doping concentration. The way we
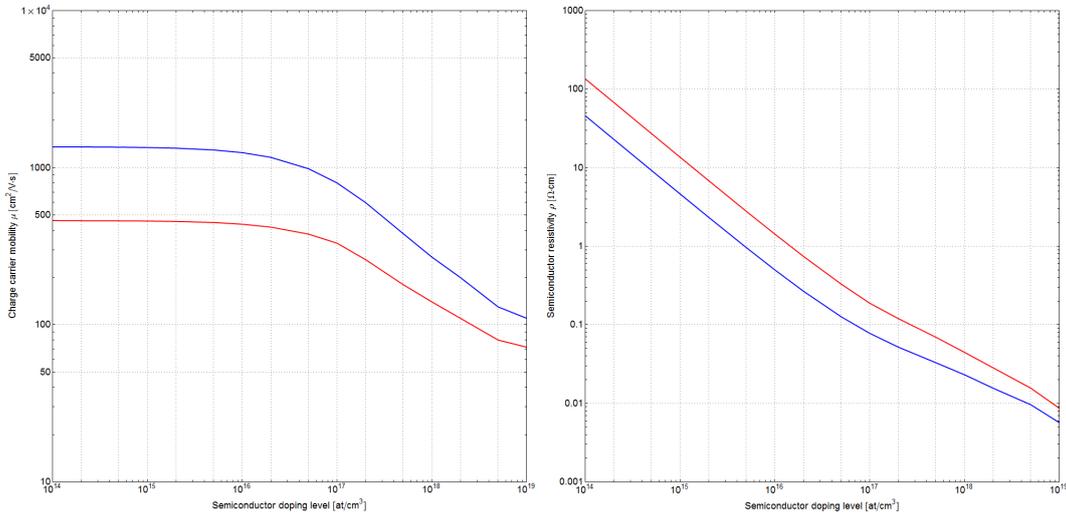
Figure 6: Dependence of mobility (left) and resistivity (right) on doping concentration. On the left graph, the blue line is electron mobility, the red is hole mobility for silicon. On the right graph, the red line is resistivity for n–type silicon, the blue line for p–type silicon.

can understand this decrease is that the more charge carriers we have, the more they are repelled from each other due to their like charges. For faster responses in detection, the dopant concentration should in fact be as small as possible.

2.3   p–n junction and reverse biasing

The most basic structure we can construct from two different types of semiconductors is a p–n junction, an interface between an n–type and a p–type. The idea seems simple, but we can't just put two semiconductor types together, one has to be diffused into the other. This actually creates a fairly well defined junction between the two, controlled by the amount of the used dopant. Since the p–side has excess holes and the n–side has excess electrons, they start diffusing over the junction,

leaving behind a region that has a net positive (electrons moving to p–side) or net negative (holes moving to n–side) charge. The two charges form an electric field that counteracts the diffusion process and after some time brings the semiconductor to an equilibrium state. This whole region is called the depletion layer, as we have no free charge carriers here. The depletion layer depends on the amount of doping of each side, but as a general rule: the side with higher doping will have less depletion. The before and after states are illustrated on figure 7, with corresponding energy level diagrams. The p–n junction is already a very simple ionization
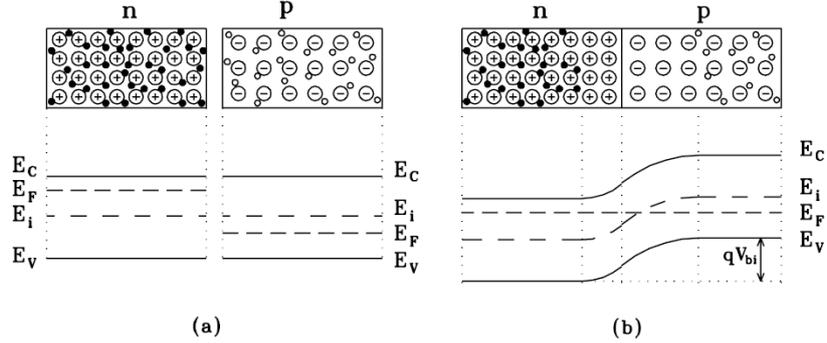

Figure 7: a) The two semiconductor types and their energy levels before forming a junction and b) a p–n junction with a depletion layer. The energy layer designations are for the highest valence energy level E_V, the lowest conduction energy level E_C, the middle E_i and the Fermi energy E_F that is connected to the donor and acceptor energy levels [7].

detector, since any particle that crosses the depletion layer and leaves part of its energy in it, will produce electron–hole pairs. Created electrons subsequently move to the net positive n–side in the presence of an electric field, while holes move in the opposite direction. If we had metal contacts on the semiconductor, we could then just measure the induced electric current.

This simple structure, however, has a very narrow depletion depth of about $100\,\mu$m and since this is the sensitive region of our detector, we need some way to extend it. Another reason to increase the depletion depth is that in such an equilibrium state, the electric field of the semiconductor will be much larger, increasing the efficiency of charge collection. To solve this problem we only need to apply an external electric field by connecting the p–side to a negative potential and the n–side to a positive potential. The negative potential will attract holes away from the junction, further depleting the semiconductor. The story is the same on the n–side, only that the positive potential attracts conduction electrons in this case. This externally applied potential is the so called reverse bias voltage. Reverse, because we are increasing the potential barrier the charge carriers face at the junction. The effect of reverse biasing a p–n junction is shown on figure 8. The width of the depletion layer after applying a reverse bias voltage V can be calculated as

$$d = \sqrt{\frac{2\epsilon(V_0 + V)}{e} \frac{N_A + N_D}{N_A \cdot N_D}}, \qquad (2.3)$$
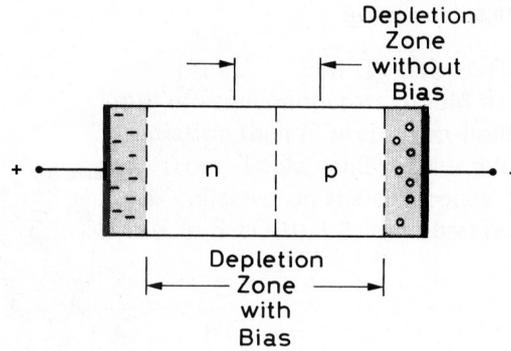
7

Figure 8: The p–n junction with an applied reverse bias (positive on n–side, negative on p–side). Note the increased depletion layer when bias is applied [8].

where V_0 is the inherent p–n junction contact potential (in the range of $1\,\mathrm{V}$ for a silicon p–n junction), ϵ the electric permittivity of the semiconductor, e the charge, while N_A and N_D are the before mentioned acceptor and donor concentrations. We now finally have a large enough sensitive volume to detect ionizing radiation with high efficiency and increased charge collection speed.

3 Passage of ionizing radiation through semiconductors

Whenever particles traverse a layer of matter, there is a non–zero cross–section for their interaction with constituents of this layer. Charged particles will in general interact via the Coulomb force or direct collisions with atoms and their bound electrons. Neutral particles, obviously, can't interact electromagnetically as they carry no charge, but can still collide with atoms in matter and cause nuclear reactions or produce other particles in collisions that in turn cause ionization. The most prominent way of losing energy for particles crossing through semiconductors is the creation of electron–hole pairs known as ionization.

3.1 Cross section and mean free path

Prior to describing different ways for particles to lose energy in matter, we should define the mean free path and cross section, both giving us some information on the probability of interaction. Cross section is a measure of probability for an interaction to occur. More precisely, when we have a flux of particles (a particle beam) heading for a target particle, there is a chance for one or more of them to scatter off the target into a region of space surrounding the target, called the solid angle $\mathrm{d}\Omega = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi$. The differential cross section can then be calculated as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{F}\frac{\mathrm{d}N_s}{\mathrm{d}\Omega}, \tag{3.1}$$

where F is the flux of incoming particles (number of particles per unit time per unit area) and N_s is the average number of scatterings per unit time. The total

cross section would then just be calculated by integrating equation (3.1) over the solid angle. The above example is clearly a generalization for whenever we are only considering a single particle as the target, but in reality this is not the case. A more realistic example involves a target with thickness δx and density of scattering centers N together with an incident particle beam of flux F and perpendicular area A. The average number of scatterings into a solid angle is thus

$$N_s(\Omega) = F \, A \, N \, \delta x \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}. \tag{3.2}$$

For a limiting case, when the beam area is smaller than the target, we can rewrite $n_{\mathrm{inc}} = FA$, where n_{inc} is the number of incident particles per unit time.

Mean free path is defined as the mean distance a particle can travel without colliding with atoms in the target and is calculated as

$$\lambda = \frac{\int x P(x) \, \mathrm{d}x}{\int P(x) \, \mathrm{d}x}, \tag{3.3}$$

where $P(x)$ is the probability for not having any interactions after distance x. This can be determined by checking what is the probability for not having any interactions after a farther distance $\mathrm{d}x$

$$P(x + \mathrm{d}x) = P(x)(1 - w \, \mathrm{d}x). \tag{3.4}$$

Here, $w \, \mathrm{d}x$ designates the probability for an interaction on interval between x and $x + \mathrm{d}x$, so $1 - w \, \mathrm{d}x$ is just the opposite (no interaction on the interval). Rewriting the probability in equation (3.4) as

$$P(x + \mathrm{d}x) = P(x) + \frac{\mathrm{d}P}{\mathrm{d}x} \, \mathrm{d}x,$$

results in a differential equation

$$\mathrm{d}P(x) = -w P(x) \, \mathrm{d}x,$$

with solution

$$P(x) = C \exp(-wx), \tag{3.5}$$

where $C = 1$, since the initial probability should be $P(0) = 1$. There are definitely no interactions before the incident particle even starts traversing through matter. Placing the solution into equation (3.3), we get a simpler description of the mean free path

$$\lambda = \frac{1}{w} = \frac{1}{N\sigma}, \tag{3.6}$$

where $w = N\sigma$ can be concluded by calculating the probability of an interaction in target thickness δx from the total number of scattered particles $N_{\mathrm{tot}} = \int N_s(\Omega) \, \mathrm{d}\Omega$.

3.2 Energy loss of particles in semiconductors

Semiconductor detectors, introduced in section 2, are usually made fairly thin (maximum thickness on the order of cm) and therefore have a low probability of atomic

collisions that are typically highly unwanted. The principal way of energy loss will be through the inelastic collisions with electrons, producing ionized pairs of electrons and holes. Ionization happens whenever a bound electron in the target receives enough energy from an incident particle for it to become free. While charged particles cause ionization via the Coulomb force, it is the neutral particles that have an indirect interaction with them, overall resulting in less interactions with matter. In general energy loss calculations, we use the Bethe–Bloch formula

$$ -\frac{\mathrm{d}E}{\mathrm{d}x} = 2\pi N_a r_e^2 m_e c^2 \rho \frac{Z}{A}\frac{z^2}{\beta^2}\left[\ln\left(\frac{2m_e\gamma^2 v^2 W_{\max}}{I^2}\right) - 2\beta^2 - \delta - 2\frac{C}{Z}\right], \qquad (3.7) $$

where $2\pi N_a r_e^2 m_e c^2 = 0.1535\,\mathrm{MeVcm^2/g}$ are purely constants and I is the mean excitation potential determined directly from measurements. Properties of the target material ρ, Z and A are density, atomic number and atomic weight, respectively. z is the atomic number of the incident particle, $\beta = v/c$ its relativistic speed and $\gamma = 1/\sqrt{1-\beta^2}$ the Lorentz factor. $W_{\max}$ is the maximum transfer of energy produced in a head–on collision, with $W_{\max} \simeq 2m_e c^2(\beta\gamma)^2$ for the case when mass of the incident particle is much larger than mass of electron, which is certainly the case here. δ is the density correction due to the polarization of atoms when an incident particle passes by, creating a sort of electromagnetic barrier and shielding bound electrons farther from the incident particle. C, on the other hand, is the shell correction arising from a comparable velocity of incident particle and orbital velocity of bound electrons. These electrons can therefore no longer be treated as stationary. Shell corrections are prominent at low energies, while density corrections are important at high energies. A typical stopping power function plotted from the Bethe–Bloch formula (3.7) is displayed on figure 9 along with both corrections.

For incident electrons and positrons, the Bethe–Bloch formula can be simplified,

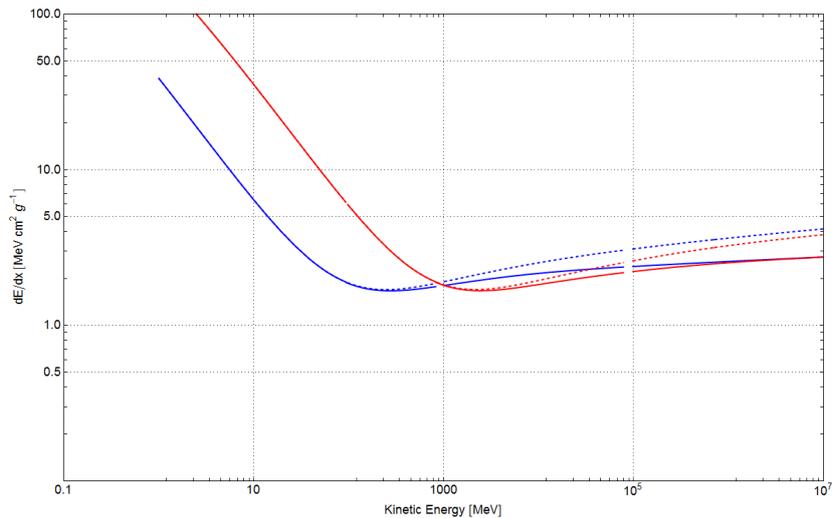Figure 9: Stopping power for a muon (blue) and proton (red) in silicon. The dashed lines are the same stopping powers without density δ and shell corrections C. Kinetic energy calculated from total relativistic energy and rest mass is $T = E - mc^2$.

due to the fact that they have mass equal to the bound electrons and will lose energy through elastic collisions with them. On the other hand, because of their

small mass they can quickly be affected by the electric field around atoms, emitting photons through Bremsstrahlung radiation (radiation due to deceleration). Its effect becomes important at high energies. The Bethe–Bloch formula for electrons and positrons is thus rewritten as

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = 2\pi N_a r_e^2 m_e c^2 \rho \frac{Z}{A} \frac{1}{\beta^2} \left[\ln \left(\frac{\tau^2(\tau+2)}{2(\frac{I}{m_e c^2})^2} \right) + F(\tau) - \delta - 2\frac{C}{Z} \right], \qquad (3.8)$$

where $\tau = T/m_e c^2$ is the kinetic energy T in units of $m_e c^2$, $z = 1$ for electrons and positrons, $F(\tau)$ are functions depending on incident particle type, kinetic energy and relativistic speed, while the rest has already been described in equation (3.7). Figure 10 shows the stopping power for electrons in copper with added Bremsstrahlung radiation and proton stopping power for comparison.

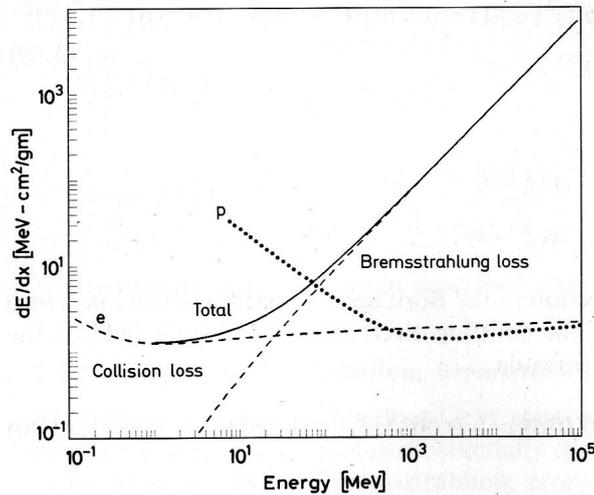When trying to detect photons with semiconductors, we need a material with

Figure 10: Stopping power for an electron (solid line) and a proton (dotted line). The dashed lines are radiation and collision loss separated, with equal energy loss at the critical energy [8].

a higher atomic number (such as germanium), because of their lower interaction probability. Photons can interact with matter in three different ways: photoelectric effect, Compton scattering and pair production. The photoelectric effect involves an absorption of a photon by a bound electron that is subsequently ejected, creating an ionized pair. Compton scattering is the scattering of a photon off a free electron, with the photon losing energy and electron acquiring kinetic energy. Due to the structure of semiconductor detectors, Compton scattering is less likely to happen, as there are no free electrons in the depletion zone. The third energy loss process is through the creation of pairs, an electron and positron, that can further cause ionizations in the semiconductor. The minimum energy of a pair producing photon needs to be at least 1.022 MeV, the rest masses of electron and positron combined. The main physics field that uses the detection of photons with semiconductors is spectroscopy, detecting high energetic X–rays and γ–rays.
Neutrons have a very small probability of interacting with matter, due to not carrying charge, but can still be detected with semiconductors. They can, in general,

collide with atoms in the matter, get captured by the atom or cause nuclear reactions if their energies are low enough and produce hadron showers at very high energies. Through the principal of ionization, semiconductors are most likely to detect neutrons according to the latter energy loss process.

4   Properties of semiconductor detectors

4.1   Ohmic contact

The depletion width in equation (2.3) can be decomposed into two parts, depletion width on the p– and on the n–side of the semiconductor junction. It stands to reason that, if one of the two sides has a higher doping, the number of free charge carriers will be larger on this side after equilibrium state is reached. This in turn means that the side with higher doping will also have a smaller depletion width. However, in the case of a metal–semiconductor junction we observe depletion only on the semiconductor side, due to an almost infinite reserve of electrons in metal. The potential barrier that is created in such a structure will block the current flowing from semiconductor to metal, not producing the ohmic contact needed for it to pass from detector to the cable connected to it. Luckily, the solution involves only a thin layer of highly doped semiconductor, designated as n$^+$ or p$^+$, placed in–between the semiconductor and metal. The depletion zone on the new metal–semiconductor junction will now be smaller, creating an ohmic contact through the tunnelling effect. Although, this seems like a simple approach, the problem is that the n$^+$ and p$^+$ layers have virtually no sensitive region and will not contribute to sensing particles passing through the semiconductor. This non–sensitive layer is called the dead layer and is one of the main properties, when discussing different types of semiconductor detectors.

4.2   Average ionization energy, energy resolution, leakage current and radiation damage

Semiconductor detectors have many advantages over other types of detectors due to their structure and superior sensing abilities of ionizing radiation. Starting with the average energy per electron–hole pair and comparing it to the average energy per electron–ion pair (naming both as average ionization energy w) of gases, we see that the lower limit for particle detection in gases is higher than in semiconductors, with the values around $w_{\text{semi}} \simeq 3\,\text{eV}$ for semiconductors and $w_{\text{gas}} \simeq 30\,\text{eV}$ for gases. A particle with energy E would on average produce E/w electron–hole or electron–ion pairs.

Directly tied to average ionization energy w, is the detector resolution R. In general, two Gaussian peaks can be resolved if the separation between them is greater than their full widths at half maximum (FWHM). If we want to determine the energy resolution of a detector, with peak width ΔE, then the resolution can be given as $R = \Delta E/E$. With the use of the Gaussian distribution, this can be rewritten as

$$R = 2.35\sqrt{\frac{Fw}{E}}, \tag{4.1}$$

where E is the energy of an incident particle, the factor $2.35 = 2\sqrt{2\ln 2}$ relates FWHM and standard deviation values and F is the Fano factor. The reason to include the Fano factor is because energy loss of a particle passing through matter is not purely statistical in nature, ionization is actually limited by discrete values in electron shells of an atom. While the Fano factor is still not well known, it has a value on the order of 0.12 for semiconductors and 0.2 for gases. Just as an example, table 1 holds average ionization energies and resolutions for semiconductor and gaseous detectors. In the table, the properties for germanium are only described at liquid nitrogen temperature (77°C), because its band–gap is smaller ($\sim 0.7\,\text{eV}$), causing electrons to pass to the conduction band at room temperature. Another

Table 1: Average ionization energies and detector resolution for silicon, germanium (semiconductors), argon and argon–krypton mixtures (gases). Resolution is given for an incident particle with energy $E = 3\,\text{keV}$ [8].

Element	w (eV)	T (K)	Fano factor F	R@3 keV (%)
Si	3.81	77	0.115	2.84
	3.62	300	0.115	2.77
Ge	2.96	77	0.13	2.66
Ar 100%	26	—	0.2	9.78
Ar 80% + Kr 20%	~ 25	—	< 0.12	7.43

property that determines the quality of a detector is leakage current. While we might expect that a reverse biased p–n junction doesn't let any charge carriers cross the junction, this is certainly not the case in practice. Leakage current, or more commonly known as detector noise, is a small current passing through the junction, limiting the lowest current that we can effectively measure. In reality, the limiting value is actually even higher, if we want to say that what we are measuring is not just noise. A measure of how well a signal from a passing particle separates from the noisy background is called signal–to–noise ratio (SNR) and, evidently, we want it to be as high as possible (with SNR = 1 meaning that signal and noise levels are equal).

An important property to note, when using semiconductors as detectors, is that they are not as resilient to radiation damage as other detector types are. The main issue here is that incident particles can collide with lattice atoms and break the crystal structure. When that happens, we get a defect, causing additional energy levels to form inside the band–gap, increasing leakage current of the detector and degrading detector resolution.

4.3 Pulse shape

For the calculation of a pulse shape, we take the simplest structure that we can find, a p–n junction with the depletion region entirely on the p–side (highly doped n–side). A positive bias voltage is applied to the n$^+$ layer, while the p–side is grounded. This kind of structure reminds us of two capacitor plates and a dielectric in the middle (the depleted region), so we can simply use equations for a parallel

plate capacitor

$$\mathrm{d}W = \frac{e\varphi}{d}\,\mathrm{d}x, \quad \mathrm{d}W = \varphi\,\mathrm{d}Q \quad \rightarrow \quad \mathrm{d}Q = \frac{e}{d}\,\mathrm{d}x, \tag{4.2}$$

where Q is the collected electric charge, W the potential energy, φ the potential, e the charge of a single charge carrier and d the separation between the capacitor plates (width of depletion zone, in our case). The geometry of this problem is shown on figure 11. The connection between potential φ and charge distribution ρ
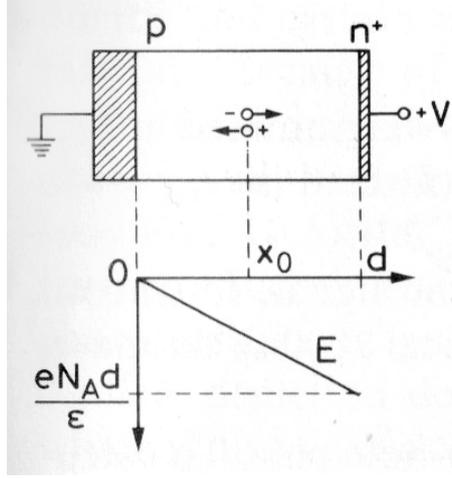

Figure 11: Geometry of a semiconductor detector for the purpose of calculating the pulse shape. The bottom graph shows the electric field inside the depletion zone [8].

is described by the Poisson equation

$$\frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2} = -\frac{\rho(x)}{\epsilon}, \tag{4.3}$$

with $\rho(x)$ being $\rho(x) = -eN_A$ for our simplified case with depletion on the p–side. Calculating the electric field inside the detector is just a matter of integrating equation (4.3) once to get

$$E = -\frac{\mathrm{d}\varphi}{\mathrm{d}x} = -\frac{eN_A x}{\epsilon} = -\frac{\sigma x}{\epsilon\mu_h} = -\frac{x}{\mu_h\tau}, \tag{4.4}$$

where μ_h, N_A, e and ϵ are the usual hole mobility, acceptor concentration, electric charge of the charge carrier and electric permittivity ($\epsilon = \epsilon_r\epsilon_0$), respectively. To replace some values with easier–to–measure properties of a semiconductor, we've used a relation for the conductivity σ of a p–type semiconductor $\sigma = eN_A\mu_h$ and defined $\tau = \epsilon/\sigma$.

If we now have a particle passing through our semiconductor detector and creating one electron–hole pair at a point x_0 inside the depleted zone, each will have a different drift speed towards the respective electrode. Using equations (2.1) with an added negative sign for an electron (according to figure 11, it's moving in a direction, opposite to the electric field direction), combining it with equation (4.4)

14

and integrating them both, results in a time evolution of the path taken by the two created charge carriers

$$x_e(t) = x_0 \exp\left(\frac{\mu_e t}{\mu_h \tau}\right) \qquad \text{and} \qquad x_h(t) = x_0 \exp\left(-\frac{t}{\tau}\right), \qquad (4.5)$$

with index e designating an electron and h a hole. With the use of equation (4.2), we can now finally calculate the charge collected at both electrodes

$$Q_e(t) = -\int_0^t \frac{e_0}{d} \frac{\mathrm{d}x_e(t)}{\mathrm{d}t}\,\mathrm{d}t = \frac{e_0}{d}x_0\left(1 - \exp\left(\frac{\mu_e t}{\mu_h \tau}\right)\right),$$

$$(4.6)$$

$$Q_h(t) = \int_0^t \frac{e_0}{d} \frac{\mathrm{d}x_h(t)}{\mathrm{d}t}\,\mathrm{d}t = -\frac{e_0}{d}x_0\left(1 - \exp\left(-\frac{t}{\tau}\right)\right),$$

where e_0 is the elementary charge (for electron $e = -e_0$, for hole $e = e_0$). Evidently, the electron will accelerate when coming closer and closer to the positive electrode, all along inducing a charge in the electrode. On the other hand, the hole will never reach the negative (or grounded) electrode, but will still induce a charge in this electrode, contributing to the total collected charge $Q_{\mathrm{tot}} = Q_e + Q_h$. For easier visualization, figure 12 displays the collected charge from one electron–hole pair with typical semiconductor detector parameters for d and τ. This calculation
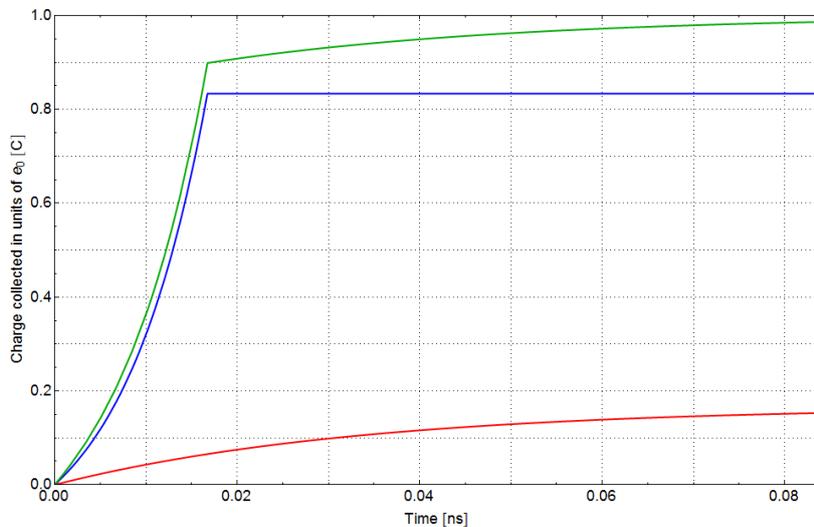

Figure 12: The collected charge for an electron–hole pair for a typical semiconductor detector with $d = 300\,\mu\mathrm{m}$ and $\tau = 3.6 \times 10^{-2}\,\mathrm{ns}$. The ionization happens at $x_0 = 50\,\mu\mathrm{m}$. The plot shows the charge collected on the positive electrode Q_e (blue), on the grounded electrode Q_h (red) and the total collected charge Q_{tot} (green).

is only for a single electron–hole pair, but a passing particle will produce many more, making the simulation harder and with added effects, like charge repulsion and variation of mobilities. We must also take into account that figure 12 displays only the collected charge at the contacts. In reality, we should observe a discharge similar to a capacitor, resulting from current passing through the connector cable and further towards the amplifier.

5 Semiconductor detector types

5.1 Diffused junction and ion–implanted diodes

The main difference between these two types of silicon detectors is the fabrication process, bestowing each some advantages and drawbacks.

A diffused junction diode is made by using a p–type semiconductor, placing it into a diffusion furnace and letting a gas flow of n–type dopant diffuse into it. For the diffusion process to work properly, we need to heat up the semiconductor to around 1000°C, producing junctions with depths between $0.1\,\mu m$ and $2\,\mu m$ and a depletion region with thickness of $\sim 300\,\mu m$. The diffused junction diodes are thus quite resilient to any surface contamination and are easier to produce. However, the high temperature creates defects inside the detector and increasing detector noise, while reducing lifetime. The thick dead layers set a lower limit to the incident particle energy.

An ion–implanted diode, on the other hand, is made by accelerating dopant atoms and bombarding a p–type semiconductor. The structure of a crystal lattice in the semiconductor is damaged in the process and must be annealed at $\sim 500°C$. We can see quite quickly that such a detector will have lower noise than a diffused one, with dead layer thickness of only $\sim 50\,nm$ and depletion region thickness of $\sim 300\,\mu m$. Ion–implanted diodes are widely used, but are fairly expensive and harder to make than diffused junction diodes. Figure 13 shows a range of ion–implanted diode detectors.

Figure 13: A range of ion–implanted diode detectors in metal casings. The front side has a circular opening of slightly darker colour, called the detector active area [9].

5.2 Surface barrier detectors (SSB)

Surface barrier detectors are special in the sense that they don't have a p–n junction, but a metal–semiconductor junction. The production process is simple and just involves depositing a thin gold (for n–type semiconductor) or aluminum (for p–type semiconductor) layer onto the semiconductor surface through evaporation. The usual thickness of this metallic layer is $\sim 40\,\mu g/cm^2 \sim 20\,nm$ and the depletion region, completely on the side of the semiconductor, can be made up to thicknesses of $\sim 5\,mm$. SSB detectors are thus well suited for energy measurements due to their thick depletion zone and small dead layer, but are very sensitive to surface contamination and must be kept in a metal encasing, because light can pass through the thin metal layer. Figure 14 explains the structure of a typical SSB detector.
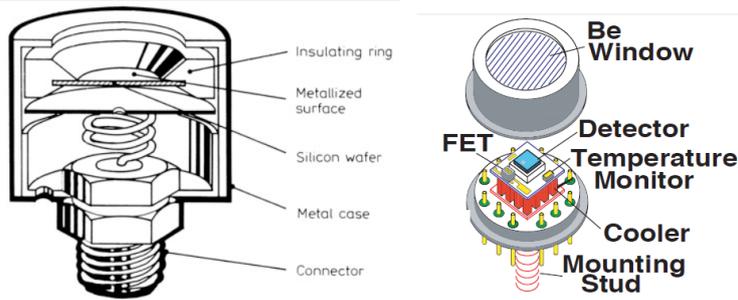
Figure 14: The cross section (left) of a typical SSB detector and an example of mounting such a detector (right) [10].

5.3 Lithium–drifted silicon diodes (p–i–n junction)

In order to measure less interacting particles, we need to have a detector with a large enough sensitive volume. The depletion width of normal p–n junction diodes and surface barrier detectors can be surpassed by using lithium–drifted detectors. The fabrication process involves placing a p–type semiconductor in a lithium bath, letting lithium ions diffuse into the semiconductor and negate the effect of p–type doping. After applying an external potential, the lithium ions will drift further into the crystal, producing an intrinsic region, sensitive to passing ionizing radiation. Such compensated regions can be made up to $\sim 10 - 15\,\mathrm{mm}$, severely increasing the sensitivity for less interacting particles (like β particles, X–rays or γ–rays) and making a p–i–n junction detector well suited for energy measurement. In exchange for a large sensitive volume, we do need to cool them with liquid nitrogen to reduce the noise coming through thermal separation of lithium ions from acceptor atoms. The geometry used is generally coaxial to maximize the sensitive volume, with several possibilities as shown on figure 15. Alternatively, for even higher resolution
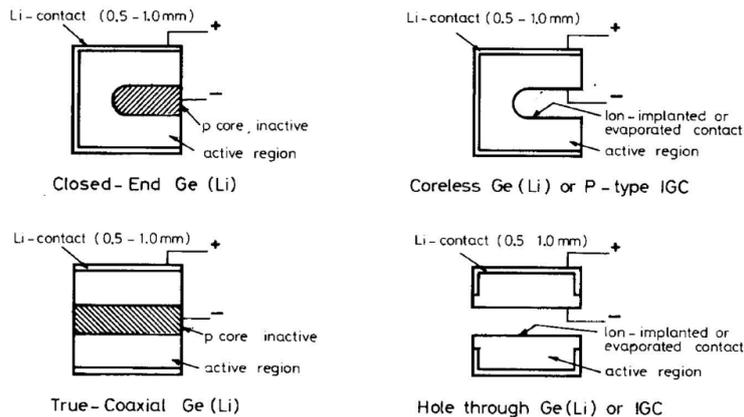

Figure 15: Coaxial geometries for lithium–drifted detectors. Plus and minus signs determine the two contacts. The white region is the sensitive volume, while the shaded one is inactive – part of the detector dead layer [8].

of a detector, we can use germanium instead of silicon as the semiconductor material, due to its higher atomic number Z (better stopping power) and higher cross section for γ–rays.

17

5.4 Position–sensitive detectors

Another use of semiconductor detectors is in determining the flight path of incident particles. Similar to proportional counters (gas detectors with wires at high voltage), they incorporate a position–sensitive readout, but take up less space and are thinner so passing particles lose only a tiny fraction of their energy, making them ideal for large detection systems.

5.4.1 Continuous detectors

In the continuous readout method, a diode has a uniform resistive electrode on one side and a low–resistive electrode on the other side, as shown on figure 16. The resistive electrode has two contacts, one for grounding (A) and the other for bias voltage and measuring the signal from the passing particle (B). The other side has
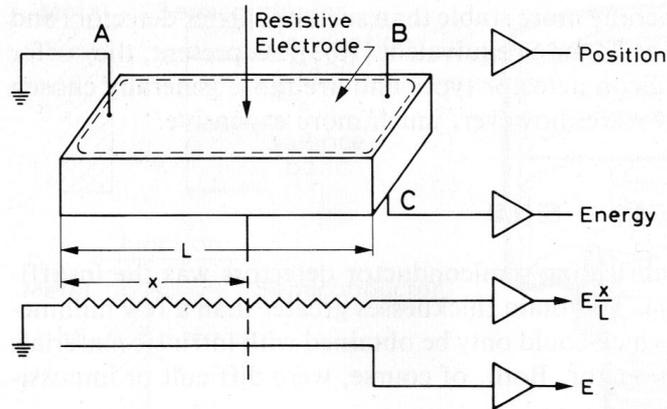

Figure 16: A continuous position–sensitive detector with resistive charge division. The upper electrode is uniformly resistive, while the bottom is low–resistive (also shown on the simplified circuit representation). Contacts B and C are vertically at the same point [8].

another contact for measuring the signal (C) at the same vertical point as contact B. A passing particle will induce charge on both electrodes, by producing electron-hole pairs, with charge being proportional to the energy. However, due to the resistance of one electrode, the signal from it will also depend on the distance between the contact B and point of incidence

$$B \propto E \frac{x}{L}, \quad C \propto E \quad \longrightarrow \quad x \simeq L \frac{B}{C}. \tag{5.1}$$

Essentially, we are measuring the difference between the charge collected at both contacts due to different electrode resistances, a method called resistive charge division. Such detectors have a spatial resolution on the order of $\sim 250\,\mu$m.

5.4.2 Strip and micro–strip detectors

Strip detectors or discrete detectors are semiconductor detectors that, instead of a continuous readout, have electrode strips on the surface of a diode structure. Each

of these electrodes detects a portion of the passing particle's energy and output different signal heights. By using a weighted average of the collected signal from all strips, we can then determine, in one dimension, the position where the particle passed through the detector. Furthermore, we can perpendicularly apply another set of electrode strips on the other side of the detector surface to create a two dimensional position–sensitive detector. A schematic of such a two–dimensional detector or a matrix detector is shown on figure 17. However, the number of strips and thus the number of readout channels can quickly become extremely large, so in order to reduce them, we can use a resistive or a capacitive–division readout method. In both cases, we directly read the signal from every second, third, fourth,... electrode strip. A resistive–division method essentially creates a detector similar to
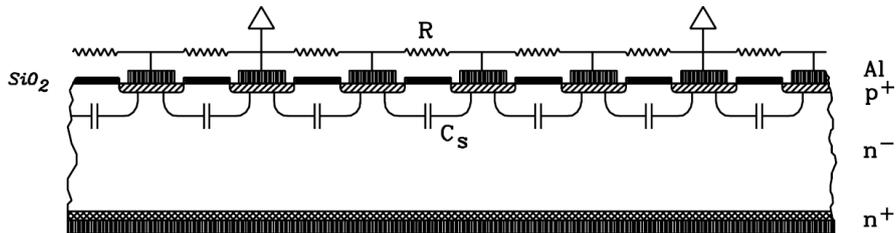

Figure 17: A two–dimensional strip detector. From top to bottom, the layers are: resistive SiO_2 layers in black, aluminum electrodes in shaded grey, p^+–type in diagonal line shading, high resistive n–type in white, n^+–type in dotted shading and then again the electrodes and oxide layers. The resistor and capacitor symbols indicate a capacitive–division readout method [7].

the continuous detector, while a capacitive–division method, shown in figure 17, uses the internal structure of the detector to its advantage. The strips of p^+–type semiconductor have an inherent capacitance between them and the oxide layers serve as resistors. That way, the signal at the electrode with direct readout will still receive a contribution from nearby strips, making use of the electrodes with no direct readout to achieve higher resolution. Whichever readout method we use, the spatial resolution of strip detectors is determined by the separation between the strips called the pitch. A quick indication of the spatial resolution can be calculated from [11]

$$R_{\mathrm{spatial}} = \sqrt{\frac{1}{p} \int_{-p/2}^{p/2} x^2 \, \mathrm{d}x} = \frac{p}{\sqrt{12}}, \tag{5.2}$$

where p is the pitch and the factor $1/\sqrt{12}$ comes from the calculation of a root–mean–square deviation from the center of the strip. Typical pitch values for strip detectors is on the order of a few hundred microns, which in turn gives us a spatial resolution up to $\sim 100\,\mu\mathrm{m}$.

Micro–strip detectors were the next logical step after technology advanced enough to create even finer structures on diode detectors. They have a similar structure and readout methods to strip detectors, but the smaller pitch $p \sim 20\,\mu\mathrm{m}$ bestows them with a remarkable spatial resolution of $\sim 5\,\mu\mathrm{m}$. The overall advantage of the strip

and micro–strip detectors, apart from its high resolution, is the detector thickness, which is around $300\,\mu\text{m}$, where particles lose only a fraction of their energy, but on average still produce 100 electron–hole pairs per micron. Figure 18 shows a silicon tracker, consisting of several micro–strip detectors, where the rectangular shape on each ladder determines one micro–strip detector.
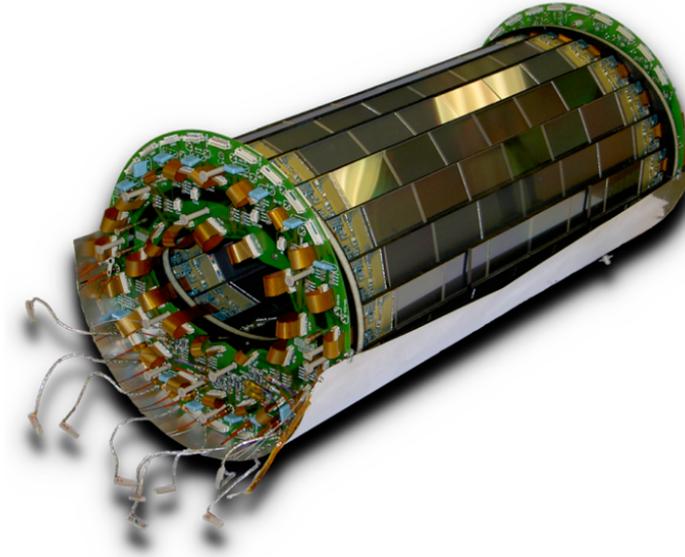

Figure 18: A silicon tracker from the H1 detector at DESY accelerator. Each small rectangular shape is a micro–strip detector arranged into a ladder structure. In total, there are 192 detectors (each of size $3.4 \times 5.9\,\text{cm}^2$) with 81920 readout channels [12, 13].

5.4.3 Pixel detectors

Pixel detectors are very similar to micro–strip detectors, but their structure is planar, spread out into arrays or detector grids. The bottom part of the detector grid is connected directly to the electronics grid by bump bonding them together. As figure 19a suggests, bump bonds are made when metal bumps (contacts) on one
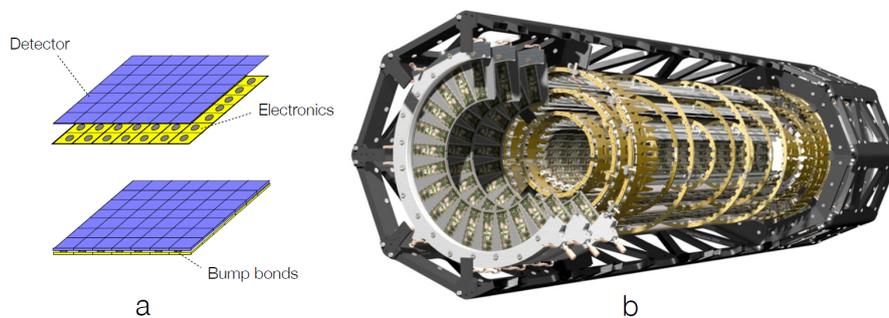

Figure 19: Schematic of a) a pixel detector [10] and b) the pixel detector mounted at the ATLAS experiment at LHC [14]. The ATLAS pixel detector has 1456 modules (each of size $62.4 \times 21.4\,\text{mm}^2$) with 46080 pixels in total.

plate are heated enough to become liquid and then another plate is placed into close proximity for the metal to connect both plates. Compared to micro–strip detectors, such a detector gives us a higher occupancy of detectors per area, comparable spatial resolution and lower noise levels due to lower capacitance of the detector. However, bump bonding technique is fairly new and complicated and the amount of readout channels further increases. The largest pixel detector currently in operation is the one mounted at the ATLAS experiment at CERN's large hadron collider, shown schematically in figure 19b.

5.4.4 Drift detectors

Precisely because micro–strip and pixel detectors have such a large number of read-out channels, a silicon drift chamber was proposed as a new position–sensitive type of semiconductor detectors. It is fabricated by depositing multiple p^+ contacts on both sides of an n–type semiconductor and then applying a single n^+ contact as the last contact instead of a p^+ on one side. The general idea behind the structure of a drift detector is shown in figure 20. The p^+ contacts on both sides create a depleted
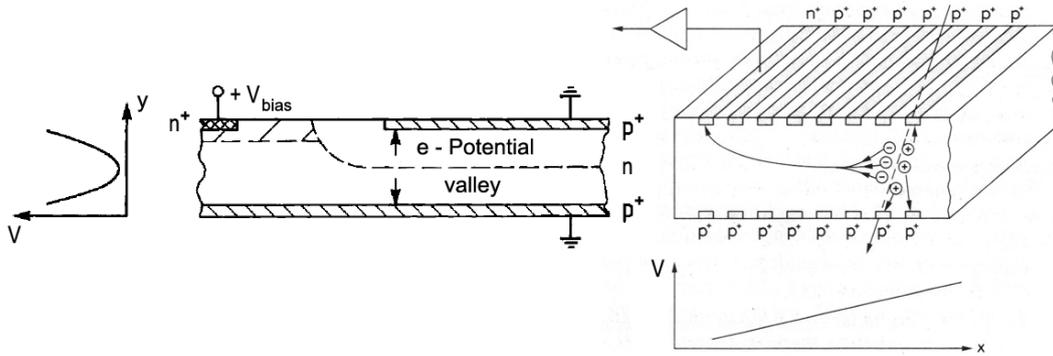

Figure 20: The structure of a drift detector. When a reverse bias is applied, the central region has a parabolic potential. From an electron–hole pair, the electron will fall to the center and drift towards the n^+ contact, while holes will find the nearest contact and give us the staring time [7, 8].

region from both sides with a central active region. Applying a reverse bias on the detector will then completely deplete the detector, forming a parabolic potential in the depleted region. The minimum lying at the center of the detector ensures electrons formed in ionizations fall down to the center of the detector and then drift towards the n^+ contact as displayed in figure 20. All we need to do now is measure the time it takes the electrons to complete this path, started by the collection of holes on all the other contacts and ended at the collection of the electron in the last contact. Apart from the smaller number of electronics needed for a drift detector, it's advantage is also lower capacitance and a spatial–resolution comparable to micro–strip detectors. The STAR vertex tracker (figure 21) that consists of silicon drift detectors, for instance, has a spatial resolution of $\sim 20\,\mu\mathrm{m}$.
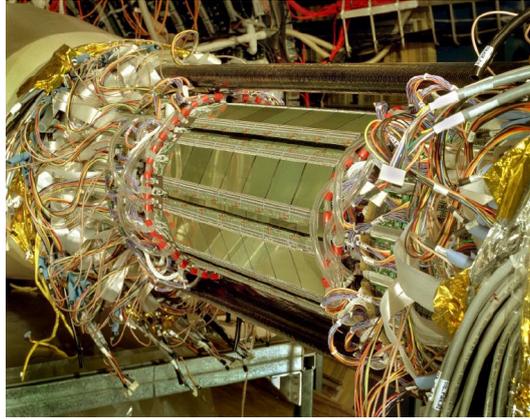
Figure 21: An example on the use of drift detectors is the STAR vertex tracker with 216 silicon drift detectors arranged on 36 ladders and 1300 readout channels in total [15].

6 Signal transmission, amplification and shaping

A signal gathered directly from a detector is usually very small and needs to be amplified to get a measurable quantity. However, transmission from the detector to the amplification stage adds some more noise to the signal and must therefore be shaped for it to be efficiently used and saved on data acquisition systems (DAQ systems).

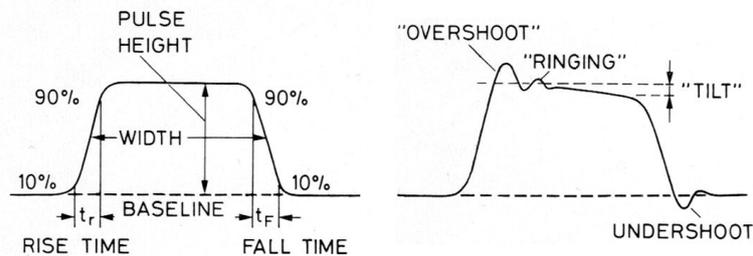A general pulse will look something like the one shown in figure 22. Baseline is

Figure 22: Terminology used for a general pulse signal [8].

the signal level to which the pulse will decay to, not necessarily zero or the lowest value of the pulse. The height of the pulse is measured as the distance between the baseline and maximum. Its width is measured as the width of the pulse at half of the pulse height (FWHM) of the signal. Rise times and fall times are measured as the time it takes the pulse to get from 10% to 90% of its value and vice versa. Some special terminology involve overshoots (value over pulse height on the leading edge), undershoots (value under baseline on the falling edge), ringing (resonant variation of the signal) and tilt (tilt of signal step).

6.1 Fourier transform and the frequency domain

Some of the information of the signal can be lost when looking at it in the time domain and can therefore be transformed to the frequency domain via the Fourier

transform

$$g(\omega) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} f(t) \exp(-i\omega t)\,\mathrm{d}t, \qquad (6.1)$$

with a reverse transformation

$$f(t) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} g(\omega) \exp(i\omega t)\,\mathrm{d}\omega. \qquad (6.2)$$

In both transformations t is time and ω is frequency as expected. For example, we might have a very complicated signal in the time domain and realize that in the frequency domain it is constructed from only a few principal frequencies — something like a chord in music is made up from a collection of notes at a given frequency, but is complicated in the time domain. Obviously, a signal from a detector is even more complex than that and thus the frequencies that make up the pulse are usually on an infinite range. By creating a cut-off of frequencies, a fault of any cable or electronic device, we are actually attenuating the frequency response by more than 3 dB on some finite range. A way to visualize this is by plotting a frequency response curve (figure 23) that tells us the frequencies that will be passed with virtually no attenuation (passband or bandwidth) and frequencies that will be stopped (stopband).
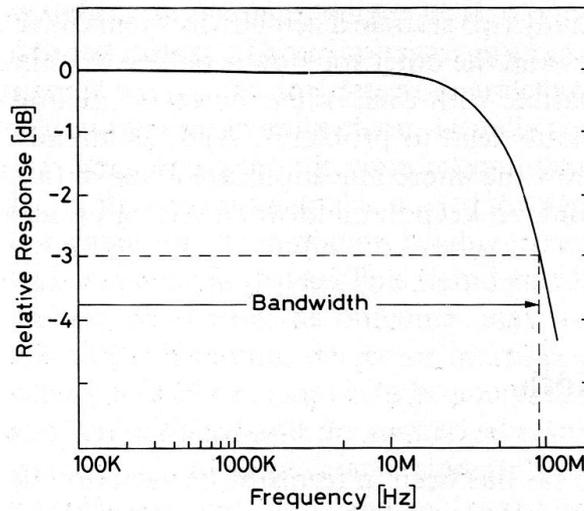

Figure 23: An example of a frequency response curve, detailing the attenuation of frequencies due to, for example, an electronic device or cable. 0 dB means an unchanged signal, positive values would give amplification and negative attenuation [8].

6.2   Coaxial cable signal transmission

As described in the previous section, the biggest problem in signal transmission is the limited frequency range of all cables and the pickup of stray electromagnetic fields in the vicinity. Therefore, selecting the right cable for the job is the difference

between a clean signal and a distorted signal riddled with noise. The principal solution to some of these problems is using the shortest possible connection between the detector and first amplification stage (preamplifier), that reduces any noise we might pick up. Warding against stray EM fields, on the other hand, is best done by using shielding around the central carrying wire. Such cables are called coaxial cables, constructed from a conducting wire, conducting screen, a dielectric between them and a plastic sheath. The conducting screen is made up from a wire braid, carries the return current and shields the wire from EM fields through the Faraday effect. The usual choices for a dielectric are polyethylene plastic or teflon, while the plastic sheath just protects the wire from any electrical contact. Figure 24 showcases a typical coaxial cable structure. There are a number of different types
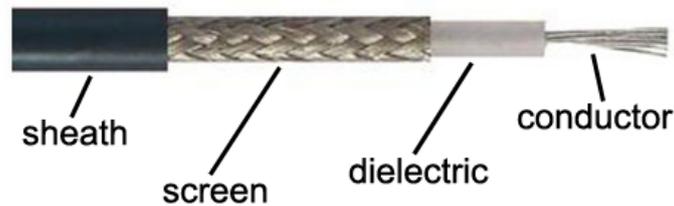

Figure 24: The structure of a coaxial cable (RG–58 as an example). A dielectric is placed in–between the central conductor and a conducting screen (wire braid). Everything is encased in a plastic sheath, preventing any electric contact with other metals [16].

of coaxial cables, depending on the applications used (signal cables, high voltage supply), maximum operating voltages and impedances. For any coaxial cable we can determine its self–inductance and self–capacitance, taking into account that the whole wire is actually a cylindrical capacitor and the braid creates a coil structure. Both can be calculated with the use of

$$
\begin{aligned}
L &= \frac{\mu}{2\pi} \ln\left(\frac{b}{a}\right) \left[\frac{\mathrm{H}}{\mathrm{m}}\right] = 0.2 K_m \ln\left(\frac{b}{a}\right) \left[\frac{\mu\mathrm{H}}{\mathrm{m}}\right], \\
C &= \frac{2\pi\epsilon}{\ln(b/a)} \left[\frac{\mathrm{F}}{\mathrm{m}}\right] = \frac{55.6\epsilon_r}{\ln(b/a)} \left[\frac{\mathrm{pF}}{\mathrm{m}}\right],
\end{aligned}
\tag{6.3}
$$

where a and b are inner and outer conductor radii, $K_m = \mu/\mu_0$ and $\epsilon_r = \epsilon/\epsilon_0$ are relative permeability and relative permittivity of the dielectric and the units used are given in square brackets. For dielectrics, the relative permeability is usually just $K_m \simeq 1$. A widely used RG–58 coaxial cable for signal transmission with polyethylene dielectric ($\epsilon_r \simeq 2.3$, $K_m \simeq 1$) has inductance $L = 0.253\,\mu\mathrm{H/m}$ and capacitance $C = 101\,\mathrm{pF/m}$.

6.2.1 Wave equation for a coaxial line

An electric circuit representation of a coaxial cable incorporates:

- a resistor R: simulating the resistance of the central conductor that should be small for a good cable,

- an inductor L: describing the magnetic field due to the coil structure of the wire braid,

- a capacitor C: gives the capacitance of the conductor–dielectric–screen structure that is basically a cylindrical capacitor,

- another resistor $1/G$: simulating the current passing through the conductor (leakage current) that should be very large (or small, if we're discussing the value of conductance G),

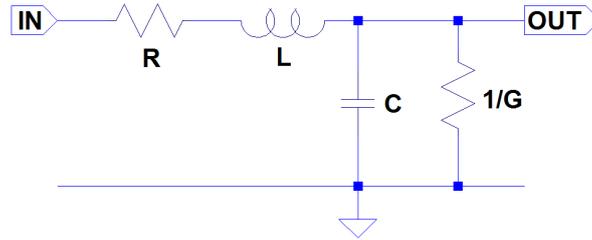and is constructed as shown on figure 25. When going from the input to the output

Figure 25: Electric circuit for a length of a coaxial cable. For $1\,\mathrm{m}$ of cable, L and C are self–inductance and self–capacitance as described by equation (6.3).

of the cable, we have a voltage drop on R and L and we lose some of the current, due to leakage through C and $1/G$. This can be described by

$$
\begin{aligned}
\frac{\partial V(z,t)}{\partial z} &= -R\,I(z,t) - L\,\frac{\partial I(z,t)}{\partial t} \\
\frac{\partial I(z,t)}{\partial z} &= -G\,V(z,t) - C\,\frac{\partial V(z,t)}{\partial t}.
\end{aligned}
\tag{6.4}
$$

Differentiating both equations with respect to z and t, we can decouple them and get one differential equation only depending on voltage $V(z,t)$ and the other on current $I(z,t)$. Showing the calculation for the voltage part gives us

$$
\frac{\partial^2 V}{\partial z^2} = -R\,\frac{\partial I}{\partial z} - L\,\frac{\partial^2 I}{\partial t \partial z} = -R\left(-G\,V - C\,\frac{\partial V}{\partial t}\right) - L\left(-G\,\frac{\partial V}{\partial t} - C\,\frac{\partial^2 V}{\partial t^2}\right),
$$
$$
\frac{\partial^2 V}{\partial z^2} = L\,C\,\frac{\partial^2 V}{\partial t^2} + (R\,C + L\,G)\,\frac{\partial V}{\partial t} + R\,G\,V,
\tag{6.5}
$$

obviously still taking into account that voltage and current depend on z and t. The second order differential equation for current is identical, but with I instead of V.

6.2.2 Lossy coaxial cables

The solution for differential equation (6.5) can be decomposed into a time dependent and position dependent part

$$
V(z,t) = V(z)\,\mathrm{e}^{i\omega t},
\tag{6.6}
$$

where the time part is a simple sinusoidal wave. Placing this into equation (6.5) will give us a simplified second order differential equation

$$\frac{\partial^2 V}{\partial z^2} = (R + i\omega L)(G + i\omega C) V = \gamma^2 V, \tag{6.7}$$

depending only on z. γ is known as the propagation constant and can be further decomposed to $\gamma = \alpha + ik$, where α is the attenuation constant and k is the wave number. Using this decomposition, we can now construct the complete solution to the wave equation

$$V(z) = V_0^+ \, e^{-\gamma z} + V_0^- \, e^{\gamma z}$$
$$V(z,t) = V_0^+ \, e^{-\alpha z} \, e^{i(\omega t - kz)} + V_0^- \, e^{-\alpha z} \, e^{i(\omega t + kz)}, \tag{6.8}$$

where V_0^+ and V_0^- are wave amplitudes. Taking a closer look at the dependence on z, we see that the first part is representing a wave moving in the forward direction (positive z direction) and the second part is a wave moving in the backward direction, a reflected wave. Again, we get an identical solution for current I.

According to the definition, the characteristic impedance is defined as

$$Z_0 = \frac{V_0^+}{I_0^+}$$

and can be calculated directly by taking the uncoupled differential equation (6.4) and using only the forward wave from equation (6.8) on it

$$\frac{\partial V}{\partial z} = -R\,I_0^+ - i\omega L\,I_0^+ = -\gamma V_0^+,$$
$$Z_0 = \frac{R + i\omega L}{\gamma} = \sqrt{\frac{R + i\omega L}{G + i\omega C}}. \tag{6.9}$$

The characteristic impedance Z_0 is one of the properties of a transmission cable and is important when connecting any cable to electronic devices.

6.2.3 Lossless coaxial cable

A lossless coaxial cable is the ideal case for signal transmission, where R and G are zero. In other words, the central conductor has no resistance and the dielectric no leakage current. Using this information in equations we already have, we see that equation (6.7) becomes

$$\frac{\partial^2 V}{\partial z^2} = -\omega^2 L\,C\,V = -k^2 V, \tag{6.10}$$

wave equation solution (6.8) is

$$V(z,t) = V_0^+ \, e^{i(\omega t - kz)} + V_0^- \, e^{i(\omega t + kz)} \tag{6.11}$$

and the characteristic impedance

$$Z_0 = \sqrt{\frac{L}{C}} \tag{6.12}$$

is completely independent of cable length. In a real case, the values of R and G are actually very small and coaxial cables are arranged by their impedances, with most used $50\,\Omega$ cables for fast signals, $93\,\Omega$ cables for slow signals and $75\,\Omega$ cables for high voltage supply. Another property we can assign to a cable is the velocity of propagation $v = \omega/k$ or its inverse, called the time delay T,

$$T = v^{-1} = \frac{k}{\omega} = \frac{\omega\sqrt{L\,C}}{\omega} = \sqrt{L\,C}, \tag{6.13}$$

measured in time per length of cable. A typical time delay for an RG–58 coaxial cable is on the order of $\sim 5\,\mathrm{ns/m}$. Figure 26 shows the difference, when comparing cables of different delay time. In the figure, the expected time delay of cables, with $16\,\mathrm{ns}$ and $128\,\mathrm{ns}$ separate delays, is $112\,\mathrm{ns}$. The measured value $110.8\,\mathrm{ns}$ is
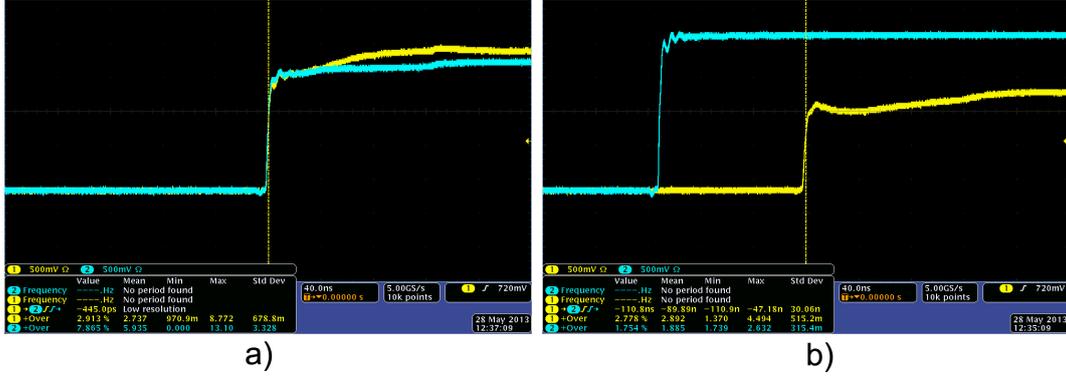

Figure 26: a) Comparison of two cables with the same time delay of $64\,\mathrm{ns}$ and b) comparison of a cable with time delay of $16\,\mathrm{ns}$ (blue) to one with $128\,\mathrm{ns}$ (yellow). The input signal is a step function with frequency $\sim 820\,\mathrm{Hz}$. Measured time delay is shown on the third line in the bottom left.

thus a fairly good measure of difference in delays. As a side note, the difference in amplitudes in this case is not dependent on time delays, but different positioning and length of cables.

6.2.4   Reflections and impedance matching

As we have seen from the solution to the wave equation (6.8), a signal travelling along a coaxial cable consists of a forward and a backward wave. The two opposite waves can give us some unwanted interference or distortion of the original signal. If the cable is terminated with an impedance

$$R = \frac{V_0^+ + V_0^-}{I_0^+ + I_0^-}, \tag{6.14}$$

the impedance of some electronic device, then we can define a reflection constant for a coaxial cable

$$\rho = \frac{V_0^-}{V_0^+} = -\frac{I_0^-}{I_0^+} = \frac{R - Z_0}{R + Z_0}, \tag{6.15}$$

where V_0^+ and I_0^+ are forward directed, while V_0^- and I_0^- are backward directed wave amplitudes. When $R > Z$ we get reflections of positive polarity and $R < Z$ means

27

we have reflections of negative polarity. However, the most interesting case is by far $R = Z$, as there are absolutely no reflections ($\rho = 0$). Cable termination is therefore very important to reduce the effect of reflections and we should make sure to match impedances for two cables or a cable–device interface. If, for instance, the first cable has impedance Z_1 and the second one has impedance Z_2, we can encounter two cases. In the first scenario, where $Z_1 < Z_2$, we need to effectively lower the impedance of the second cable by adding a resistor in parallel with resistance R

$$\frac{1}{Z_1} = \frac{1}{R} + \frac{1}{Z_2} \qquad \longrightarrow \qquad R = \frac{Z_1 Z_2}{Z_2 - Z_1}. \tag{6.16}$$

For the second scenario, $Z_1 > Z_2$, the impedance of the second cable has to be increased with a resistor in series

$$Z_1 = R + Z_2 \qquad \longrightarrow \qquad R = Z_1 - Z_2. \tag{6.17}$$

For measuring equipment, this is done by using a $50\,\Omega$ terminator that adds a resistor in parallel.

6.3 Preamplification stage

If we were to use a long cable for a signal coming from the detector, we would soon observe that the noise level is too great to effectively measure any kind of signal. To remedy such a distortion of the signal, we need to amplify it as soon as possible. This will give us a favourable signal–to–noise ratio and make further signal amplification and shaping that much easier. This stage is known as the preamplification stage and a device called a preamplifier, with a differential input, amplifies the signal while keeping the noise at relatively the same level or reducing it. There are three basic types of preamplifiers: voltage sensitive, current sensitive and charge–sensitive, but for the purpose of semiconductor detectors, we limit ourselves to charge–sensitive amplifiers. The reason for this is that it is independent on the detector capacitance C_d, which is variable for semiconductor detectors. Figure 27 shows the circuit schematic of a charge–sensitive preamplifier. At the heart of a
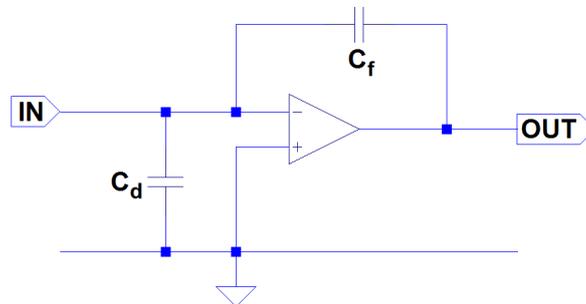
Figure 27: The schematic of a charge–sensitive preamplifier consisting of an opamp with gain A (signal goes through the inverting input) and feedback capacitor C_f. C_d is the capacitance of the detector.

preamplifier lies an operational amplifier (opamp), with gain $A = V_{\text{out}}/V_{\text{in}}$ and two

inputs, an inverting one $(-)$ and a non–inverting one $(+)$, where the inverting one obviously inverts the signal. The feedback capacitor C_f is used to integrate the charge from input Q_{in} and transforming it into output voltage V_{out}. The voltage drop V_f and charge Q_f collected on capacitor C_f are calculated as

$$V_f = V_{\text{in}} - V_{\text{out}} = V_{\text{in}} - (-V_{\text{in}}A) = V_{\text{in}}(1 + A), \tag{6.18}$$

$$Q_f = C_f V_f = C_f V_{\text{in}}(1 + A). \tag{6.19}$$

Since opamps use a differential comparison of the two inputs, there is virtually no current flowing into the negative input (for an ideal opamp, the current is in fact 0). That means that the charge collected on capacitor C_f is equal to the input charge Q_{in} and gives us a virtual capacitance of the preamplifier

$$C_{\text{in}} = \frac{Q_{\text{in}}}{V_{\text{in}}} = C_f(1 + A). \tag{6.20}$$

We can now finally determine the gain G of our preamplifier

$$G = \frac{V_{\text{out}}}{Q_{\text{in}}} = -\frac{A}{C_{\text{in}}} = -\frac{A}{1+A}\frac{1}{C_f} \simeq -\frac{1}{C_f}, \tag{6.21}$$

where the gain of the opamp A is taken as a large number, a very reasonable simplification, since it is typically $A \sim 10^3$ or larger. Obviously, some of the current will flow through C_d (detector capacitance), but with the careful choice of C_f, we can neglect this effect

$$\frac{Q_{\text{in}}}{Q_{\text{in}} + Q_d} = \frac{C_{\text{in}}V_{\text{in}}}{Q_{\text{in}}(C_{\text{in}} + C_d)} = \frac{1}{1 + \frac{C_d}{C_{\text{in}}}}. \tag{6.22}$$

Therefore, if we select a capacitor C_f such that C_{in} from equation (6.20) is much bigger than C_d, we will lose a very small fraction of the signal. As an example, a silicon detector with area $20\,\text{cm}^2$ and thickness $300\,\mu\text{m}$ will have capacitance around $C_d = 700\,\text{pF}$. Assuming the opamp has gain $A = 10^3$ and using a capacitor $C_f = 20\,\text{pF}$, we will only lose about 3% of the total charge on the detector

$$\frac{Q_{\text{in}}}{Q_{\text{in}} + Q_d} = \frac{1}{1 + \frac{700\,\text{pF}}{20\,\text{pF}\cdot(1+10^3)}} \simeq 0.966.$$

For a quicker discharge of capacitor C_f, we can add another resistor in parallel to it, which helps if we are expecting a faster signal.

6.4 Main amplification stage

With our signal properly preamplified (with a high signal–to–noise ratio), we can then connect the output to the main amplifier. The transmission cables can now be much longer, since added noise is not as problematic as before. The main amplifier differs from the preamplifier in that it not only amplifies the signal, but its main job is to also shape the signal, using many sub–components, such as CR–RC pulse shaping, pole–zero cancellation, counters, analog–to–digital converters... For this work, I will be giving a quick explanation of the first three.

6.4.1 CR–RC pulse shaping

The CR–RC pulse shaping is in fact nothing more than a high–pass filter (the CR part), followed by a low–pass filter (the RC part). Putting both together results in a filter that improves the signal–to–noise ratio and attenuates any unwanted high–frequency and low–frequency noise, producing a smoother signal shape. A typical CR–RC circuit is displayed on figure 28 with an added frequency response curve for such a case. Observing the frequency response curve, we see that the CR part
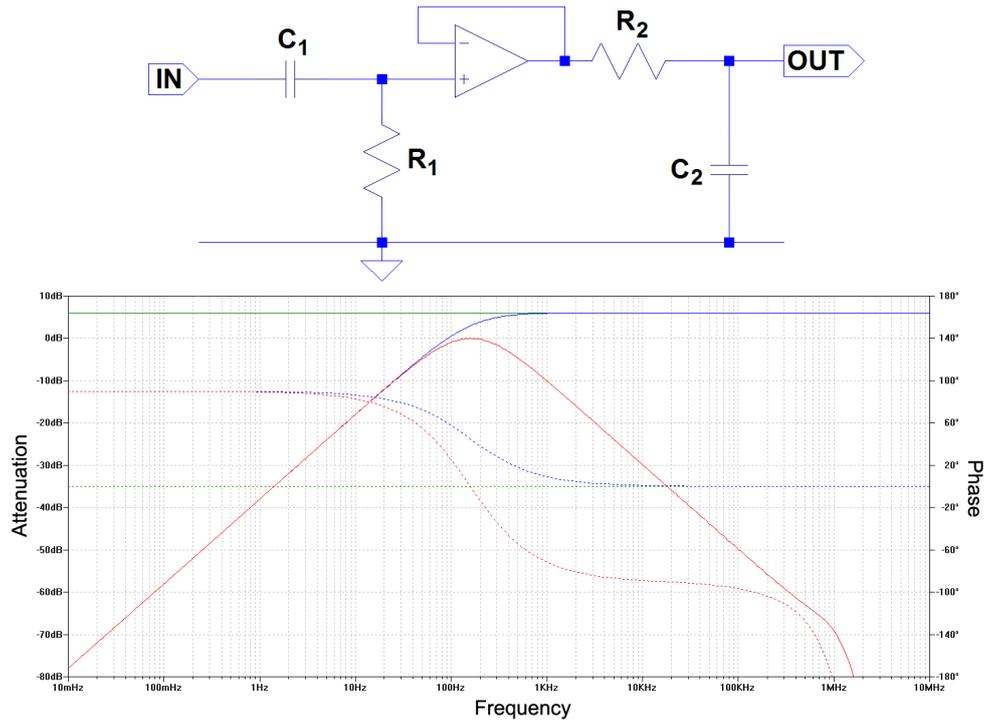

Figure 28: An electronic circuit of a typical CR–RC pulse shaping filter and a frequency response curve for input voltage (green), after CR filter (blue) and after RC filter (red). Solid line represents attenuation/amplification, dashed represents phase.

really is a high–pass filter (blue line), while the red line represents the RC part or low–pass filter. In addition to the CR–RC pulse shaping, we can add another differentiating CR shaping stage even further filtering the outcome.

6.4.2 Pole–zero cancellation

A side–effect that arises when using CR–RC pulse shaping is a noticeable undershoot of the output pulse. This can create an amplitude defect, where the amplitude of each subsequent pulse (if they are all equal) starts trending towards a lower value. Keep in mind that the actual pulse height is not getting smaller, it's just the baseline that is getting lower. To solve this problem, we should add a resistor in parallel to the capacitor in the CR part of the filter from figure 28. This method is called pole–zero cancellation and most amplifiers in use already have it in the form of a potentiometer on the front panel, through which we can restore the baseline after

a pulse. Figure 29 shows the effect of pole–zero cancellation an a trending signal. However, we have to be careful when applying pole–zero cancellation, because we
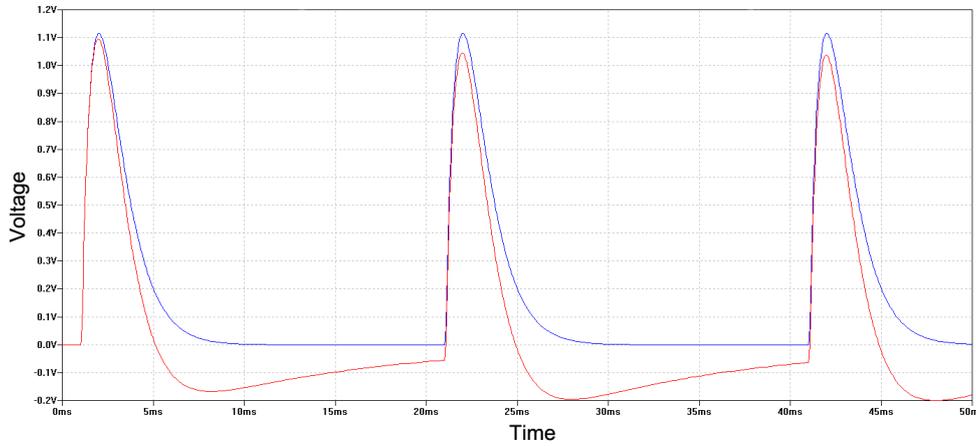

Figure 29: Comparison between a pulse with applied pole–zero cancellation (blue) and one without (red). Note the change in amplitude.

can easily overcompensate by adding a resistor with a too small resistance. In that case the pulse won't have enough time to settle to the baseline in time for the next pulse.

6.4.3 Counter or trigger

Triggers are very widely used in applications where we would like to determine, if a signal was produced with a high enough voltage, for instance, a particle passing through the detector. Their basic structure is that of a discriminator, displayed in figure 30. We create it by running the signal to the non–inverting input and
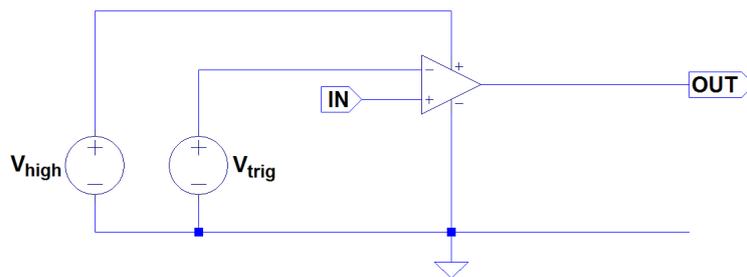

Figure 30: A schematic of a trigger, based on a discriminator. V_{trig} is the triggering voltage level and V_{high} the output trigger voltage.

the trigger voltage level to the inverting input. The trigger level determines the minimum voltage our signal has to have to trigger the counter. Below this level, there is no signal at the counter output. The two supply voltages to the opamp, called the voltage rails, set the maximum and minimum voltage of the counter output. For the case on figure 30, the high voltage rail is V_{high}, while the low voltage rail is grounded (at $0\,\text{V}$).

References

[1] http://www.marysrosaries.com/collaboration/images/c/cf/ Radioactive_Decay_(PSF).png, (23.7.2013).

[2] http://cdn.zmescience.com/wp-content/uploads/2011/07/ fission-1.jpg, (23.7.2013).

[3] http://www.atlas.ch/photos/atlas_photos/selected-photos/lhc/ 9906026_01_layout_sch.jpg, (23.7.2013).

[4] https://upload.wikimedia.org/wikipedia/commons/7/74/LHC.svg, (23.7..2013).

[5] http://www.hep.phy.cam.ac.uk/outreach/CROCS/images/ AugerPrinciple.png, (23.7.2013).

[6] http://www.answers.com/topic/energy-bands, (3.6.2013).

[7] Gerhard Lutz, *Semiconductor and radiation detectors*, 2nd printing, Springer (2007).

[8] William R. Leo, *Techniques for Nuclear and Particle Physics Experiments*, 2nd revised edition, Springer (1994).

[9] http://www.canberra.com/products/detectors/ pips-detectors-standard.asp, (16.7.2013).

[10] H.–C. Schultz–Coulon lecture notes, http://www.kip.uni-heidelberg.de/ ~coulon/Lectures/Detectors/, (16.7.2013).

[11] I. Duerdoth, *Track fitting and resolution with digital detectors*, Nucl. Instr. and Meth. in Physics research, **203** (1982), 291 — 297.

[12] B. List, *The H1 Silicon tracker*, http://hepwww.rl.ac.uk/Vertex03/Talks/ BennoList.pdf, (19.7.2013).

[13] http://www-h1.desy.de/h1det/tracker/sitracker/cst/, (19.7.2013).

[14] ATLAS, http://atlas.ch/inner_detector1.html, (19.7.2013).

[15] R. Bellwied et al. (STAR collaboration), *The STAR silicon vertex tracker: A large area silicon drift detector*, Nucl. Instr. Meth. A, **499** (2003), 640 — 651.

[16] http://2.imimg.com/data2/QO/LL/IMFCP-2825512/rg-ii-co-axial- cables-big-500x500.jpg, (21.7.2013).

[17] H. Spieler, *Solid state detectors and electronics – Electronics I*, http://www-physics.lbl.gov/~spieler/TSI-2007/PDF/ Electronics_I.pdf, (22.7.2013).