Georg Viehhauser

Particle detectors

C4 Lecture notes

Contents

Introduction 1
Properties to be measured 1
Interactions of particles with matter 2
Charged particle energy loss 2
Cherenkov radiation 4
Bremsstrahlung 6
Multiple scattering 6
Hadronic interactions 7
Interaction of photons with matter 8
Interaction of neutrinos with matter 10
Response to ionization 11
Electrical signal generation 11
Internal amplification 12
Gaseous detectors 13
Charge drift 15
Semiconductor detectors 18
Response to excitations 23
Organic scintillators 23
Inorganic scintillators 24
Emulsions, cloud and bubble chambers 24
Track reconstruction 26
Momentum measurement in magnetic fields 26
Calorimetry 28
Shower dimensions 28
Energy resolution 29
Homogenous and sampling calorimeters 30
Hadron showers 31
Particle flow analysis 33

Introduction

Particle detectors are devices which detect subatomic particles. The size of the detector can range from 100s of μm to 100s of km, although typically they will be at the lower end of this range, up to a few metres. This has two important consequences: First, particle detectors are significantly larger than the Compton wavelengths (h/mc) of the particles they need to detect. It is therefore entirely justified to designate them as 'particle' detectors, and it is usually sufficient for the understanding of the detector to treat the particles as quasi-classical particles.

Detector physics is therefore the 'true' particle physics, whereas what is usually called particle physics should be called 'experimental quantum field theory'...

The second important consequence is that most of the subatomic particles cannot be directly detected, because they just don't live long enough to leave a trace that could be detected. Even with a significant boost, only particles with a lifetime of about 10^{-13} s can be detected directly.

Direct detection here means observation of a signal generated in the detector by the particle itself.

	in standard experiments	in high granularity detectors
Leptons	e, μ, ν_e, ν_μ, ν_τ	τ
Mesons	$\pi^\pm$, $K^\pm$, K_L	K_S
Baryons	p, n	Λ^0, $\Sigma^\pm$, $\Xi^{0,-}$, Ω^-
Bosons	γ	

Table 1: Directly detectable sub-atomic particles (plus their anti-particles, where applicable). High granularity detectors are detectors like emulsions or bubble chambers.

Other directly detectable sub-atomic particles are bound states of the two most stable baryons, the proton and the neutron (for example deuterons or αs, but also other nuclei). Finally, some mesons containing c and b quarks could potentially be directly observed in high granularity detectors, although I am not aware of experiments doing this systematically.

All the other subatomic particles can only be observed indirectly, through their decay products, and their existence can only be inferred from a resonance in the invariant mass reconstructed from the four-momenta of their daughter particles.

The fundamental particles that are most obviously directly unobservable are quarks and gluons, because of confinement. However, the hadronization of these particles produces jets. Because of their abundance in collider experiments (in particular at hadron colliders), detectors are needed to measure their kinematic properties. This is usually done for the jet as a whole.

Jets are bundles of closely spaced parallel particles resulting from the hadronization.

Properties to be measured

There are cases when the only information required is evidence for the existence of particles, like for example when the amount of radiation harmful to life in an area needs to be measured (the classical example being a Geiger-Müller particle counter). Then there are cases where the location of incoming particles needs to be observed (for example photons in digital imaging). However, in particle physics it is usually the complete 4-momentum of individual particles that needs to be determined, because of the need of full kinematic reconstruction. This can be done by measuring any combination of its 3-momentum (magnitude

and direction), energy, and mass. The latter can be inferred if one identifies the type of particle, either by observing a process or experimental signature only possible for that particle type, or by other means.

Interactions of particles with matter

The biggest challenge for the detection of subatomic particles is obviously the small amount of energy these particles carry. A T*e*V particle has about 10^{-7} J of energy, far below what can be detected directly as a thermal energy, and the momentum is about 10^{-15} kgm/s, so that direct mechanical measurements are not possible.

However, calorimetric particle detectors have been used in the early days of particle physics (but not for the detection of individual particles). See for example the confirmation of the continuous beta decay spectrum by Meitner & Orthmann (W.Orthmann, Zeitschrift für Physik 60 (1930) 137).

However, subatomic particles do carry the charges of the electroweak or the strong interaction, and with these charges they can couple to similar charges in the detector material and transfer some or all of their energy to the detector.

Charged particle energy loss

For charged particles the primary means of interaction with the detector material is by electromagnetic interaction. Even though the detector material is on the whole electrically neutral, it provides plenty of electrical charges at the subatomic level. Of those, the most relevant for particle detectors are electrons, which are relatively mobile and thus will readily respond to the stimulus of the passing particle.

It is instructive to look at this first using a classical approach for a very massive incoming particle, originally developed by Bohr. We start with modelling the interaction of the incoming particle of mass M and charge ze with a single electron of mass m_e, where we make two assumptions: First, that the mass of the incoming particle is high so that its motion is not disturbed by the collision, and second, that the electron is free to move and changes in its electrostatic potential energy are negligible. The electron will then get a radial kick

N.Bohr, Philos. Mag. 25, 10 (1913).

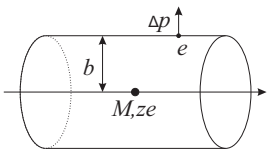


Figure 1: Geometry used in the derivation.

$$\Delta p = \int F \,\mathrm{d}t = e \int E_\perp \frac{\mathrm{d}t}{\mathrm{d}x}\,\mathrm{d}x = \frac{ze^2}{2\pi\epsilon_0 b v},$$

where we have used Gauss' law $\int E_\perp 2\pi b \,\mathrm{d}x = ze/\epsilon_0$. The energy gained by an electron at radius b is then

$$\Delta E(b) = \frac{\Delta p^2}{2m_e} = \frac{z^2 e^4}{8\pi^2 \epsilon_0^2 m_e v^2 b^2}.$$

If the particle passes through a slab of matter of thickness $\mathrm{d}x$, with an electron density n_e, the energy loss will be

$$\mathrm{d}E = -\Delta E(b) n_e \,\mathrm{d}V = -\frac{z^2 e^4 n_e}{4\pi\epsilon_0^2 m_e v^2} \frac{\mathrm{d}b}{b}\,\mathrm{d}x.$$

The energy loss per path length (also called the 'stopping power') is then obtained from an integration over the distance b,

$$\frac{\mathrm{d}E}{\mathrm{d}x} = -\frac{z^2 e^4 n_e}{4\pi\epsilon_0^2 m_e v^2} \ln \frac{b_{\max}}{b_{\min}},$$

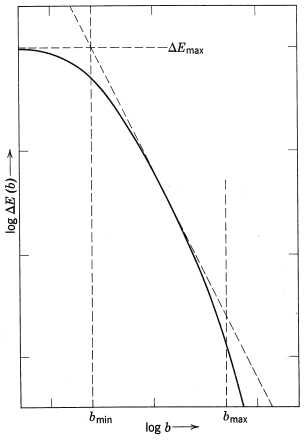


Figure 2: Energy transfer as a function of impact parameter (from Jackson).

where we have introduced physical integration boundaries $b_{\min}$ and $b_{\max}$. These are given by the limits where our assumptions break down. The lower boundary $b_{\min}$ reflects the fact that the maximum energy transfer is limited and occurs in a head-on collision. In that case (for $M \gg m_e$) $\Delta E_{\max} = 2\gamma^2 m_e v^2 = \Delta E(b_{\min})$. The upper boundary is given by the distance beyond which the interaction loses its ability to affect the electronic state. To estimate this, we assume that the electron is bound in an atomic orbit. If the interaction time, which can be estimated as $t(b) = b/(\gamma v)$, becomes longer than the orbital period ω^{-1} of the electron, the influence on the electronic state will be adiabatic, and no energy will be transferred.

The exact expression is

$$\Delta E_{\max} = \frac{2\gamma^2 m_e v^2}{1 + 2\gamma m_e/M + (m_e/M)^2}.$$

The difference to the approximate expression becomes more relevant at higher energies, and results typically in a difference of a few % for $\mathrm{d}E/\mathrm{d}x$ at 100 G*e*V.

The combined result of the classical derivation is thus

$$\frac{\mathrm{d}E}{\mathrm{d}x} = -\frac{z^2 e^4 n_e}{4\pi\epsilon_0^2 m_e v^2} \ln \frac{4\pi\epsilon_0 \gamma^2 m_e v^3}{z e^2 \omega}.$$

There are several things to note:

- The energy loss is independent of the mass of the incoming particle;
- it is proportional to the electron density of the detector material, otherwise it is only weakly depending on the detector material (through ω);
- as it originates in the interaction with shell electrons in the material, it is proportional to $z^2 e^4$;
- for low velocities ($\beta\gamma \lesssim 1$) it is dominated by the falling $1/v^2$ function, whereas for large velocities it will increase slowly due to the logarithmic part.

The rising section is referred to as the 'relativistic rise'.

More refined derivations of the charged particle collision energy loss consider the effects of the dielectric properties of the detector material on the electromagnetic field accompanying the incoming particle. A commonly used form to describe this energy loss is

$\frac{1}{\rho}\frac{\mathrm{d}E}{\mathrm{d}x}$ is called the 'mass stopping power'.

$$\frac{1}{\rho}\frac{\mathrm{d}E}{\mathrm{d}x} = -K z^2 \frac{Z}{A}\frac{1}{\beta^2}\left(\frac{1}{2}\ln\frac{2 m_e c^2 \beta^2\gamma^2 \Delta E_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2}\right),$$

This equation is the 'Bethe equation' after Hans Bethe, who was the first to develop the quantum mechanical treatment of the problem. Sometimes it is also referred to as the 'Bethe-Bloch equation'. However, Bloch's approximation of the mean excitation energy is not used any more in modern parametrizations.

where the constant $K = 4\pi N_A \alpha^2 \hbar^2/m_e \simeq 0.3\ \mathrm{M}e\mathrm{V mol}^{-1}\mathrm{cm}^2$, I is the mean excitation energy and $\delta(\beta\gamma)$ a correction due to the density effect. The latter describes the reduction of the relativistic rise due to polarisation of the detector material. Another modification of the relativistic rise in real detector geometries takes place when the maximum energy transfer is limited to a value $E_{\mathrm{cut}} < E_{\max}$. In that case the rise is stopped, and the energy loss becomes constant. In practice, this occurs in thin detectors, because of the low probability of high energy transfer interactions.

This is called the 'Fermi Plateau'.

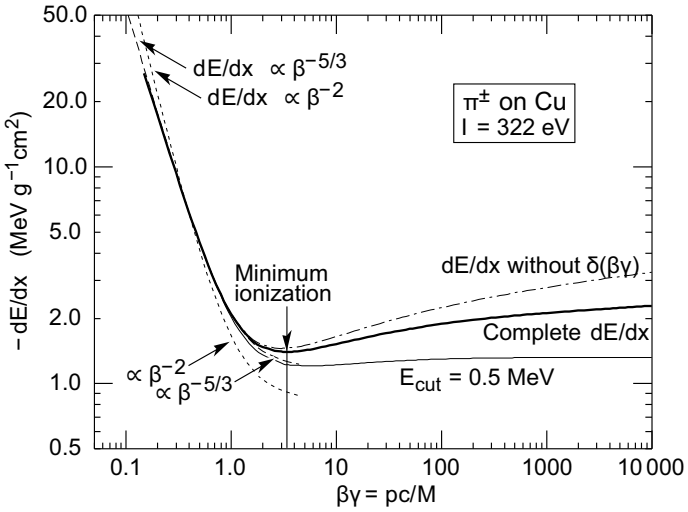


Figure 3: Mass stopping power as a function of $\beta\gamma$ (modified from PDG, Phys. Rev. D50, 1173 (1994)).

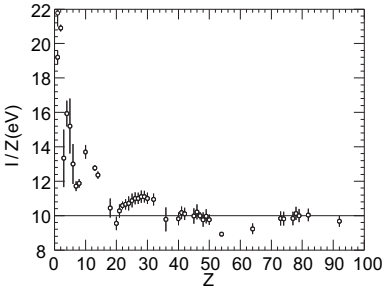


Figure 4: Mean excitation energies divided by Z (modified from PDG, Phys. Rev. D50, 1173 (1994)).

An important feature of this curve is the minimum, which occurs at $\beta\gamma \simeq 3-4$ with a mass stopping power of about 1-2 MeVg^{-1}cm^2. We call a particle with such a velocity a 'minimum ionizing particle'.

The calculations discussed so far average over many interactions with electrons in the detector material. In reality, the interactions involve individual particles in different kinematic conditions, so that ionization events along the track are discrete, and the amount of energy transferred fluctuates between different ionization events. Individual events can comprise several ionizations and/or excitations of atoms. In the extreme case, the interaction can result in the transfer of a sizeable amount of energy to a single electron, which then travels for macroscopic distances in the detector material. Such an electron is called a 'δ-electron'.

As a consequence of these statistical fluctuations the energy deposited in a finite slab of detector material varies from event to event. In a thick layer the distribution of the deposited energy will acquire a Gaussian shape, in accordance with the central limit theorem. For thin layers the distribution is skewed with a long tail due to rare high-energy clusters. Such a distribution is called a 'Landau distribution'. The discrete nature of the ionization loss and the associated cluster size distribution is particularly observable in gaseous detectors, where the lower interaction density along the track leads to a wide separation of the clusters.

To bring a particle to a complete stop typically significant amounts of material are required, so that the relative fluctuations of the energy loss become smaller. Consequently, the range of charged particles of a given energy in a material is a well-defined quantity. Furthermore, because of the increased energy loss at low velocities, the energy deposition is highest shortly before the end point of the particle track. This profile for the deposited energy is exploited in particle beam therapy.

Cherenkov radiation

The density effect introduced in the previous section can be calculated looking at the electromagnetic fields accompanying the incoming particle. The calculation shows the strong attenuation of the field in the dielectric. In QFT the interaction is described by the exchange of virtual

These ionization events are also called 'clusters'.

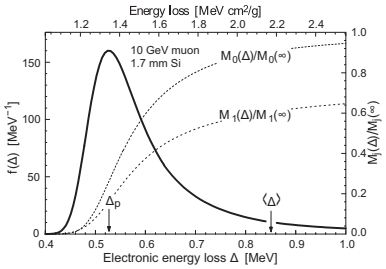


Figure 5: Landau distribution of energy loss. Δ_p is the most probable energy loss. $\langle\Delta\rangle$ is the energy loss given by the Bethe equation. M_0 and M_1 are the cumulative 0^{th} moment (mean number of collisions) and 1^{st} moment (mean energy loss) (modified from PDG, Prog. Theor. Exp. Phys. 2020, 083C01 (2020)).

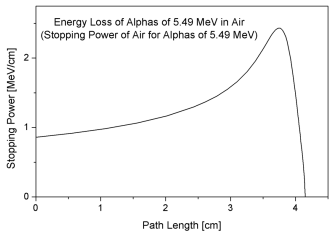


Figure 6: Energy deposition as a function of penetration depth. Such a curve is called a 'Bragg curve', and the region of maximum energy deposition close to the end point the 'Bragg peak' (from Wikipedia).

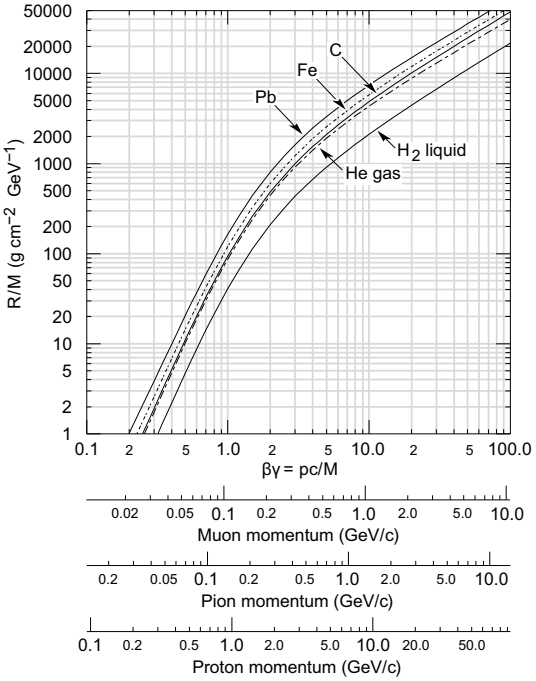


Figure 7: Range of heavy charged particles in different materials (modified from PDG, Phys. Rev. D50, 1173 (1994)).

photons with a limited range. There is, however, a condition under which real photons are emitted, which can travel macroscopic distances. This type of radiation is referred to as 'Cherenkov radiation'.

A simple picture of this effect is a luminar boom. The particle is seen as the source of spherical electromagnetic waves, which expand with the speed of light in the medium, c/n. If the particle moves faster than this, the spherical waves will build up to a conical wavefront with an opening angle given by $\cos\theta_C = (c/n)/v = (n\beta)^{-1}$. It should be obvious that for a given material with index of refraction n, Cherenkov radiation will only be emitted if the particle's velocity is above a threshold $\beta_{\min} = 1/n$.

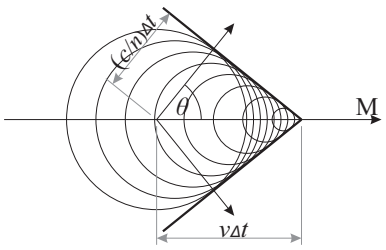


Figure 8: Cherenkov cone.

The energy lost to Cherenkov radiation is given by the Frank-Tamm equation

$$\frac{dE}{dx} = \frac{(ze)^2}{4\pi\epsilon_0 c^2}\int_{\beta>n(\omega)^{-1}}\left(1-\frac{1}{\beta^2 n(\omega)^2}\right)\omega\, d\omega, \tag{1}$$

The integral extends over the energies where photons are observed, i.e. where Cherenkov radiation is emitted, but also where the photon detector is sensitive.

where z is the charge of the incoming particle. The energy density is to good approximation proportional to the frequency, as long as the index of refraction does not change significantly. This is the reason why Cherenkov radiation from nuclear reactors immersed in water appears blue. The integral stays finite, as the index of refraction approaches 1 for $\omega \to \infty$.

The number of photons $N = E/\hbar\omega$ is then

$$\begin{aligned}\frac{dN}{dx} &= \frac{\alpha z^2}{\hbar c}\int\left(1-\frac{1}{\beta^2 n(E_\gamma)^2}\right)dE_\gamma,\\ &\simeq 370 z^2\int \sin^2\theta_C(E_\gamma)\, dE_\gamma\, eV^{-1}cm^{-1},\end{aligned} \tag{2}$$

where E_γ is the photon energy.

The amount of energy lost to Cherenkov radiation is only a small fraction of the charged particle energy loss (between 10^{-2} and 10^{-3}), but it is the long range of the Cherenkov photons, and the correlation of their

direction with the direction and speed of the incoming particle that leads to this effect being used in a whole range of particle detectors.

Bremsstrahlung

The energy loss for electrons (positrons) has in principle a similar speed dependency as the Bethe equation, despite the fact that some of the assumptions we used to justify that equation do not apply. There is, however, one big difference for electrons interacting electromagnetically with matter: down to very low energies the dominant energy loss mechanism is bremsstrahlung. To conserve 4-momentum, Bremsstrahlung requires the electromagnetic field of a charged particle, typically a nucleus. The calculation of the cross-section for this process is fairly intricate, as it has to take into account screening of the nuclear charge by the shell electrons. However, for large energies the energy loss per track length is proportional to the energy of the incoming particle

$$\frac{\mathrm{d}E}{\mathrm{d}x} = -\frac{E}{X_0},$$

where we have introduced the proportionality constant X_0. For a given type of particle, this constant depends on the detector material, and is called the 'radiation length' of the material. It indicates the length at which the energy of the incoming particle has dropped by a factor $1/e$. The detailed calculation yields

$$\frac{1}{\rho}\frac{1}{X_0} = 4\frac{\hbar^2\alpha^3}{m_e^2c^2}\frac{N_A}{A}\left\{Z^2\left[L_{\mathrm{rad}} - f(Z)\right] + ZL'_{\mathrm{rad}}\right\},$$

where L_{rad} and L'_{rad} can be approximated by $\ln(184.15Z^{-1/3})$ and $\ln(1194Z^{-2/3})$, respectively, for elements with $Z > 4$. These are the contributions from bremsstrahlung in the field of the nuclei ($\propto Z^2$), and the electrons ($\propto Z$). $f(Z)$ is a small correction $\mathcal{O}(\alpha Z^2)$. A simple approximation for X_0 in g/cm^2 is $X_0 \sim 180A/Z^2$ (good to 20% for 13≤Z≤92).

As the radiative losses increase linearly with the energy of the incoming particle, they will dominate at high energies. The energy at which the radiative energy loss exceeds the loss to ionization is called the critical energy E_c. It can be parametrized as 710 MeV/$(Z + 0.92)$ (for gases) or 610 MeV/$(Z + 1.24)$ (for solids) to within a few %.

The bremsstrahlung radiation spectrum is relatively uniform in energy, and limited at low energy (large impact parameter) by shielding of nuclear charge by atomic electrons, and at high energy by the maximum incident particle energy. In the frame of the electron the radiation is classically a dipole radiation, with the maxima perpendicular to the deceleration. Lorentz transformation into the lab frame gives a narrow, but not sharp cone of blue-shifted radiation with a cone angle of $\theta = 1/\gamma$.

Multiple scattering

So far we have discussed the energy loss of charged particles due to inelastic collisions with the electrons in the detector material, which ultimately generates the signal we require for observation. However,

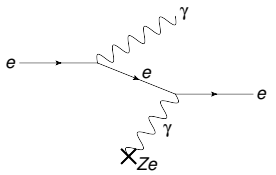


Figure 9: A Feynman diagram for bremsstrahlung.

The rate is proportional to the matrix element squared, where the matrix element $M \propto Ze^2$, differently than the interaction with the shell electrons, which will have a rate $\propto Z|M|^2$, with $M \propto e^2$.

For other particles m_e needs to be replaced by the mass of the incoming particle M. Hence radiation effects for other particles happen on length scales which are M^2/m_e^2 larger than for electrons. For muons bremsstrahlung becomes relevant at the scale of T*e*V.

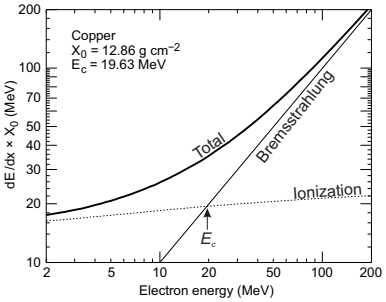


Figure 10: Electron energy loss due to bremsstrahlung compared to energy loss due to ionization (modified from PDG, Phys. Rev. D50, 1173 (1994)).

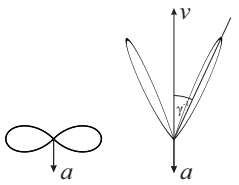


Figure 11: Bremsstrahlung emission direction in the electron frame (left), and in the lab frame (right).

charged particles also interact elastically with the nuclei in the detector. While this does not lead to a detectable energy deposit, it can have significant effects on the particle trajectory, with associated degrading of the knowledge of the original direction of the particle.

Individually, the scattering events can be described by Rutherford scattering, with the usual high prevalence of small scattering angles, and rare large-angle collisions. Multiple ($n > 100$) small scattering angle events will result, in accordance with the central limit theorem, in a Gaussian probability distribution of the overall scattering angle, with non-Gaussian tails at larger scattering angles, due to the rare large-angle scattering events. It is customary to ignore the tails, and focus on the Gaussian core of the distribution.

A widely used description of the width of this distribution for scattering in thin slabs of material by Highland is

$$\theta_0 = \theta_{\text{plane}}^{\text{rms}} = \frac{1}{\sqrt{2}}\theta_{\text{space}}^{\text{rms}} = \frac{13.6\,\text{MeV}}{\beta c p} z \sqrt{\frac{x}{X_0}} \left[1 + 0.088 \log_{10}\left(\frac{z^2}{\beta^2}\frac{x}{X_0}\right)\right], \quad (3)$$

where x is the distance travelled through the material. The appearance of the radiation length X_0 in this expression is coincidental, and not justified by the underlying physics. However, it is an accessible parameter to describe the scattering centre density, which is reasonably correlated.

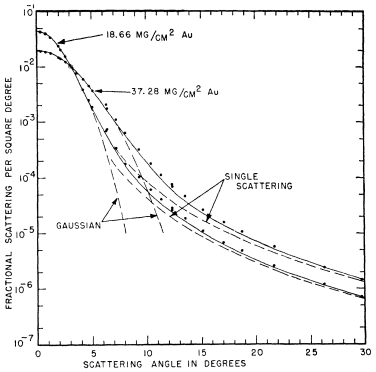


Figure 12: Angular scattering distribution of 15.7 MeV electrons from thick and thin gold foils (A.O.Hanson et al., Phys. Rev. 84 (1951) 634).

$\theta_{\text{space}}^{\text{rms}}$ is the rms of the 3D spatial scattering angle distribution, which for small angles can be approximated by

$$P(\theta)\,\mathrm{d}\Omega = \frac{2\theta}{\left(\theta_{\text{space}}^{\text{rms}}\right)^2} \exp\left(\frac{-\theta^2}{\left(\theta_{\text{space}}^{\text{rms}}\right)^2}\right) \mathrm{d}\theta.$$

$\theta_{\text{plane}}^{\text{rms}}$ is the rms of the angle in a projection of this distribution onto a plane containing the direction of the incoming particle.

Hadronic interactions

Incoming hadrons can also interact with the nuclei in the detector material via the strong interaction. Here the limiting factor is that the strong interaction is mostly contained within the hadrons, and the residual effect of the interaction is limited to a range of about 10^{-13} m. Such interactions thus require a lot of material to take place. Again, there is a characteristic scale length, the 'nuclear interaction length' λ, which describes the distance over which the number of particles in a beam that do not undergo a hadronic interaction is reduced by a factor $1/e$. It is given by

$$\lambda = \frac{A}{N_A \rho \sigma_{\text{total}}}. \quad (4)$$

To obtain a simple approximation it is customary to use the black disk limit $\sigma(hN) = \sigma(hp)A^{2/3}$, and to use the observation that $\sigma(hp)$ is only weakly depending on the projectile and the energy of the incoming particle (and about 37.5 mb), so that a commonly used approximation in the energy range 1 to 1000 GeV is

$$\lambda = \frac{35\,\text{cm}}{\rho} A^{1/3}, \quad (5)$$

with the density ρ in g/cm^3. The error associated with these approximation can be estimated from fig. 13.

For low energies (eV to keV), capture reactions are dominant, with the subsequent emission of a photon (radiative capture), or a very small nucleus fragment (a nucleon, hydrogen isotope, or α), or even fission. All these capture reactions have a cross-section that drops with $1/v$, and possibly a number of narrow resonances.

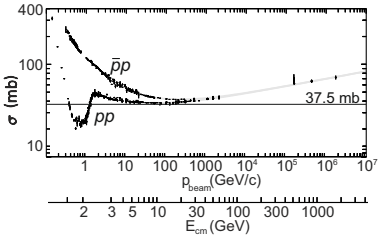


Figure 13: Total cross-section for the scattering of pp and $\overline{p}p$ (modified from PDG, Phys. Rev. D50, 1173 (1994)). Cross-sections for other hadronic projectiles look similar (see https://pdg.lbl.gov/2021/hadronic-xsections/hadron.html).

At higher energies (MeV), there is enough energy to excite the nucleus, with subsequent de-excitation typically via γ emission. However, the principal energy loss mechanism in this energy regime is elastic scattering, with the energy used in the recoil of the hit nucleus.

Finally, hadronic collisions of incoming particles with an energy above the pion mass will result in hadronic secondaries, which will either decay electromagnetically (π^0), lose energy to ionization (charged hadrons), or in further hadronic interactions (charged and neutral hadrons).

Hadronic interactions are primarily important for the detection of neutral hadrons, as they do not directly lose energy to ionization. It is rather the ionization generated by the secondaries created in the hadronic interactions, or the nuclear recoil, which we can use for detection. The absence of ionization loss also accounts for the high penetration capabilities of neutral hadrons. The most important neutral hadron, of course, is the neutron.

The nuclear interaction length and the radiation length are comparable for small nuclei, but become more and more different for larger nuclei.

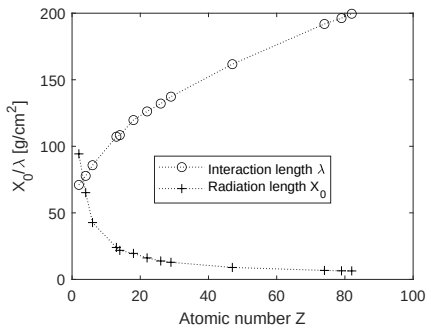


Figure 14: Nuclear interaction length λ and radiation length X_0 for different elements (data from https://pdg.lbl.gov/2019/AtomicNuclearProperties/). To get to a length in metres, the value in the plot needs to be multiplied with the density in g/cm^3.

Interaction of photons with matter

Photons themselves don't carry an electric charge, they therefore also do not lose energy gradually in ionization events via exchange of virtual photons, but they are destroyed in a single interaction event with matter. Depending on the energy, there are three major interaction mechanisms.

We are ignoring several less important interaction types here: Raleigh and Thomson scattering are types of elastic scattering mostly relevant at low energies. Photonuclear reactions take place in the medium energy range (~MeV). All of these have a cross-section at least an order of magnitude less than the processes discussed here, and have no practical use in detectors.

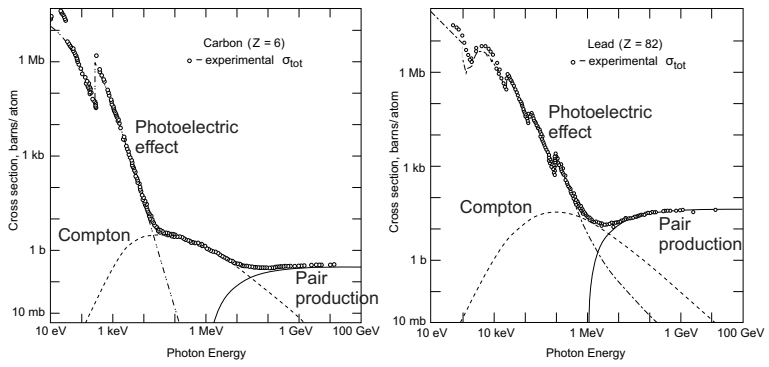


Figure 15: Cross-sections for the interaction of photons for a low-Z and a high-Z element (modified from PDG, Phys. Rev. D50, 1173 (1994)). We have ignored coherent scattering, photonuclear reactions and pair production in the electron fields.

The mean free path is given by λ[cm] = $1.66A/(\rho \times \sigma$[b]), where ρ is the density in g/cm^3.

Photoelectric effect At low energies (up to 1 MeV) the dominant interaction is photoelectric effect. The photon gets absorbed by an atom and its energy frees one of the electrons, which escapes with an energy $E = E_\gamma - E_{\text{bind}}$, where E_{bind} is the binding energy of the electron. At very low energies this equation constitutes a threshold. For a single atom the binding energy equals the ionization energy. For solids, and in particular for metals, the energy required to free an electron is lower than that, and is called the 'work function' of the material. For semiconductors it usually is sufficient to lift the electrons across the band gap between the valence and the conduction band (for Silicon about 1.12 eV). Typically, the threshold for photoelectric effect is a very few eV.

$\lambda[\mu\text{m}] = 1.24/E[e\text{V}]$, so a 1 eV photon has a wavelength of 1.24 μm, in the near infrared. Visible light is in the range from 1.75 to 3.1 eV.

The cross-section for photoelectric effect drops with energy (with decreasing wavelength), because the matrix element for the photon interaction becomes smaller as it is integrated and thus gets averaged over more and more periods of the electric field. For higher energies, the shell structure of the atom becomes visible with an absorption edge for each new energetically accessible shell, with the K shell the last and highest energy shell. Because of the complexity of shell and shielding effects, there is no closed form description of the cross-section for photoelectric effect, but typically it scales with Z^m with m between 3 and 5, and E_γ^{-n}, with n around 3.

The strong dependence on Z is the reason why high-Z materials (e.g. lead) are preferred shielding materials for γ radiation.

Compton scattering In the intermediate energy range (about 1 MeV) Compton scattering is the most prevalent interaction. The cross-section for Compton scattering on a free electron has been one of the first results from QED and is given by the Klein-Nishina equation

$$\sigma^{\text{KN}} = 2\pi\left(\frac{\alpha\hbar c}{m_e c^2}\right)^2\left[\frac{1+x}{x^2}\left(\frac{2(1+x)}{1+2x}-\frac{\ln(1+2x)}{x}\right)+\frac{\ln(1+2x)}{2x}-\frac{1+3x}{(1+2x)^2}\right],$$

with $x = E_\gamma/(m_e c^2)$. For low energies this approaches the cross-section for Thompson scattering, $\sigma^{\text{Th}} = (8\pi/3)\alpha\hbar c/(m_e c^2)^2$.

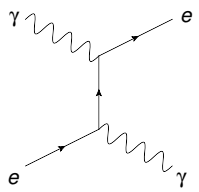


Figure 16: Feynman diagram for Compton scattering.

Thomson scattering is the elastic scattering of photons by a free charged particle, and can be described by classical electromagnetism.

For scattering on atomic electrons the cross-section at high energies ($x > 1$) approaches the Klein-Nishina result, which decreases as

$$\sigma^{\text{KN}} \underset{x\to\infty}{\longrightarrow} \frac{3}{8}\sigma^{\text{Th}}\frac{1}{x}\left(\ln 2x + \frac{1}{2}\right),$$

and the cross-section for Compton scattering on the atom becomes $\sigma = Z\sigma^{\text{KN}}$. For low energies, however, binding effects are important. If the photon does not dislodge the bound electron it will scatter coherently, without energy loss. Thus the cross-section for Compton scattering decreases once the photon energy becomes comparable to, or smaller, than the binding energies of the electrons in the atom.

Ignoring binding effects, the scattered photon energy can be found from 4-momentum conservation, $E'_\gamma/E_\gamma = [1 + x(1-\cos\theta)]^{-1}$, where θ is the scattering angle for the photon. The kinetic energy of the scattered electron is thus

$$E_e = E_\gamma\frac{x(1-\cos\theta)}{1+x(1-\cos\theta)},$$

which has a maximum for head-on collisions of $E_e^{\max} = [2x/(1+2x)]E_\gamma$. The scattered electron typically loses its energy to ionization in the detector until it comes to a stop.

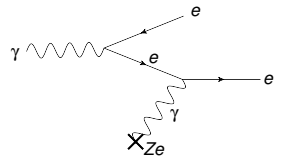


Figure 17: A Feynman diagram for pair production.

Pair production Once the energy of the photon exceeds $2m_e c^2$, it can create an electron/positron pair. To maintain 4-momentum conservation it can only do this in the electric field of a nucleus, which absorbs the recoil. The Feynman diagram for pair production is similar to Bremsstrahlung, and thus the result for the cross-section is similar. For high energies the absorption coefficient tends to the constant value

$$\mu = n\sigma = \frac{7}{9}\frac{1}{X_0},$$

where n is the density of nuclei.

Pair production can also take place in the field of a shell electron, but because of the smaller charge the atomic cross-section for this type of pair production is smaller by a factor Z.

The excess energy above the energy required to create the electron-positron pair is carried away as kinetic energy by the two leptons. They in term lose their energy to bremsstrahlung or ionization, until they come to a stop. At this point the positron will annihilate with an electron in the material, with the emission of two 511 keV photons. Those can be absorbed in the detector by photoelectric effect or Compton scattering, or, if the detector is small, escape the detector.

Decay into two photons is required to conserve 4-momentum.

Interaction of neutrinos with matter

Neutrinos can only interact weakly with matter. They can do that either via charged current or neutral current interactions. In the former case a charged lepton is generated, which can be detected through its electromagnetic energy loss. Typically, this is done using the Cherenkov radiation in a large water tank, which allows for a large detector volume with moderate readout costs, but there are also detectors using emulsions (to detect the very short-lived taus created from ν_τ), and ionization and scintillation detectors. For charged current reactions the neutrino must have enough energy to create the massive lepton, which implies a threshold for the neutrino energy for these type of detectors.

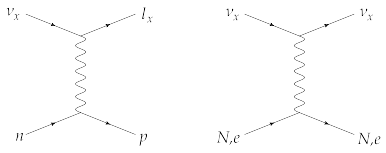


Figure 18: Neutrino charged current (left) and neutral current (right) reactions.

Neutral current reactions do not have that threshold, but the signal in the detector has to be generated by the recoil of the massive fermion on which the neutrino has scattered. This can be an electron, or a nucleus. The recoil in the latter case is typically small, so that nuclear reactions are required to provide a signal.

Response to ionization

In most cases the interaction of subatomic particles with a detector results ultimately in charged particle collisions with the electrons in the detector, which leads to excitation or ionization. We are therefore presented with two typical types of signals, electrically with pairs of ions and electrons (or holes and electrons in a semiconductor), or optically with low energy photons (from IR to UV) from the decay of excited atomic states.

To collect the charges for electrical readout, they need to be movable, with a low probability for capture and recombination. Viable detector materials are therefore non-polar pure gases or liquids, or semiconductors. Gaseous detectors have the advantage of low density, and hence low multiple scattering, whereas liquids and semiconductors create significantly more ionization charge per track length due to their higher density, and the reduced distance between ionization events. Liquids and gases have the advantage that they can be easily replaced, so contamination and radiation damage effects can be reduced, as long as these are not solidifying inside the detectors.

In case of semiconductors this is enhanced by the narrow band gap.

Electrical signal generation

Typically, the ionization charges are separated by an electrical drift field in the detector. When the electrodes generating the drift field are connected, it is the movement of the charges that produces an instantaneous current signal according to the Shockley-Ramo theorem

$$I(t) = q\vec{E}_{\mathrm{w}}\vec{v}(t), \tag{6}$$

where q is the charge of the moving particle, $\vec{v}(t)$ its instantaneous velocity, and $\vec{E}_{\mathrm{w}}$ is called the weighting field. The latter is obtained by grounding all electrodes in the system safe the one on which the induced current is measured, and placing that electrode at unit potential. Note that a current is induced from the moment the ionization charge is created, for as long as the charge is moving, but it ends when the charge reaches the electrode and its movement stops. The current is a result of the change of the density of the electrostatic flux lines on the electrode, caused by the change of the position of the charge.

The connection needs to be for high frequencies only, via a capacitor.

W.Shockley, J. App. Phys. 9 (10): 635–636 (1938) and S.Ramo, Proc. of the IRE. 27 (9): 584–585 (1939)
The Ramo-Shockley theorem is derived under the assumption that magnetic effects are negligible and the electric field propagates instantaneously.

The weighting field depends only on the geometry of the electrodes and determines how the motion of the charge couples to a specific electrode to induce a signal. It has no relation to the actual electrical field. Only in geometries with two electrodes do the weighting field and the electric field have the same shape.

We will illustrate this here with a simple parallel plate capacitor and a single electron-ion pair created in its gap.

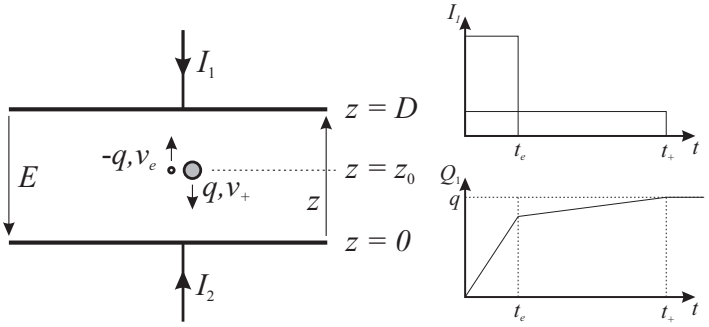


Figure 19: Ionization in a parallel plate capacitor. It is assumed that the capacitor has an area that is large against it's thickness.

For this geometry the electric field is $E = V_0/D$. The weighting field is $\vec{E}_w^1 = -(1/D)\hat{z}$ for the upper electrode, while it is $\vec{E}_w^2 = (1/D)\hat{z}$ for the lower electrode. Typically, the drift velocity for the charges is proportional to the drift field ($v = \mu E$, with the mobility μ) and thus constant in this geometry. Therefore the induced currents on the two electrodes are

In gases the mobility for electrons is typically by about a factor 100 higher than for ions.

$$I_1(t) = \underbrace{(-q)\left(-\frac{1}{D}\right)v_e}_{\text{from } e} + \underbrace{q\left(-\frac{1}{D}\right)(-v_+)}_{\text{from ion}} = \frac{q}{D}(v_e + v_+), \quad \text{and} \quad I_2(t) = -I_1(t).$$

$Q = \int_0^t I\,\mathrm{d}t$, and thus the total charge induced on the upper plate is

$$Q_1^{\text{tot}} = \int_0^\infty I_1\,\mathrm{d}t = \frac{q}{D}(v_e t_e + v_+ t_+) = \frac{q}{D}\left(v_e \frac{D - z_0}{v_e} + v_+ \frac{z_0}{v_+}\right) = q.$$

Internal amplification

Often, the ionization charge produced by collisions in the detector material is too small to be sensed in an electronics readout system amongst the irreducible electronic noise that is present in such a system. This is particularly true for gaseous detectors, where the primary ionization can be as low as a few ten to a hundred electron-ion pairs, but even in semiconductor detectors, where the ionization from a minimum ionizing particle can be a few thousand electron-hole pairs, a primary signal increase can be beneficial for small signals (for example in response to low intensity visible light).

In these cases internal charge multiplication is required that amplifies the charge generated in the detector before collection in the readout electronics. This is achieved by impact ionization in high electric fields in the detector. The primary electrons get accelerated between collisions with atoms inside the detector in the high fields, so that they get enough energy to ionize an atom in the next collision.

A detector, which exploits this on a macroscopic scale is the photomultiplier tube (PMT).

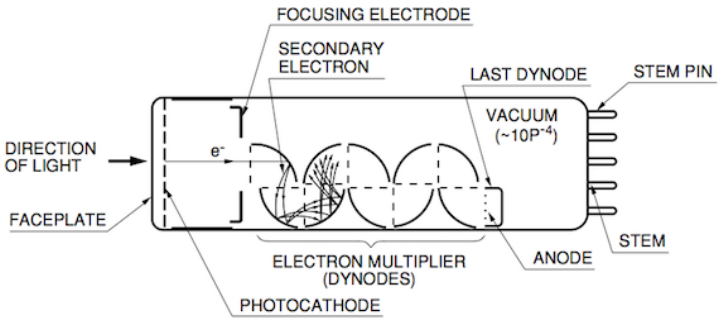


Figure 20: Photomultiplier tube. http://analyticalprofessional.blogspot.co.uk/2013/05/photomultiplier-tube.htm

In this detector, incoming photons generate photoelectrons in a photocathode. The electrons are then guided by electric fields onto a set of high voltage dynodes, which are held at progressively lower potential by a voltage divider network. The electric field between the dynodes accelerates the electrons, so that more electrons are ejected in the impact on the following dynode. To allow for the travel of the electrons and to protect the photocathode and dynode surfaces the photomultiplier is contained

A photocathode is typically a metallic electrode that is covered with an alkali with low work function.

in an evacuated glass tube. Photomultipliers are cheap, yet efficient and fast photon detectors with a large acceptance. They are therefore ideal for the detection of Cherenkov photons in large neutrino experiments, or for the optical readout of large calorimeters. Their main drawbacks are their large size and hence limited position resolution, and that they are strongly affected by magnetic fields. They are therefore often enclosed in a layer of 'μ metal', a soft, ferromagnetic nickel-iron alloy.

Gaseous detectors

Due to the low pressure, the amount of primary ionization in gaseous detectors is small (typically a few 10 electron-ion pairs/cm). Internal amplification is therefore a must. This is achieved by very strong electrical fields (> 100 kV/cm), so that an electron drifting through the gas acquires enough energy between collisions with the gas molecules that it will ionize the next molecule it collides with. Once this condition is met, the number of charge carriers will grow exponentially into an avalanche of electrons.

Gas	Density, mg cm^{-3}	E_x eV	E_I eV	W_I eV	$dE/dx\|_{\text{min}}$ keV cm^{-1}	N_P cm^{-1}	N_T cm^{-1}
He	0.179	19.8	24.6	41.3	0.32	3.5	8
Ne	0.839	16.7	21.6	37	1.45	13	40
Ar	1.66	11.6	15.7	26	2.53	25	97
Xe	5.495	8.4	12.1	22	6.87	41	312
CH_4	0.667	8.8	12.6	30	1.61	28	54
C_2H_6	1.26	8.2	11.5	26	2.91	48	112
iC_4H_{10}	2.49	6.5	10.6	26	5.67	90	220
CO_2	1.84	7.0	13.8	34	3.35	35	100
CF_4	3.78	10.0	16.0	54	6.38	63	120

Figure 21: Properties of typical detector gases at NTP. E_x: first excitation energy, E_I: ionization energy, W_I: average energy per electron/ion pair, $dE/dx|_{\text{min}}$:differential energy loss, N_P number of primary clusters and N_T total number of electron-ion pairs per cm, for unit charge minimum ionizing particles (PDG, Prog. Theor. Exp. Phys. 2020, 083C01 (2020)).

$$\mathrm{d}N = N\alpha\,\mathrm{d}x, \text{ and } N(x) = N_0 e^{\int \alpha(E(x'))\,\mathrm{d}x}, \tag{7}$$

where the proportionality constant α is called the 'Townsend coefficient'. $A = N_{\text{final}}/N_0$ is called the 'gas gain'. As it depends on the field and the mean free path between collisions, it scales with E/p, where E is the field strength, and p the pressure of the gas.

Ionization is not the only possible outcome of the collisions in the avalanche. Sometimes, a hit atom will end up in an excited state, and subsequently de-excite with the emission of a photon. Gaseous detectors are typically optically transparent, and this photon can travel macroscopic distances. At some point, it will interact with the gas or the detector walls, releasing a photoelectron, which in turn will become the start of a new avalanche. When this process repeats, it leads to a continuous discharge in the detector, rendering the detector useless. There are two ways around this issue:

This is called 'photon feedback'.

- The high voltage for the amplification field is supplied via a resistor. When the detector goes into discharge, the increase in current will increase the voltage drop across this resistor, so that the field in the detector drops to levels below what is required for avalanche multiplication. This is the principle of the Geiger-Müller counter.

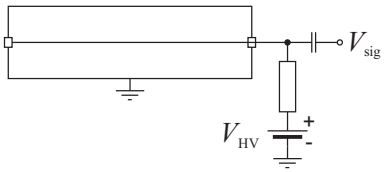


Figure 22: Discharge-limiting resistor in a Geiger-Müller counter.

- A poly-atomic gas like CH_4 or iso-C_4H_{10} is added, that has a large absorption coefficient for photons in the few-eV range, by excitation of rotational and vibrational levels. Ultimately, the energy is then safely absorbed in dissociation or elastic collisions of these molecules.

This is called a 'quencher gas'.

Equation 7) implies the proportionality of the signal charge with the primary charge generated in the gas volume by the incoming particle. Typical gas gains for proportional avalanche multiplication are 10^4 to 10^6. However, at higher fields so many secondary electron-ion pairs can be created in the avalanche that their collective space charge reduces the

local electric field. At this point the proportionality starts to break down. At even higher fields and thus gas gains space charge effects together with photon feedback can create macroscopic discharge channels or 'streamers', for which the signal becomes completely independent of the amount of primary charge. In this mode the typical gas gains are between 10^7 and 10^9. Finally, at extremely high fields the detector will go into plasma discharge, at which operational stability can only be restored by external reduction of the voltage (Geiger-Müller mode).

This is called the 'limited proportionality' regime.

Another reason for keeping the gas gain small is that ageing effects in the detector typically scale with the accumulated charge on the electrodes. Lower gas gain thus can help to delay the onset of ageing effects.

While gaseous ionization chamber type detectors do exist, the classical geometry for gaseous detectors is the 'wire chamber', where the amplification field is generated close to thin anode wires ($E \propto 1/r$). The electrons from the primary ionization first drift to the wire, where avalanche multiplication takes place in the high field close to the wire. The ions generated in the multiplication process then drift back from the wire, which generates the biggest part of the signal.

An example is the Resistive Plate Chamber (RPC), in which the field electrodes are made of resistive material, so that the field drops locally in response to a discharge.

Avalanche multiplication is again a stochastic process, and thus the amount of secondary charge for a given primary charge will vary from event to event. The gain distribution is exponential for lower amplification fields, but becomes more peaked for higher field and for small anode wires.

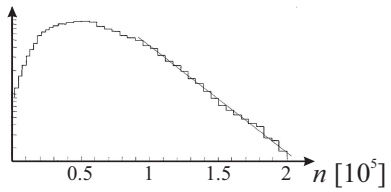


Figure 23: Distribution of number of ions generated by single electrons in a parallel plate geometry in Methylal. (H. Schlumbohm, Z.Phys.151 (1958) 563.)

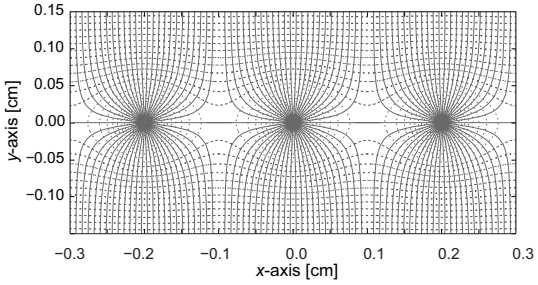
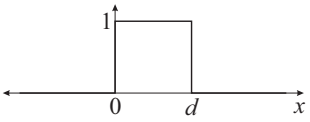


Figure 24: Field and equipotential lines in a MWPC (modified from PDG, Prog. Theor. Exp. Phys. 2020, 083C01 (2020).

The simplest implementation of a wire chamber for position detection is the Multi-Wire Proportional Chamber (MWPC), which is an array of cells with one anode sense wire each. The spatial information is given by the binary information cell hit/not hit. Thus the position resolution of such a detector is $d/\sqrt{12}$, where d is the cell size. As the cell size of a practical detector is limited to a few mm due to mechanical stability of the wires at high voltage the position resolution is in the order of 1 mm.

The spatial resolution of a binary (hit/no hit) detector with cell size d can be described by the RMS of the positions covered by the cell.



$$\begin{aligned}\Delta x &= \sqrt{\langle x^2\rangle - \langle x\rangle^2} \\ &= \sqrt{\frac{1}{d}\int_0^d x^2\,\mathrm{d}x - \left(\frac{1}{d}\int_0^d x\,\mathrm{d}x\right)^2} \\ &= \frac{d}{\sqrt{12}} \simeq 0.29d.\end{aligned}$$

The obvious way to improve the spatial resolution of MWPCs is to reduce the cell size. One way to deal with the electrostatic instability issues of thin conductors is to mount them on an insulating substrate. Advances in micro-fabrication led to the development of Micro-Strip Gas Chambers (MSGCs), where the electrodes are metal traces deposited on an insulating substrate like glass. However, this type of detector suffers from catastrophic discharges between the electrodes and charges generated during the avalanche multiplication that stick to the insulating substrate. Alternative charge multiplication geometries had to be found, while maintaining the high segmentation of the readout electronics.

Detectors of these type are sometimes referred to as 'Micro-Pattern Gas Detectors' (MPGDs).

Today, two main technologies are used: 'Gas Electron Multipliers' (GEMs) are Kapton foils which are stretched above the readout plane. The two sides of the foil are coated with a metallic layer. Through this assembly a tight grid of small holes has been etched. For operation, a voltage is applied between the two faces of the foil, which generates a high electric field, which is sufficient to induce impact ionization when electrons pass the hole. Typical gas gains of ~100 are achieved with one foil and higher gas gains can be achieved by stacking several foils. Ultimately, the amplified charge is then registered on a grid of readout electrodes on an insulating layer behind the GEM foils.

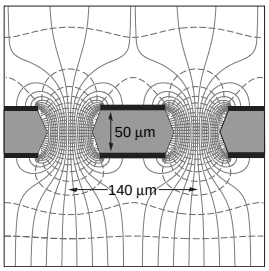


Figure 25: Field lines in a GEM (PDG, Prog. Theor. Exp. Phys. 2020, 083C01 (2020)).

A similar, but slightly different approach are 'Micromegas', which are stacks of single layer perforated metallic foils. Again, electrons pass through the holes, but the amplifying field is this time achieved by the potential difference between consecutive foils.

A very powerful detector can be built using such amplification structures with a highly segmented Silicon pixel readout chip.

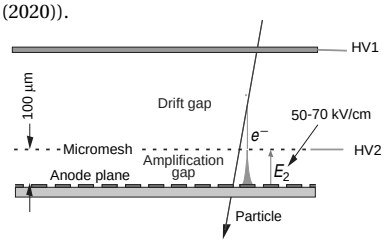


Figure 26: Schematics of a Mircomegas detector (PDG, Prog. Theor. Exp. Phys. 2020, 083C01 (2020)).

Charge drift

A different approach to improve the position resolution of gaseous detectors is to exploit the finite speed, with which the electrons move through the gas. By measuring the time it takes the primary electrons to move from the point of the ionization to the point where the actual electric signal is generated by the massive increase of charge in the avalanche, one can deduce the drift distance, if the drift velocity is known.

The motion of the charge carriers in the gas at a microscopic level is the result of the acceleration of the particles in the electric drift field, and collisions with the gas molecules. Macroscopically, this is described by the Langevin equation

$$m\frac{\mathrm{d}\vec{u}}{\mathrm{t}} = e\vec{E} + \vec{A}(t), \tag{8}$$

where $\vec{u}$ is the macroscopic drift velocity, and $A(t)$ is a force term describing the interaction between the drifting particle and the gas molecules. On average, the latter can be described by $\langle\vec{A}(t)\rangle = -m\vec{u}/\tau$, where τ is the average time between collisions. In the steady state this leads to

$$\vec{u} = \frac{e\tau}{m}\vec{E} = \mu\vec{E}, \tag{9}$$

where we have introduced the mobility μ. For the ions the drift velocity is to good approximation proportional to E/p, implying that the mobility is constant at a given pressure. Typical ion mobilities are between 1 and 10 $\mathrm{cm}^2/\mathrm{Vs}$.

For electrons the situation is different, as the collision cross-section and the fraction of energy lost per collision are both strongly dependent on the energy of the electron and the type of molecule. Hence there are large variations of the drift velocity for different gases and different values of E/p. Typical electron drift velocities are between 0.1 and a very few 10 $\mathrm{cm}/\mu\mathrm{s}$.

Such a detector is called a 'drift chamber'.

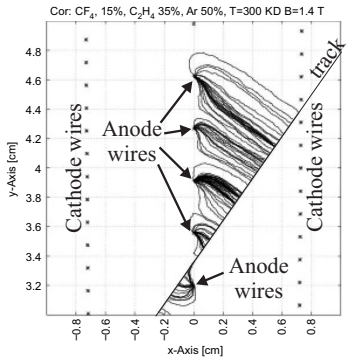


Figure 27: Simulated electron drift in the CDF outer tracker (modified from CDF, NIMA 526 (2004) 249).

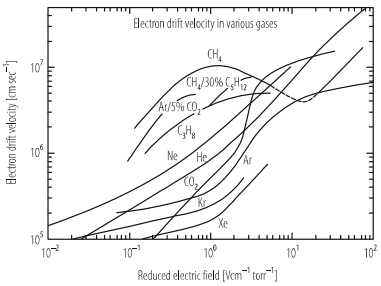


Figure 28: Drift velocities for different gases (A.Jeavons et al., NIM 176 (1980) 89).

The collisions with the gas molecules are stochastic. Macroscopically, this results in slight variations of the drift velocity, which manifest themselves as diffusion for an ensemble of drifting charge carriers. From the diffusion equation a Gaussian distributed cloud of charge carriers will expand in time in one dimension as

$$n(x) = \frac{n_0}{\sqrt{4\pi D t}} \exp\left(-\frac{x^2}{4Dt}\right), \tag{10}$$

where D is the diffusion constant. The width of the cloud can be described by its RMS and increases as $\sigma = \sqrt{2Dt}$. This spreading also applies if there is an average macroscopic drift velocity, so that $x \to x - ut$.

The diffusion coefficient is given by

$$D = \frac{\langle v^2 \rangle \tau}{3} = \frac{2\langle \epsilon \rangle}{3m}\tau,$$

where τ is the mean time between collisions, and $\langle \epsilon \rangle$ the average energy of the electrons. In the case that the energy of the electrons is given by the thermal limit $\langle \epsilon \rangle = 3/2 k_\mathrm{B} T$,

$$D = \frac{k_\mathrm{B} T}{m}\tau = \frac{k_\mathrm{B} T}{e}\mu,$$

and the width of the diffusion cloud is growing like

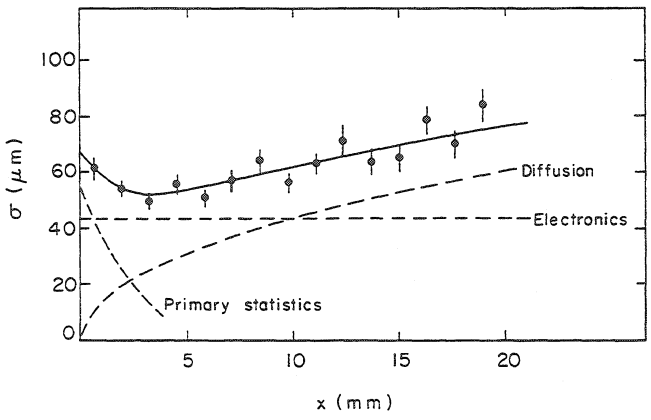
$$\sigma = \sqrt{\frac{2 k_\mathrm{B} T \mu t}{e}} = \sqrt{\frac{2 k_\mathrm{B} T l}{eE}},$$

where l is the drift distance. In real gases at high E/p, $\langle \epsilon \rangle \gg k_\mathrm{B} T$ and

$$\sigma = \sqrt{\frac{4 l \langle \epsilon \rangle}{3eE}}.$$

Gases, which stay close to the thermal limit up to high values of E/p (for example CO_2) are called 'cold' gases, the opposite being 'hot' gases (for example argon).

The increase in diffusion with drift distance entails a degradation of the position resolution for increasing drift distances. For short drift distances, on the other hand, it is the finite spacing between primary ionization clusters in the gas that dominates the position resolution.



A more careful analysis reveals that the energy differences between the front and tail of a drifting ensemble lead to a difference in mobility. Consequently, while the diffusion equation still holds in principle, the transverse and the longitudinal diffusion coefficient D_t and D_l are not equal.

After $N = t/\tau$ collisions, the displacement is given by $x = \sum_{i=1}^{N} \xi_i$. Hence

$$\sigma^2 = \langle x^2 \rangle = N\langle \xi^2 \rangle = N\langle v_x^2 \rangle \langle t^2 \rangle =$$
$$= \frac{N}{3}\langle v^2 \rangle \langle t^2 \rangle = \frac{2N}{3}\langle v^2 \rangle \tau^2 = \frac{2}{3}\langle v^2 \rangle \tau t,$$

and thus $D = \frac{\langle v^2 \rangle \tau}{3}$.

This equation is known as the Nernst-Townsend or Einstein equation.

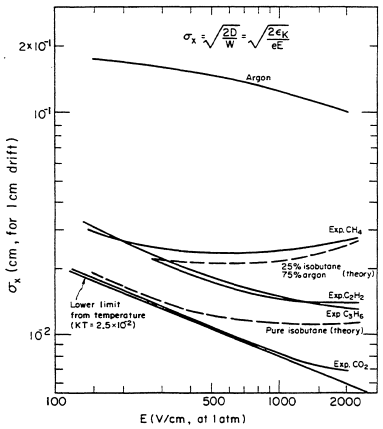


Figure 29: Diffusion in different gases (F. Sauli, CERN-77-09).

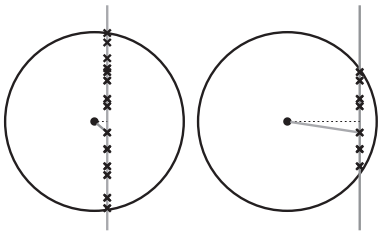


Figure 30: Effect of primary statistics on resolution. Crosses represent primary ionization clusters. Ideal distance (dashed) and distance to first ionization cluster (full).

Figure 31: Typical drift distance dependence of the position resolution in a drift chamber (F. Sauli, CERN-77-09).

In a magnetic field the Langevin equation becomes

$$\frac{\mathrm{d}\vec{u}}{\mathrm{d}t} = \frac{e}{m}\vec{E} + \frac{e}{m}\left(\vec{u}\times\vec{B}\right) - \frac{1}{\tau}\vec{u}.$$

In steady state ($\mathrm{d}\vec{u}/\mathrm{d}t = 0$), and the solution for the drift velocity is

$$\vec{u} = \frac{e}{m_e}\frac{\tau}{1+\omega^2\tau^2}\left(\vec{E} + \frac{\omega\tau}{B}\vec{E}\times\vec{B} + \frac{\omega^2\tau^2}{B^2}\left(\vec{E}\cdot\vec{B}\right)\vec{B}\right),$$

with the cyclotron frequency $\omega = eB/m_e$. In tracking detectors the E and B fields are often perpendicular, as the electron drift, which is providing the accurate position information, and the bending of the incoming particle track are both in the same direction. In this case the solution reduces to (using $\vec{E} = (E_x, 0, 0)$ and $\vec{B} = (0, 0, B_z)$)

$$u_x = \frac{e}{m_e}\frac{\tau}{1+\omega^2\tau^2}E,$$
$$u_y = -\frac{e}{m_e}\frac{\tau}{1+\omega^2\tau^2}\omega\tau E$$
$$u_z = 0.$$

Hence, the motion of the electron is rotated around the B field by an angle $\Psi = \mathrm{atan}(u_y/u_x) = \mathrm{atan}(-\omega\tau)$. Ψ is called the 'Lorentz angle'. The speed of the drifting electron is given by $u = (e/m_e)\tau E\cos\Psi$, hence it is determined by the electric field in the drift direction. Similarly to the drift velocity, the Lorentz angle depends on the drift gas and E/p.

Another important field configuration is a parallel E and B field. In that case the drift velocity is not modified by the magnetic field. However, microscopically the trajectories of the drifting electrons between collisions become helices around the field direction. When the magnetic field is strong enough this will reduce the transverse diffusion as $D_\mathrm{t}(B)/D_\mathrm{t}(0) = (1+\omega^2\tau^2)^{-1}$. The longitudinal diffusion is unaffected.

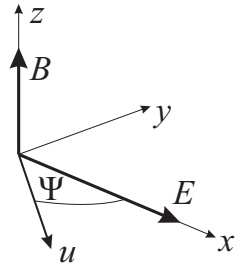


Figure 32: Lorentz angle in orthogonal electric and magnetic field.

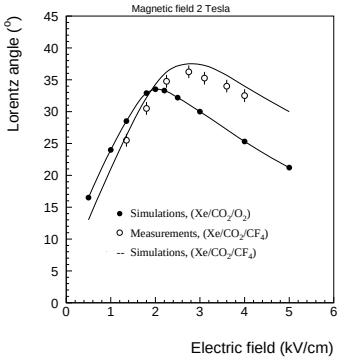


Figure 33: Lorentz angle in different gases (ATLAS TRT, 2008 JINST 3 P02013).

Typical drift chamber geometries are either individual tubular or square sections for integration into planar structures, or larger systems which are designed to fill the central cylindrical tracking cavity in a collider detector.

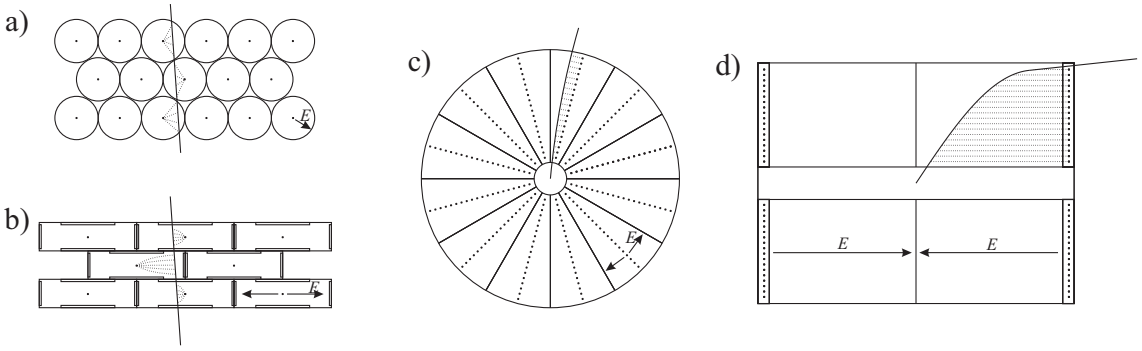


Figure 34: Drift chamber geometries. a) Planar made of tubes. b) Planar made of square cells. c) Jet cells (endview). d) Time Projection Chamber (side view). Not to scale.

A historically important drift chamber for central tracking is the 'jet chamber', where cells or radial layers of anode wires provide a large number of points (about 1 hit per cm) for tracks from the vertex. These type of detectors had good two-track resolution (~100 μm), good pattern recognition capabilities, and measurement of $\mathrm{d}E/\mathrm{d}x$. To compensate for the Lorentz angle the cells were often tilted.

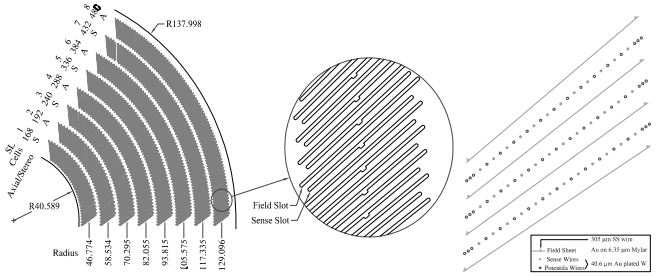


Figure 35: Example for a jet cell geometry: CDF Central Outer Tracker (CDF, NIMA 526 (2004) 249).

A gaseous drift chamber technology that provides 3D information over a large volume with minimum material is the Time Projection Chamber (TPC). It is a cylindrical gas filled volume split into two halves by a central cathode disk at very high voltage (100 kV in the ALICE TPC). Primary electrons drift to the endplates, which are instrumented with a 2-dimensional readout. In the case of existing and past TPCs this uses of grid of anode wires for avalanche multiplication. The position is determined from the ion signal induced on a plane of pads. To prevent ions drifting back into the detector volume a gating grid is used, which closes their path electrostatically after an avalanche on the anode wires. The $r\phi$ information in this detector is maintained despite the very long drift distance by the reduced diffusion transverse due to the parallel E and B fields.

In the TPC the z information is obtained from the drift time of the electrons to the readout plane. The long drift time makes this detector impractical for hadron colliders, but they are used for heavy ion experiments and are under consideration for future lepton colliders, where collision rates are lower. For future colliders, detector planes with better position resolution, for example based on MPGDs are being investigated.

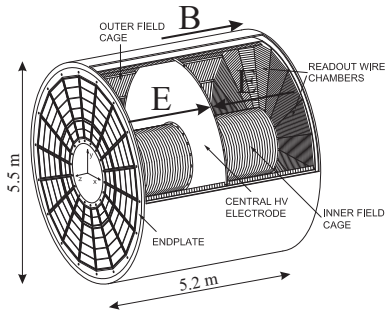


Figure 36: ALICE TPC (modified from ALICE, NIMA 622 (2010) 316).

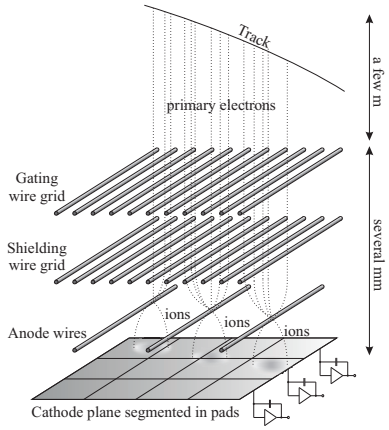


Figure 37: Schematics of TPC readout.

Semiconductor detectors

In solids the stopping power for a charged particle is higher due to the large density. The key challenge is then to separate the electron from the positive ion or hole, and collect it on a readout electrode.

In solids higher energy electron states occur in bands. Typically most valence electrons reside in the valence band, where they are stationary. To move, they have to be in the higher-energy conduction band. We distinguish three cases:

There are additional benefits to solid detectors: The range of δ electrons is smaller, and the detector itself is mechanically stable and does not need a container.

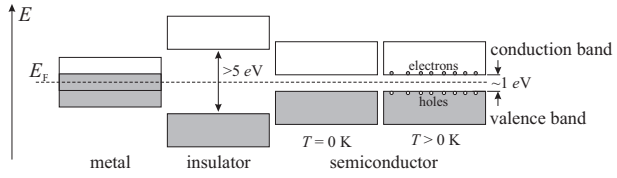


Figure 38: Band structure in solids.

- In metals, the conduction and the valence band overlap. The number of mobile electrons far exceeds the number of electrons generated by ionization, and the ionization signal is unobservable.

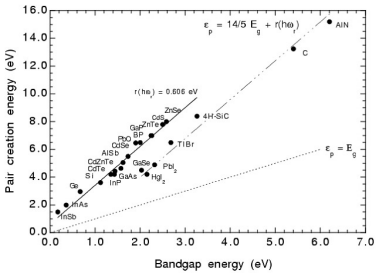


Figure 39: Energy used per electron hole pair vs band gap in different materials (https://sci.esa.int/s/wVMvnRA). The difference between the energy used per electron-hole pair and the band gap is due to lattice excitations (phonons) to conserve 4-momentum.
In silicon, the band gap is 1.12 eV, but the average energy used per electron-hole pair is 3.6 eV.

- In semiconductors, there is a finite, but small band gap (a few eV). Some electrons are lifted to the conduction band thermally, but this population can be kept small by lowering the temperature or reducing charge carrier concentrations by doping and biasing. Electrons can then be moved to the valence band by the interaction with the incoming particle, where they can move to the readout electrodes.
- In insulators the band gap is larger (>5 eV). Insulators can be used as particle detectors (for example diamond with a band gap of 5.5 eV), but only if the material is very pure. Normal insulators have lots of interstitial states, which will trap the electrons and result in local de-excitation, strongly limiting the range of the electrons, so that a signal cannot be collected.

In the following we will focus on silicon detectors, as the most common semiconductor detector material. Undoped silicon has an intrinsic charge carrier density of

$$n_i = \sqrt{N_\mathrm{c} N_\mathrm{v}} \exp\left(-\frac{E_\mathrm{g}}{2k_\mathrm{B}T}\right) \text{ with } N_\mathrm{i} = 2\left(\frac{m_\mathrm{i}^* k_\mathrm{B} T}{2\pi\hbar^2}\right)^{3/2}, \tag{11}$$

where E_g is the band gap energy (1.12 eV in silicon) and the N_i are the numbers of effectively available states for electrons per unit volume in the conduction band (N_c) and for holes in the valence band (N_v). For pure silicon at room temperature $n_i \simeq 10^{10}$ cm^{-3}. For a detector element with an area of 1 mm^2 this gives $n_i/d \simeq 10^8$ electron/hole pairs/cm. If we compare this with the ionization produced by the passage of a charged particle, we get

$$\frac{n_\mathrm{ion}}{d} = \frac{\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_\mathrm{mip}}{E_\mathrm{e/h}} = \frac{2\ \mathrm{M}e\mathrm{Vg}^{-1}\mathrm{cm}^2 \cdot 2.33\ \mathrm{gcm}^{-3}}{3.6\ e\mathrm{V}} = 1.2 \times 10^6\ \mathrm{cm}^{-1},$$

two orders of magnitude smaller. Thus, to make a viable silicon detector, we need to remove the intrinsic carrier concentration. This we can achieve by joining n- and p-doped silicon. In n (p) doping small amounts of a period 3 (period 5) element are infused into the silicon lattice. These atoms act as electron acceptors (donors), increasing the number of holes in the valence band (electrons in the conduction band). Doping also changes the Fermi levels in the two regions according to

$$\Delta E_\mathrm{F}^\mathrm{n,p} = k_\mathrm{B} T \ln\left(\frac{N_\mathrm{a,d}}{n_i}\right), \tag{12}$$

where $N_\mathrm{a,d}$ is the acceptor or donor density, respectively. p doping brings the Fermi energy closer to the conduction band, n doping brings it closer to the valence band.

At the junction of n- and p-doped silicon the surplus charge carriers in each doping region diffuse into the other doping region and neutralize until thermal equilibrium is reached. As a result, the area around the junction gets depleted of charge carriers. The ions remain and create a space charge and an electric field stopping further diffusion.

The depth of the depletion zone can be found from integrating the Poisson equation, $\mathrm{d}^2\Phi/\mathrm{d}x^2 = -\rho/\epsilon$, where ϵ is the permittivity of silicon. In a simple model the number of free charge carriers is assumed to

For semiconductors at room temperature

$$\begin{aligned} n(E)dE &= g(E)\frac{1}{\exp\left(\frac{E-E_\mathrm{F}}{k_\mathrm{B}T}\right)+1}\mathrm{d}E \simeq \\ &\simeq g(E) e^{-\frac{E-E_\mathrm{F}}{k_\mathrm{B}T}}\mathrm{d}E \\ &\simeq \frac{(2m^*)^{3/2}}{2\pi^2\hbar^3}\sqrt{E-E_\mathrm{c}}\, e^{-\frac{E_\mathrm{c}-E_\mathrm{F}}{k_\mathrm{B}T}} e^{-\frac{E-E_\mathrm{c}}{k_\mathrm{B}T}}, \end{aligned}$$

where E_c is the start of the conduction band. After integration over the energy, and using $E_\mathrm{c} - E_\mathrm{F} = E_\mathrm{g}/2$

$$N = 2\left(\frac{m^* k_\mathrm{B} T}{2\pi\hbar^2}\right)^{3/2} \exp\left(-\frac{E_\mathrm{g}}{2k_\mathrm{B}T}\right).$$

m^* is the effective mass of the electrons and holes in the conduction and valence bands, respectively.

As we will see this only works for thin detectors. However, if we want to use a semiconductor detector to detect gamma rays (~ 1 MeV) with high efficiency we need a large active volume because of the long absorption length. There, the only way to reduce intrinsic charge carrier density is low temperature. The finest gamma spectroscopy detectors are germanium detectors (E_g = 0.67 eV) cooled by liquid nitrogen.

At this point the Fermi energy in the two regions becomes equal.

The permittivity is independent of doping, $\epsilon = 11.7\epsilon_0$, where ϵ_0 is the permittivity of the vacuum.

vanish over the depth of the depletion zone ($d = d_\mathrm{n} + d_\mathrm{p}$), but constant outside. As a result of the first integration $d_\mathrm{p}/d_\mathrm{n} = N_\mathrm{d}/N_\mathrm{a}$. The second integration then yields

$$d = \sqrt{\frac{2\epsilon}{e}\frac{N_\mathrm{a}+N_\mathrm{d}}{N_\mathrm{a}N_\mathrm{d}}\Delta V}. \tag{13}$$

For an unbiased pn junction, ΔV is the difference of the Fermi levels in the two doping regions. However, the depth of the depletion zone can be increased by applying a DC bias voltage (negative pole connected to the n doped, and positive pole to the p doped region). This is known as a reverse-biased diode. In that case $\Delta V = \Delta E_\mathrm{F}/e + V_\mathrm{bias} \simeq V_\mathrm{bias}$, i.e. a bias of 100 V will increase the depletion depth by more than a factor 10.

A typical particle physics silicon detector comprises a thin heavily p-doped (p^+) region on a lightly n-doped bulk material. For example, $N_\mathrm{a} \simeq 10^{15}$ cm^{-3} and $N_\mathrm{a} \simeq 10^{12}$ cm^{-3}, so that $\Delta E_\mathrm{F} \simeq 0.42$ eV and $d_\mathrm{p} = 0.02$ μm and $d_\mathrm{p} = 23$ μm without bias. With a bias voltage of 100 V this becomes $d_\mathrm{p} = 0.4$ μm and $d_\mathrm{p} = 363$ μm.

In practice, it is the depletion depth that is set (typically, for charge collection the aim is to maximize the depletion depth within the thickness of the silicon material), and the depletion voltage is chosen to achieve that depletion depth. This also limits the sensible thickness of silicon detectors of the reverse biased diode type to a few 100 μm.

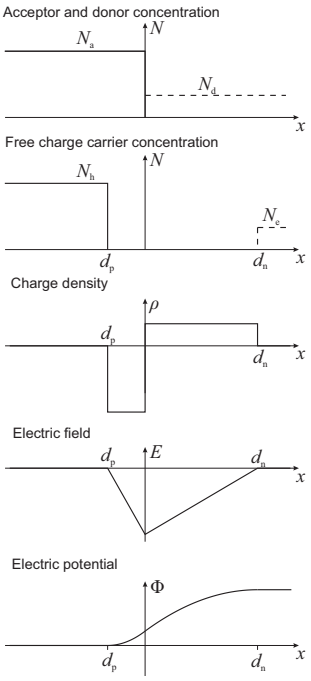


Figure 40: Simple pn junction model.

Even though the pn-junction is depleted of charge carriers there is a leakage current from thermally created electron hole pairs. According to eq. (11) the reverse bias leakage current has a strong temperature dependence.

$$I_\mathrm{leak} \propto T^2 \exp\left(-\frac{E}{2k_\mathrm{B}T}\right). \tag{14}$$

E in this equation is slightly larger than the band gap, 1.2 eV for silicon

As a rule of thumb, the leakage current doubles for an increase in the temperature of 7°C. Three things need to be considered about the leakage current:

It is common to quote the leakage current at a reference temperature (often 0°C). It can then be scaled to the appropriate temperature using eq. (14).

- The readout electronics needs to accommodate the constant current. Typically, the readout is capacitively coupled, so that this is taken care of;
- Statistical fluctuations due to the discrete nature of the moving charge carriers lead to shot noise, which scales with $\sqrt{I_\mathrm{leak}}$;
- The leakage current and the bias voltage across the pn junction results in ohmic heating of the silicon. Because of the finite heat capacity this will increase the temperature of the silicon, which will increase the leakage current and thus the leakage power, and so on. If the heat is not removed at low temperature, the silicon will go into 'thermal runaway'. Cooling of silicon sensors is therefore essential, both with sufficient power removal capacity and at sufficiently low temperature.

Thermal management of silicon sensors needs to take into consideration power from other heat sources on the detector (electronics, cables, etc.), the temperature of the local heat sink (typically a cooling fluid), and the thermal conductivity between the sensor and the heat sink.

A major consideration for current silicon detectors in high-rate environments is radiation damage. In particular, it is hadrons that in non-ionizing energy loss (NIEL) interactions damage the lattice structure

of the silicon. The faults in the lattice create interstitial states, which have two major consequences: First, the leakage current will increase (while maintaining the temperature dependence), with the associated challenges for the thermal management and increase in shot noise. Second, the lattice faults will manifest themselves as changes in the doping concentrations, even to the point that they can change the doping type. Effectively, this will require adjustment of the bias voltage to maintain full depletion, until the point when this is technically not possible any more, and degradation of the signal has to be accepted. The long-term results of irradiation are affected by temperature-dependent annealing processes, of which some are beneficial, reducing the effects of the radiation damage, and some are detrimental, making them worse. Designs for silicon systems in such environments thus need to take the temperature profile of the operation into account.

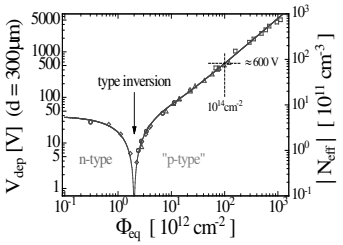


Figure 41: Effective doping concentration (depletion voltage) as a function of particle fluence for a standard float zone (FZ) n-type silicon detector (CERN Yellow report CERN-2021-001).

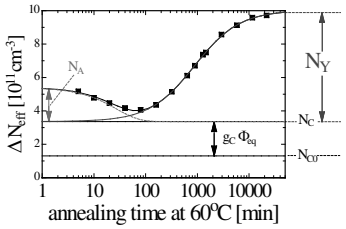


Figure 42: Evolution of the effective doping concentration as function of annealing time. Irradiation at room temperature while the annealing at 60°C (CERN Yellow report CERN-2021-001).

The most common type of silicon detectors are silicon strip detectors, which are segmented into strips that are typically of the order of 100 μm wide and several cm long.

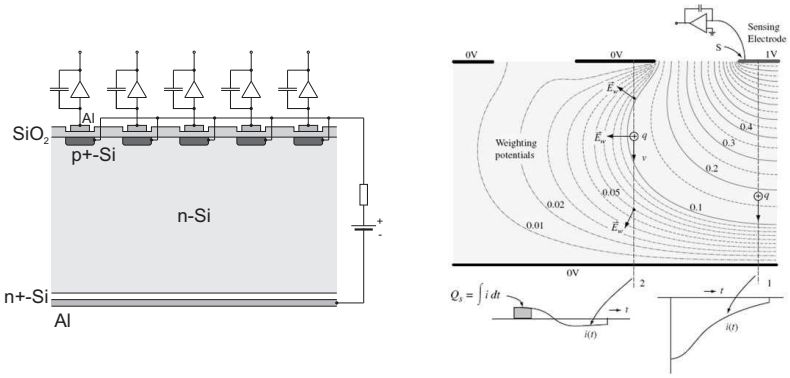


Figure 43: Left: Schematics of a p-type AC coupled silicon strip detector. Right: Weighting field in a strip detector. The curves at the bottom show the current signal for the two drift trajectories indicated above (V. Radeka, Particle Physics Reference Library, https://doi.org/10.1007/978-3-030-35318-6_10).

The drift of electrons and holes can be described using similar principles as for gaseous detectors. In particular, for electrons $v_n = -\mu_n E = -(e\tau/m_n)E$, and for holes $v_p = \mu_p E = (e\tau/m_p)E$, where τ is the average time between collisions. For semiconductors, τ decreases for high fields, and the drift velocity saturates. Diffusion again increases like the square root of the drift distance, and the diffusion coefficient can be found from the Nernst-Townsend equation. Finally, in a magnetic field perpendicular to the bias field, the drift direction of the charge carriers is again rotated by the Lorentz angle.

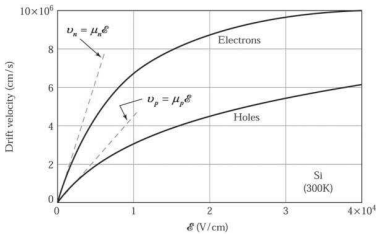


Figure 44: Charge carrier drift velocities in silicon (S.M.Sze, Semiconductor devices).

For vertex detectors, to cope with the high track densities at small radius and for improved vertex resolution, smaller size detector elements are needed. The challenge here is to make the connection to the readout electronics. Pixel sensors have a typical element size of 50 to a few 100 μm per side.

A hybrid pixel detector is made by bump-bonding the sensor to a readout chip with similar segmentation. The bump-bonds are small metallic balls (diameter few 10 μm) for each channel between the two layers of silicon, which are heated ('reflowed') to make a solder interconnect.

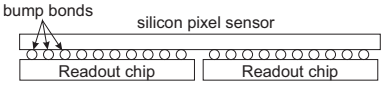


Figure 45: Schematics of a hybrid pixel detector with bump bonding.

Bump-bonding is a demanding operation, and high yields are a challenge. It therefore suggests itself to integrate at least the first stages of the

analogue readout into the pixel sensor. Such a detector is called a Monolithic Active Pixel Sensor (MAPS). Current approaches exploit the wide availability of CMOS technology for this approach. However, MAPS are very complex devices and design of large devices as needed in particle physics and achieving radiation hardness are ongoing challenges.

The small band gap in semiconductor detectors makes them attractive for the detection of photons. The simplest such detector is a photodiode. Silicon photodiodes are typically sensitive to light in the spectral range from about 200 nm (near UV) to about 1100 nm (near IR), thus exceeding the spectral response of the human eye. Photosensor responsivity is measured in Amperes of photocurrent generated per Watt of incident light power, and varies with the wavelength of the incident light, with peak values from 0.4 to 0.7 A/W.

Other photodiode materials are GaAs or InGaAs, which have a faster response and are therefore used in telecommunications via optical fibres.

Photodiodes can be used in two configurations:

- Photovoltaic mode: The photodiode is unbiased. (This is the mode of operation for a solar cell.)
- Photoconductive mode: An external reverse bias is applied. This reduces the response time because the reverse bias increases the depth of the depletion layer, which decreases the junction's capacitance. But the bias also increases dark current (a current that is not associated with incoming photons), which leads to higher noise levels.

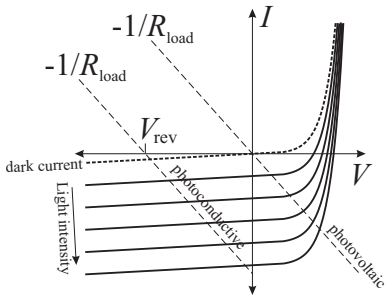


Figure 46: Typical I-V characteristic of a photodiode

Quantum efficiency is defined as the ratio of the number of produced charge carriers per incident photon.

For visible light the mean free path in silicon is about 1 μm, much smaller than the depletion depth, and thus photodiodes are capable of a high quantum efficiency (up to 80%). However, for low intensities (low photon incidence rate), the signal becomes too small for detection. The answer is again to employ an internal amplification process based on impact ionization (avalanche multiplication). In the Avalanche Photo Diode (APD) a strong local field is generated at the junction by the choice of doping and a high reverse bias voltage. This field is sufficient to accelerate the charge carriers to achieve impact ionization. The typical gain in such devices is about 100, but APDs with gains of several 10^3 are available. At very high gain single photon detection is possible, but proportionality of the signal is lost.

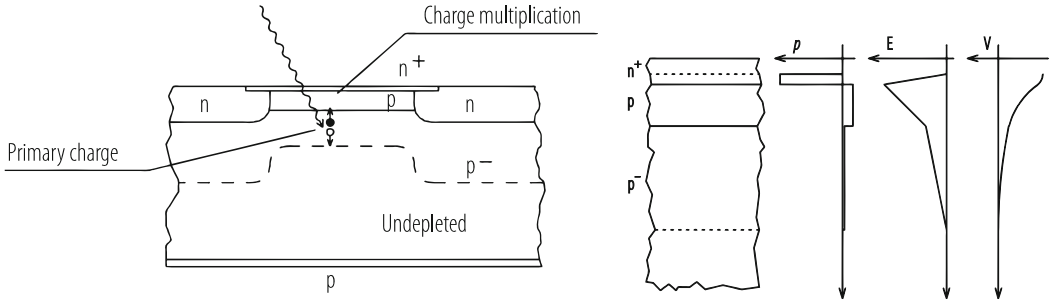


Figure 47: Avalanche photodiode (G.Lutz, R.Klanner, Particle Physics Reference Library, https://doi.org/10.1007/978-3-030-35318-6_5).

Response to excitations

So far, we have discussed detectors that relied on the collection of ionization charge produced by an incoming charged particle. We have seen that the size of these detectors has been driven by the dimensions of the field which can be provided for the collection of the charge. In semiconductor detectors this dimension could be made very small, so that such a detector has an excellent space resolution. However, sometimes position resolution is not the primary target, but a large dense detector volume or cheap instrumentation of a large volume is required. In that case it is often easier to collect photons, which have a long range in optically transparent materials. Such photons can be generated in detector materials called 'scintillators' from the relaxation of excited states created by the electromagnetic interaction of the incoming charged particle with the detector.

Organic scintillators

In organic scintillators it is the de-excitation of molecular electrons that generates the photons. Organic substances which contain aromatic rings, such as polystyrene (PS) and polyvinyltoluene (PVT), do display this effect. These are polymers, which can be conveniently shaped and machined. The molecules typically have a spectrum of electronic singlet and triplet states, with a typical separation at *e*V level, plus a fine structure of vibrational levels with a ten times smaller spacing. At room temperature most electrons are in the S_0 singlet ground state. The electromagnetic interaction with an incoming particle lifts the electrons into an excited state somewhere in this spectrum. The singlet excitations decay very quickly (<10 ps) to the S* first excited state (internal degradation), which subsequently decays to the electronic ground state, but typically to one of the vibrationally excited S_0 levels, with the emission of a photon. Re-absorption of the emitted photon is suppressed, as its energy is less than the transition ground state $S_0 \rightarrow S^*$. To achieve even larger macroscopic transmission distances, the trick is to add additional organic fluorescing compounds, which degrade the photon energy further. The first step in this degradation cascade happens on such small distance scales, that the energy is not transferred by photons, but by a resonant dipole-dipole interaction (Förster energy transfer). Because organic scintillators involve relaxation of electronic states within one molecule the emission process can be fast (a few ns), although delayed photon emission can occur from triplet states, for which transitions to the ground state are forbidden by selection rules.

Plastic scintillators can be conveniently made in many different shapes and sizes. One special geometry are scintillating fibres (SCIFI) with diameters down to a few 100 μm. Multi-clad fibres have an appropriately doped polystyrene-based core and a two-layer cladding of polymethyl methacrylate (PMMA) and a fluoroacrylic to improve the trapping of light in the fibre by total internal reflection to about 5%.

Because scintillation in organic scintillators is happening within molecules, the actual state of the material is not that relevant, and there are also crystalline, liquid and gaseous organic scintillators, although they are not commonly used in particle physics.

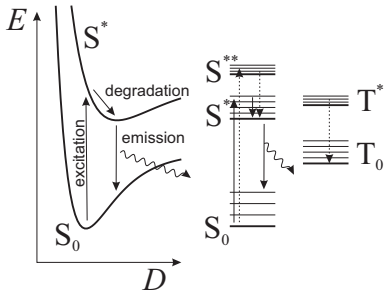


Figure 48: Organic scintillator. Left: Potential energy vs atomic separation. Right: Energy levels.

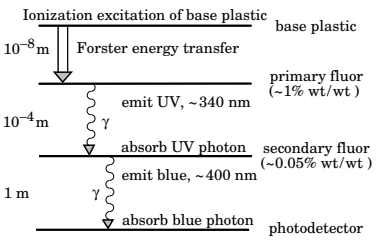


Figure 49: Photon energy degradation in an organic scintillator (PDG, Prog. Theor. Exp. Phys. 2020, 083C01 (2020)).

Inorganic scintillators

Inorganic scintillators are transparent crystals, usually with some dopant (for example NaI, doped with Tl). The interaction with the incoming charged particle lifts electrons well above the band gap and leaves deep holes in the valence band and below. These lose energy in inelastic collisions, until all charge carriers are contained in the valence and conduction bands. These states then thermalize to the lower edge of the conduction band (electrons) and the top of the valence band (holes). The electrons then form excitons (electron/hole states just inside the band gap). They travel together until they get captured as a whole by impurity centres (activators). These usually have been deliberately introduced. Finally, the excited activator states decay, emitting a visible or UV photon, which can travel long distances as their energy is smaller than the band gap.

Because the process requires movement over significant distances, crystal scintillators are usually not very fast detectors.

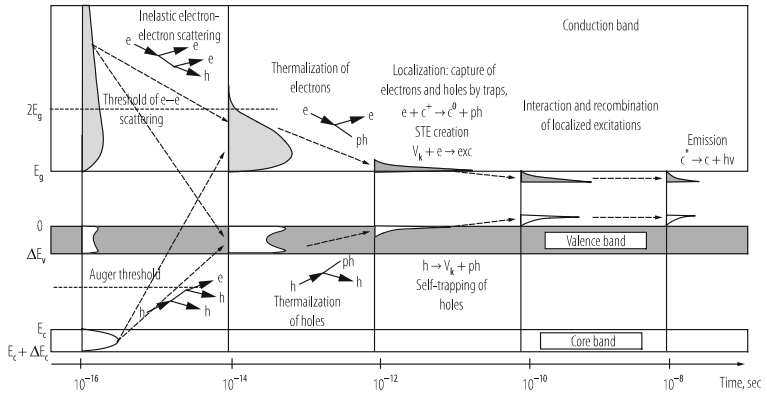


Figure 50: Relaxation scheme for electronic excitations in an insulator: e, electrons; h, holes; ph, phonons; hν, photons; V_k, self-trapped holes; STE, self-trapped excitons; c^n , ionic centers with charge n. Density of states is represented by grey and white areas for electrons and holes respectively (Lecoq P. (2020) Particle Physics Reference Library. https://doi.org/10.1007/978-3-030-35318-6_3).

Not all crystals require doping: Bismuth Germanium Oxide (BGO), for example, is a pure inorganic scintillator without any activator impurity. There, the scintillation process is due to an optical transition of the Bi^{3+} ion, a major constituent of the crystal. In $CaWO_4$ and $CdWO_4$ the photon emission is due to radiative decay of self-trapped excitons.

These crystals are particularly interesting for calorimetry because of their short response time, high density and short radiation length.

The energy of the emtitted photons has a broad spectrum because of distortions of the emission centres due to lattice deformations and by temperature broadening of the optical transitions. Typically, about 10 to 100 *e*V are required per photon produced, depending on the band gap. Practically, the energy per detected photon is larger, because of photon loss due to absorption and on surfaces, and inefficiencies in the detection. Photon detectors commonly used are PMTs for large segmentation, and APDs for small segmentation. Coupling of the scintillator to the photon detector is critical.

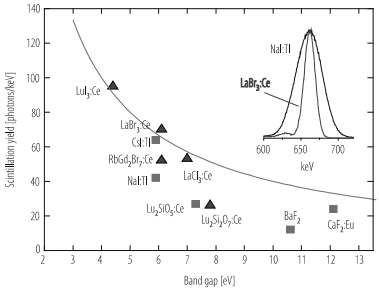


Figure 51: Photon yield/k*e*V of several inorganic scintillators vs band gap (Lecoq P. (2020) Particle Physics Reference Library. https://doi.org/10.1007/978-3-030-35318-6_3).

Emulsions, cloud and bubble chambers

So far, we have exclusively discussed particle detectors that ultimately produce an electrical signal. Nowadays, these are the most widely used detectors, because they are responsive and fast, and their data can be easily processed, both, in fast decision online decision logic (triggers), and for offline analysis. However, historically, high precision particle

Responsive here means that they are ready to record an event at any time, without preparation, and at high rate/

detectors have been built, which did not rely on charge collection, but made use of the local change of the molecular structure due to ionization and excitations caused by traversing particles.

An example for this is the photographic emulsion. This consists of silver halide crystals typically immersed in a gelatinous carrier. If a halogen atom in the crystal gets ionized, the electron can combine with a silver ion to a neutral silver atom. Several silver atoms then form a small cluster. Exposure to a development solution will make the silver clusters grow, until they cover a macroscopic area, which can be resolved by eye. Photographic emulsion has been used historically in particle physics (the first observation of subatomic particles has been made by Becquerel by accidentally exposing photographic plates to beta radiation from Uranium salts), but is still in use today, when very high resolution is needed, but at low speed and cost (for example in τ neutrino experiments).

Other historic types of detectors are cloud and bubble chambers. Here, a homogenous detector medium is prepared in a metastable state (supercooled vapour or superheated liquid), and the excitations and ionizations created by the passage of a charged particle act as nuclei for the retarded phase transition. The pattern of droplets or bubbles were then recorded using photography and analysed after scanning of the image.

At this point we speak of a 'latent image').

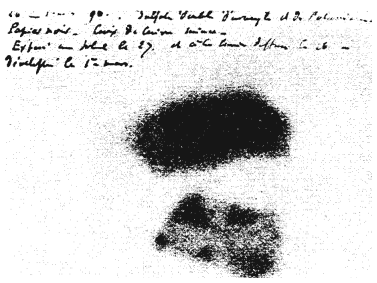


Figure 52: Image of photographic plate after exposure to beta radiation from Uranium salts (A.H.Becquerel, Nobel Lecture 1903).

Track reconstruction

A powerful way to find the kinematic properties of charged particles even in an environment with many particles is the measurement of the position at multiple locations, and the combination of the individual space-points to a track. Depending on the detector used the tracks are typically reconstructed from either a limited number of points with excellent space point resolution (typical for silicon trackers), or many points with a somewhat poorer resolution (e.g. in a TPC). Independent of the nature of the track, the usual steps in the reconstruction are

1. Find combinations of space points, which are likely to belong to one track ('pattern recognition');
2. Find for each track the track parameters, and, to estimate the errors on these parameters, calculate the associated covariance matrix;
3. Test the track hypothesis and identify outliers and wrongly assigned points. This includes detection of kinks in the track, which are indicative of a decay vertex.

For a straight track these are the direction and offset, but the parametrization can be much more complicated, for example in a non-uniform magnetic field.

The quality of the track parameters will be affected by measurement errors of the particle position in the different detector layers, errors on the knowledge of the position of the sense elements with respect to each other and a global coordinate system, and disturbances of the particle's flight path, like for example by multiple scattering, decays and energy loss of the particle, or magnetic field variations.

In complex detector systems the track is often extrapolated forward and backward, to match it with signals in other sub-detectors, most importantly the calorimeters, but also possible particle identification detectors or other tracking systems. Also, often an important information is the geometrical relation of the track to a known primary interaction vertex. If the back-extrapolation of the track misses the primary vertex, this indicates that this track likely originated from a secondary vertex. This is particularly important for the reconstruction of particles that have too short a lifetime to reach any active part of the detector. In collider experiments these are typically mesons containing b and c quarks, or τ leptons. These particles usually propagate for less than 1 mm before they decay, so for good vertex reconstruction it is important to provide detector layers with very good position resolution very close to the region of the primary vertices, and to minimize potential disturbance of the track by multiple scattering.

The distance of the primary vertex to the back-extrapolated track is called the 'impact parameter'.

Typically, the resolution on the impact parameter d_0 can be described by $\sigma(d_0) = A \oplus B/p$, where p is the momentum of the particle.

Momentum measurement in magnetic fields

In a homogenous magnetic field B the trajectory of a particle with charge ze becomes a helix with radius R and pitch angle λ. The momentum is then given by

$$p_\mathrm{t} = p\cos\lambda = zeRB.$$

Often a dimensional version of this equation is used

$$p_\mathrm{t}[\mathrm{GeV}] = 0.3zR[\mathrm{m}]B[\mathrm{T}].$$

In the case of three equidistant position measurements, the sagitta of the track is measured

$$s = R - R\cos\frac{\Phi}{2} \simeq R\frac{\Phi^2}{8} \simeq \frac{R}{8}\frac{L^2}{R^2} = \frac{L^2}{8R} = \frac{zeBL^2}{8p_t}.$$

If the sagitta measurement is used to determine the momentum, then the error on the momentum measurement is related to the error on the sagitta

$$\frac{\sigma_{p_t}}{p_t} = \frac{\sigma_R}{R} = \frac{L^2}{8sR}\frac{\sigma_s}{s} = \frac{\sigma_s}{s} = 8R\frac{\sigma_s}{L^2} = \frac{8}{ze}\frac{p_t}{BL^2}\sigma_s.$$

The sagitta error is closely correlated with the position measurement error in each layer. For the simple case used here with three equidistant stations $s = x_2 - (x_1 + x_3)/2$ and therefore, assuming the same position measurement error σ_x in the three stations,

$$\sigma_s = \sqrt{\sigma_x^2 + 2\frac{\sigma_x^2}{4}} = \sqrt{\frac{3}{2}}\sigma_x.$$

It should be obvious that for a good momentum measurement a large tracking volume, a large magnetic field and good space point resolution are required.

At low momenta, the effects of multiple scattering must also be considered. If we stay with the simple system of three thin stations for the sagitta measurement with no scattering material between, then for the three stations

$$\Delta x_1 = 0, \quad \Delta x_2 = \frac{L'}{2}\Delta\theta_1, \quad \Delta x_3 = L'\Delta\theta_1 + \frac{L'}{2}\Delta\theta_2,$$

where $L' = L/\sin\Theta$, with Θ being the angle of the track with the magnetic field. Therefore

$$\Delta s = \frac{L'}{2}\Delta\theta_2 \text{ and } \sigma_s = \frac{L'}{2}\sigma_\theta \simeq \frac{L'}{2}\theta_0,$$

where θ_0 is given by eq. (3). For a uniform distribution of scattering material the integration yields $\Delta s = (L'/3)\theta_0$.

To simplify eq. (3) further (which, as discussed, already ignores the large-angle scattering tails), we drop the logarithmic part to get for uniform scattering material

$$\left(\frac{\sigma_{p_t}}{p_t}\right)_{\text{MS}} = \left(\frac{\sigma_s}{s}\right)_{\text{MS}} \simeq \frac{\frac{L'}{3}\frac{13.6\,\text{MeV}}{\beta cp}z\sqrt{\frac{L'}{X_0}}}{\frac{zeBL^2}{8p\sin\Theta}} \propto \frac{1}{\beta B\sqrt{LX_0\sin\Theta}}.$$

The combined momentum resolution can thus be described by

$$\frac{\sigma_{p_t}}{p_t} = C_1\frac{p_t\sigma_x}{BL^2} \oplus C_2\frac{1}{\beta B\sqrt{LX_0\sin\Theta}}.$$

Hence, multiple scattering will be the dominant contribution at small momenta, while it will increase linearly for high momenta, due to the increasing effect of the space point position resolution. A good resolution requires large B and L, and good space point position resolution and small amounts of scattering material.

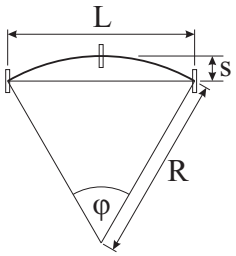


Figure 53: Sagitta geometry.

L is sometimes called the 'lever arm' of the momentum measurement.

For a large number ($n > 10$) of equidistant position measurements

$$\sigma_s = \frac{1}{8}\sqrt{\frac{720}{n+4}}\sigma_x.$$

This estimation is just a justification for the functional form of the result. The challenge of all calculations of multiple scattering is the integration of the effect for thick slabs of material. The theoretical apparatus on this is daunting, and the equations not very instructive, even if the results are highly accurate. The alternative is to obtain the multiple scattering distributions from Monte Carlo.

If the material is concentrated in discrete layers, the resulting dependence is more like L^{-1}.

For very low momenta energy loss fluctuations in can also be relevant.

Note that a large L usually implies a large magnet, and the stored energy of the magnet is proportional to the volume.

Calorimetry

Calorimetry in particle physics has little to do with the original meaning of the name. The term comes from early nuclear physics, where indeed the heat created in radioactive decays was measured. However, this was done for the radiation from sources comprising a large number of particles. The name was then applied to detectors that detected neutral particles through the shower processes described in the main text. The refinement for particle physics is then that the particle's energy is converted into a proportional signal. This is the meaning of the term in particle physics today. Some of the energy is indeed transferred into heat, but this means it is actually lost for the detection and a limiting factor of the measurement in the calorimeter.

The continuous small energy loss of charged particles enables the tracking of these particles. For neutral particles a different approach is needed. We have seen that in electromagnetic and hadronic interactions of high energy particles secondary particles can be created. If these secondaries have enough energy they in turn can produce further secondaries, so that a cascade of secondaries (a 'shower') is generated. Because the energy transferred to the secondaries is a significant fraction of the energy of the interacting particles, the energy of the original incoming particle is relatively quickly degraded in this process, so that the original particle is brought to a complete stop, and all its energy transferred to the detector material. This is the principle of a particle physics calorimeter.

Charged particles will usually suffer the same fate in a calorimeter. The only charged particles which are usually not stopped in a calorimeter are muons, as they are not interacting hadronically and bremsstrahlung for them only becomes relevant at high energies (>1 TeV). The other almost unstoppable particle (within a sensible calorimeter dimension) is of course the neutrino.

Shower dimensions

As discussed above, the interactions responsible for the cascading process do have characteristic length scales, the radiation length X_0 for electromagnetic interactions, and the interaction length λ for hadronic interactions. It thus makes sense to express the longitudinal development of the shower in these units. Assuming that for large energies the probability for an interaction within one scale length l is a constant, the number of secondaries will grow exponentially as a function of the depth of the shower, $n(d) = (n_2 l)^{d/l}$, where n_2 is the average number of secondaries produced in one scale length. The shower development typically stops once the energy of the secondaries is insufficient to create further secondaries. If this happens at an energy scale E_f, and we assume that the energies are shared equally between the further secondaries, then

$$E_f = \frac{E_0}{n(d_f)} = \frac{E_0}{(n_2 l)^{d_f/l}},$$

where d_f is the total depth of the shower. Thus, the depth of the shower scales logarithmically with E_0.

Thus, to achieve shower containment in an experiment the size of the calorimeter needs to scale logarithmically with the scale of the energy of the experiment. To maintain the momentum resolution in a spectrometer the size of the spectrometer needs to scale with L^2. Thus calorimeter are powerful detectors in particular for high energy experiments.

A good rule of thumb is about $25X_0$ to contain 99% of a shower at ~ 100 GeV.

In electromagnetic showers the scale length is the radiation length X_0. Because the number of secondaries generated per X_0 is small (~ 1), it requires a significant number of X_0 to contain the shower. Laterally, the scattering angle of the secondaries in electromagnetic interactions is typically small (typically it's only in the tails of the shower that photons have a significant probability to be Compton-scattered by larger angles), so that the lateral shower dimensions are dominated by multiple scattering of the electrons and positrons in the shower. The scale length for this is called the Molière radius, which contains 90% of the deposited energy of a shower. From integrating eq. (3) one gets

$$R_{\mathrm{M}} \simeq \frac{21.2\,\mathrm{M}e\mathrm{V}}{E_c} X_0. \tag{15}$$

Note that the radius of containment is thus independent of energy.

Several versions of this equation using approximate expressions for E_c and X_0 are in use.

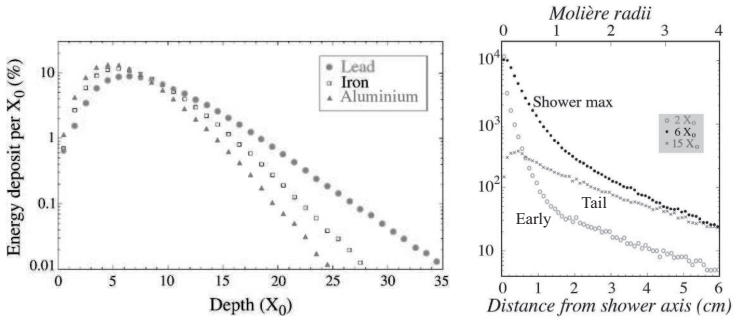


Figure 54: Electromagnetic shower dimensions (R.Wigmans, Sci. Act. 2, (2008) 18).

The scaling of the longitudinal profile with X_0 is not perfect. This is because of the different response for low-energy photon interactions (Compton scattering).

In hadronic showers a smaller number of interaction lengths λ is required. Because the scattering angle in hadronic interactions is larger, the transverse shower dimensions are larger than for electromagnetic showers, and the transverse full width at half (but not the non-Gaussian tails) maximum scales with the interaction length.

Containment of hadronic showers is also less important for reasons explained below.

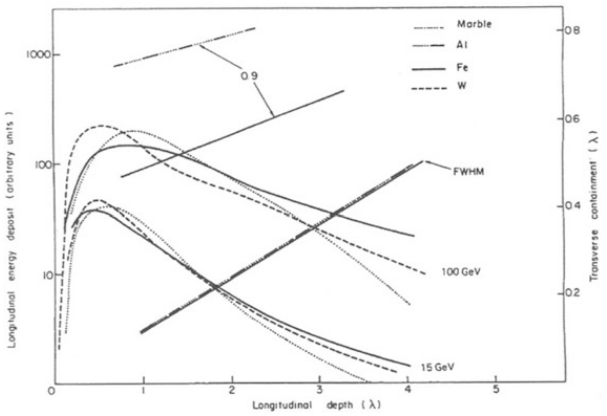


Figure 55: Hadronic shower dimensions (C.W.Fabjan, D.Fournier, Particle Physics Reference Library, https://doi.org/10.1007/978-3-030-35318-6_6).

In general purpose particle physics experiments there are usually two different calorimeters, an electromagnetic and a hadronic calorimeter. In terms of their interaction with the incoming particle there is no fundamental difference between the two. However, because of the differences of the characteristic lengths for dense solids, electromagnetic showers will be mostly contained in the first calorimeter section, whereas the hadronic showers need both for full containment. The first section, the electromagnetic calorimeter, is usually more finely instrumented to fully exploit the smaller transverse shower size and better energy resolution for electromagnetic showers.

For a short X_0 large Z is required, whereas for a short λ it is a smaller A (but high density). Hence, to separate electromagnetic and hadronic showers a large ratio of Z/A for the material in the electromagnetic calorimeter is desirable.

Energy resolution

In calorimeters the signal, which is proportional to the energy of the incoming particle, is the sum of a large number N of more or less independent small energy transfers (to create ionization charges, or scintillation photons, etc.). Because in all cases these energy transfer mechanisms do not transfer all the energy, and the amount of energy lost in the transfer varies, there is a stochastic error on the total signal reconstructed.

Assuming that the energy measurement is a simple counting of the particles created in the energy transfer, it is sensible to assume that this number is Poisson distributed, and thus

$$\frac{\sigma_E}{E} \propto \frac{\Delta N}{N} = \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}} \propto \frac{1}{\sqrt{E}},$$

and indeed a stochastic term $\propto E^{-1/2}$ is usually a major contribution to the energy resolution of a calorimeter. Other contributions to the stochastic term come from particles in the cascade escaping detection, because they traverse sections without instrumentation, or because of incomplete containment of the shower in the calorimeter.

Another reason to strive for complete shower containment is to prevent 'punchthrough', where the tail of a shower escapes the calorimeter and creates a signal in the muon detectors, which would fake a muon.

Generally, the energy resolution of a calorimeter can be parametrized as

$$\frac{\sigma_E}{E} = \frac{A}{\sqrt{E}} \oplus \frac{B}{E} \oplus C.$$

The first term contains the stochastic contribution, and is often (but not always) the most relevant. The second term, which corresponds to an energy-independent error typically is caused by readout noise, the final term is associated with inhomogeneities in materials, calibration errors and some forms of radiation damage.

In addition to the energy measurements calorimeter sometimes also need to provide information about the position of the shower (for track shower matching, or to reconstruct the 4-momentum of a neutral particle). In general, the position resolution has a similar energy dependence as the energy resolution, and in particular is often dominated by a $1/\sqrt{E}$ term caused by stochastic fluctuations. The scale of the position resolution is given by the transverse scale parameter (R_M for electromagnetic, and λ for hadronic showers), and the segmentation of the calorimeter, whatever is larger.

Finally, for some calorimeters also the time of the shower signal can be important, if it is needed to correlate the shower with a signal in another part of the detector. The time resolution in that case can usually also be described with a similar functional form.

Homogenous and sampling calorimeters

To minimize the depth of the material required to contain a shower, thick layers of a dense solid (usually a metal) are desirable as absorber. However, these cannot provide a signal as particle detectors. Hence, a common approach is to intersperse layers of dense absorber with less dense detector layers. This is called a 'sampling calorimeter'. Because of the size of the required instrumentation, the typical readout technologies are layers of scintillators, or ionization gaps filled with liquid noble gases. Wire chambers are less often used, because of the small amount of active material, and additional fluctuations in the internal avalanche amplification process. Calorimeters with semiconductor readout so far have been considered prohibitively expensive, but as will be discussed later there are benefits from the fine segmentation such readout can provide, which leads to their consideration in modern detector systems.

Although the obvious choice is to arrange these in layers, other geometries also work, as long as size of the sampling is smaller than the shower size.

The price to pay for the reduction of size are additional fluctuations between the energy deposited in the absorber, and in the detector layers, which increase the stochastic term in the energy resolution.. If we look at a simple electromagnetic 'counting' calorimeter, where the signal is obtained from the number of electrons or positrons crossing from the absorber into the detector layers, then the number of crossings is given by

These are called 'sampling fluctuations'

$$N = \frac{x_{\text{tot}}}{x} = \frac{E_0}{E_c}\frac{X_0}{x} = \frac{E_0}{\Delta E},$$

where x and x_{tot} are the thickness of one layer and the total track length, respectively, and ΔE the energy loss in one layer. The error on the energy measurement is now given by the error on N, which is again assumed to be Poisson distributed, so that

$$\frac{\sigma_E}{E} \propto \frac{1}{\sqrt{N}} = \sqrt{\frac{\Delta E}{E_0}}.$$

Typical stochastic terms for the energy resolution of electromagnetic sampling calorimeters are $\sim 10\%/\sqrt{E}$, whereas good homogenous calorimeters can achieve $\sim 1\%/\sqrt{E}$.

Hadron showers

While electromagnetic showers consist of a simple sequence of photons and electrons/positrons, handronic showers are more complex due to the large variety of outcomes of strong interaction reactions, and the subsequent hadronization.

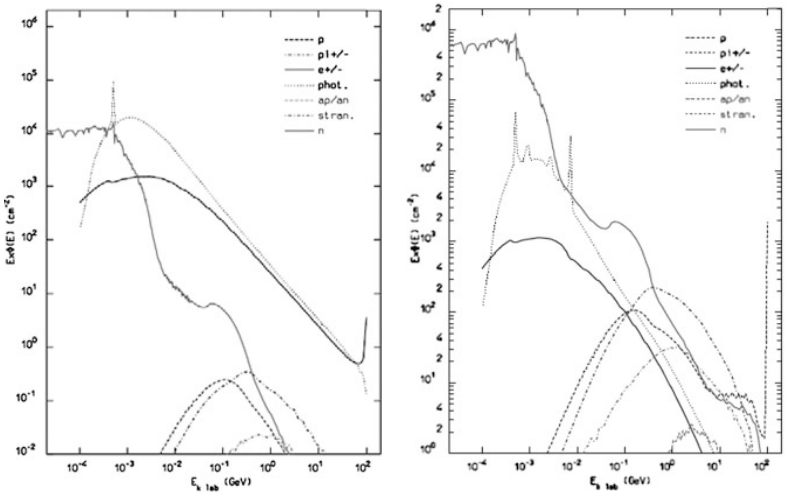


Figure 56: Simulation results of particle spectra produced in electromagnetic cascades initiated by 100 G*e*V electrons (left), and in hadronic cascades initiated by 100 G*e*V protons absorbed in lead (right). The y-axis shows the energy times the track length for a given energy (C.W.Fabjan, D.Fournier, Particle Physics Reference Library, https://doi.org/10.1007/978-3-030-35318-6_6).

The energetic components of the hadronic shower are dominated by charged pions, protons, neutrons, and photons that originate from decays of neutral pions. Due to the charge invariance of the strong interaction a third of the pions produced in the cascade are neutral pions, which decay quickly into two photons.

As these are typically high-energy photons they will start an electromagnetic shower component within the hadronic shower.

In addition to the creation of these secondaries, nuclei in the material will be excited. They will subsequently de-excite, often through spallation (emission of a number of neutrons and M*e*V photons) resulting in a

large component of low energy neutrons. The neutrons will be captured, but the relaxation times of the compound nuclei formed in the capture before the emission of a photon are long ($\sim \mu$s), and will this thus evade detection on the timescales of usual particle physics experiments. Hence, the signal measured for the hadronic component of the shower is smaller than for an electromagnetic shower of the same energy.

This is often described by the ratio e/h or e/π, where e is the response to an electromagnetic shower, and h or π the response to a hadronic shower.

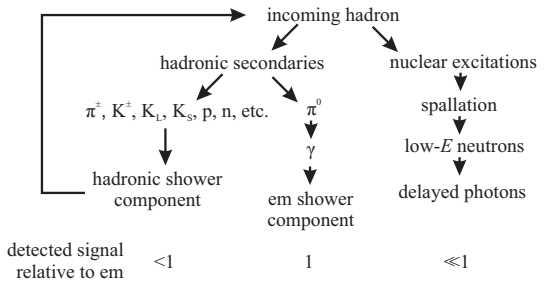


Figure 57: Components of a hadronic cascade.

Because of the stochastic nature of the outcomes of the hadronic interactions, the energy shares of the different components for individual showers will fluctuate. If the response ratio e/h differs strongly from 1, this will introduce large fluctuations of the recorded energy signal, and thus the energy resolution will be degraded.

Another consequence of $e/h \neq 1$ is that such a calorimeter is non-linear, as the em component is increasing for larger showers.

Because of the large amount of material needed to contain a hadronic shower, and because of the limits on the energy resolution imposed by the fluctuations between the different shower components, hadronic calorimeters are usually sampling calorimeters.

The goal in building a good hadronic calorimeter is to achieve $e/h = 1$. Several strategies are pursued to achieve this.

A calorimeter that achieves this condition is called a 'compensated' calorimeter.

- Reduce the electromagnetic response: Cross-sections for electromagnetic interactions are scaling with Z^2, thus a high-Z absorber will contain a larger fraction of the electromagnetic component, so that less of it develops in the detector layers. Interspersing additional thin low Z-absorber layers further reduces the electron component reaching the detector layers, because they will have a higher critical energy than the high-Z layers, and thus remove electrons with an energy between the critical energies of the two absorber materials.

- Improve the hadronic response: Detector layers with a large number of hydrogen atoms improve the energy response for the slow neutrons in the shower detector layers. The neutrons scatter from these protons, which due to the similar mass acquire significant recoil energy, which they deposit as ionization/scintillation signal contributing to the calorimeter response. Plastic scintillator (scintillating fibres) are well suited for this.

A drawback of achieving compensation using these two approaches is that the optimum sampling fraction for compensation is rather small. Thus the energy resolution for purely electromagnetic showers is strongly degraded.

- Software compensation: As the radiation length for dense absorbers is significantly smaller than the interaction length, the electromagnetic showers within the hadronic cascade will be fairly localized. In the reconstruction regions with a high energy deposition are de-weighted, so that the electromagnetic contribution to the reconstructed energy is more strongly penalized.

While uncompensated hadronic calorimeters can have energy resolutions of $100\%/\sqrt{E}$, the energy resolution of compensated calorimeters can be as low as $20\%/\sqrt{E}$.

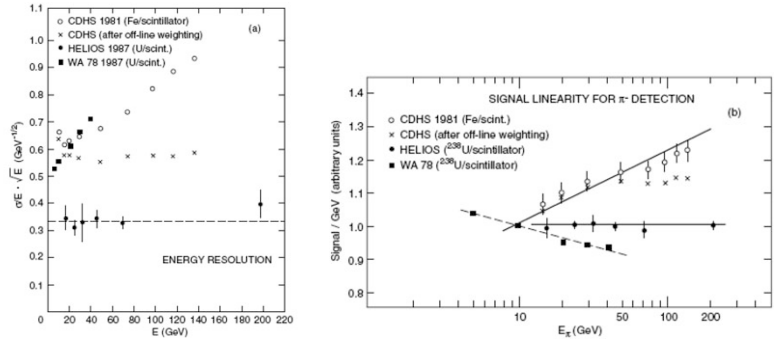


Figure 58: Energy resolution and signal linearity for different hadronic calorimeters (CDHS undercompensated, HELIOS compensated, WA78 overcompensated) (C.W.Fabjan, R.Wigmans, Rep. Prog. Phys. 52 (1989) 1519).

Particle flow analysis

Jets are the product of the hadronization of quarks and gluons produced in the hard scattering physics process. The particles generated during the hadronization propagate in a narrow cone, with the jet as a whole retaining the kinematics of the originating parton. Typically 60% of the particles in the jet are charged hadrons, 30% are photons and about 10% neutral hadrons. Because of the neutral component, and the challenges of detecting closely spaced charged tracks, measurement of jets in general purpose collider experiments have for a long time relied on the combination of electromagnetic and hadronic calorimeters. In that case the quality of the measurement of the jet energy was limited by the performance of the hadronic calorimeter.

Recent jet reconstruction algorithms strive to use the best measurement of the kinematic properties of the various particles in the jet. For charged hadrons this is usually the momentum measurement by the tracking system, for the photons the energy measurement in the electromagnetic calorimeter and only the measurement of the energy of the small neutral hadron component has to rely on the hadronic calorimeter. This approach can provide an improvement of the jet energy measurement by a factor 2 to 3. The challenge is to avoid double-counting, by identifying energy depositions in the calorimeter that are associated with charged tracks. This identification is facilitated by a fine calorimeter granularity, and the most ambitious new calorimeter project today aim for instrumentation with silicon detectors. The limiting factor of the jet energy measurement now becomes 'confusion', the misassignment of calorimeter energy deposits.

The tracking system also has to have good two-track separation capabilities.

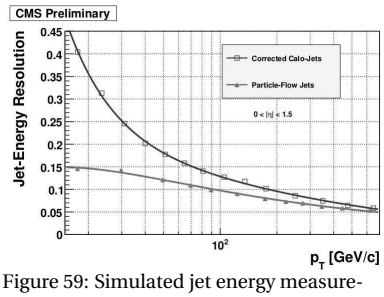


Figure 59: Simulated jet energy measurement performance in CMS (CMS PoS (ICHEP 2010) 002).

C4 Accelerator Lectures

Prof. Tony Weidberg

Michaelmas Term 2020

Table of contents

Why We Need Particle Accelerators?

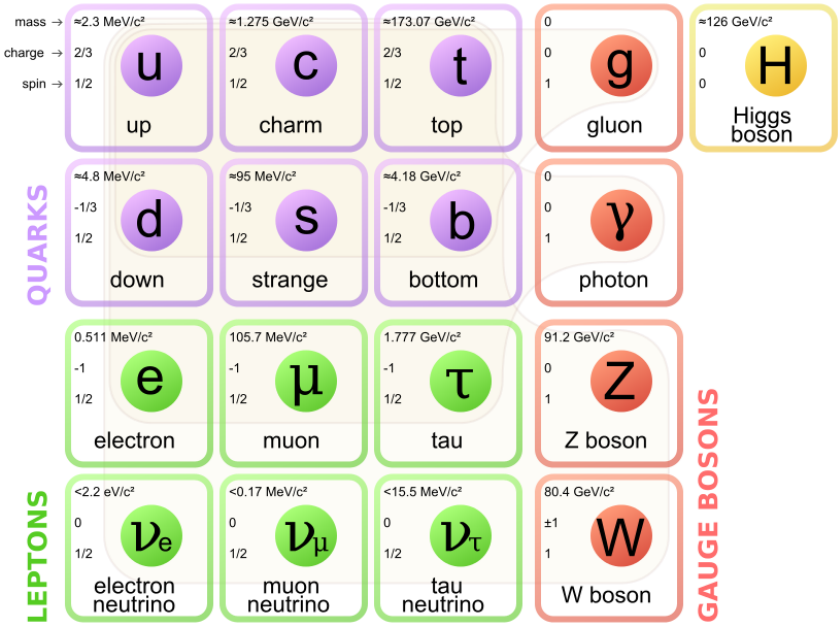
RF Acceleration

Magnetic bending and focusing

Luminosity and Colldiers

Why We Need Particle Accelerators?

- How much of the Standard Model would we know without particle accelerators?



Particle accelerators are essential tools for understanding nearly all particle physics

- Many other practical applications, e.g. ion implantation (Chips), radioactive isotopes for medicine, cancer therapy, synchrotron radiation.

Introduction

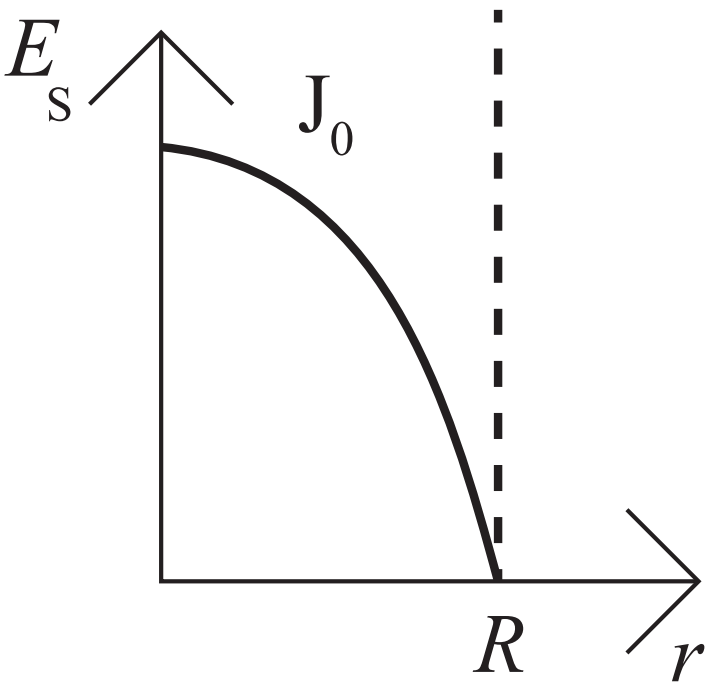
- Source of charged particles, e.g. protons (how?).
- Acceleration: Need E field (why not B field?)
- Bending of beams, e.g. for synchrotron, need magnetic dipoles.
- Beam focusing, need magnetic quadrupoles ("lenses").

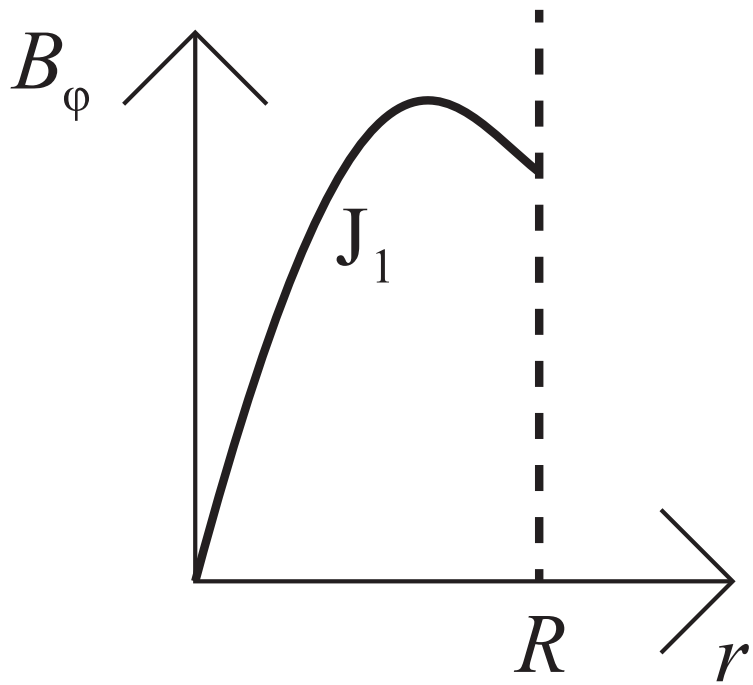
Acceleration -1

- $\mathbf{E} = -\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t}$
- DC fields no use for high energies (why?).
- RF E fields accelerate charged particles.
- In free space EM waves have only transverse E and B fields $\rightarrow$ no use for acceleration of charged particles!
- inside a cavity: different modes, TE and TM allowed.
- Need TM modes because these have longitudinal E fields.
- Solve Maxwell's equations subject to B.C. on surface of conductor: $E_{\parallel} = 0$ $B_{\perp} = 0$. Bessel functions.
- In general we can have TM_{nlm} modes. n, l, m refer to number of nodes in ϕ r s directions.
- need $E(r = 0) \neq 0$ and $E(r = R) = 0 \rightarrow l = 1$

Acceleration -2

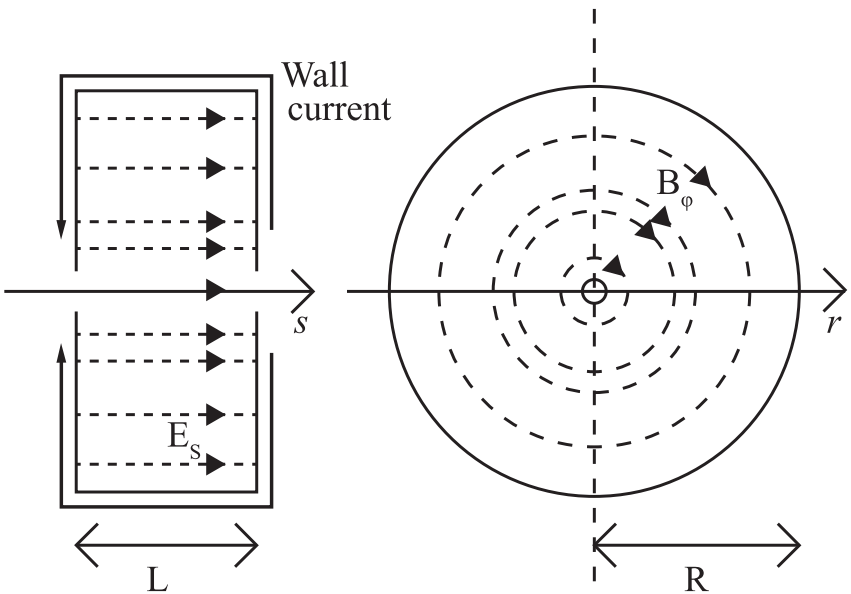
- How to select optimal modes?
- Use $n, l, m = 0, 1, 0$
- $E_s \sim J_0(kr)\exp(i\omega t)$, $\quad B_\phi \sim J_1(kr)\exp(i\omega t)$
- $J_0(kR) = 0 \;\rightarrow\; kR = 2.62$ sets allowed spatial frequency k.





RF caviteis -1

- Simplest: pill-box cavity
- Is this efficient in terms of energy required compared to acceleration achived?

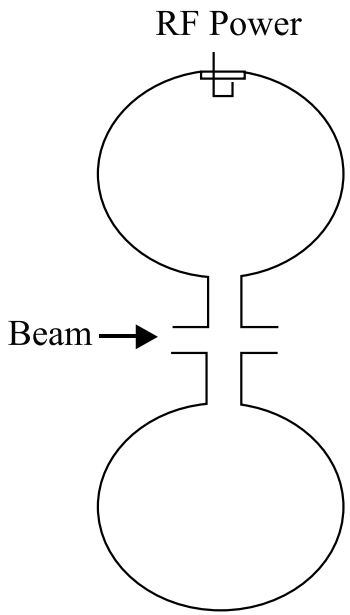


RF caviteis -2

- Define Q-value: $Q = U/\Delta W$
- Want highest Q-value possible to minimize electrical power dissipation for given acceleration.
- Energy stored in B field: $U = (\mu_0/2) \int \mathbf{H}_0^2 \, \mathrm{d}^3 r$.
- Power dissipated: $|\mathbf{H}| = j_s$ (surface current density). Current flows close to surface in good conductor
  $P = (1/2) \int \mathbf{H}_0^2 R_{\mathrm{surf}} \, \mathrm{d}s \quad R_{surf} = 1/(\sigma\delta)$ and $\delta = \sqrt{(2/\mu\sigma\omega)}$ is skin depth.
- Energy dissipated over one cycle: $\Delta W = \pi \sqrt{(\frac{\mu}{2\sigma\omega})} \int \mathbf{H}_0^2 \, \mathrm{d}s$.
- Need to maximize volume: surface area $\rightarrow$ large cavity.
- High frequency (typically 100 MHz - 1 GHz) and conductivity (superconductor?).

RF caviteis -3

- Also need to maximize useful acceleration
- $E_s = E_0 \cos(\omega t)$.
- Ultra-relativistic particle: $s = ct$.
- Energy gain: $\Delta W = \int_{-G/2}^{G/2} E_0 \cos(\omega z/c)\, dz = \frac{\sin(\omega G/2c)}{\omega G/2c}\, E_0\, G$.
- Efficient acceleration: $G \ll \lambda$.
- Very schematic RF cavity:



Real RF Cavities



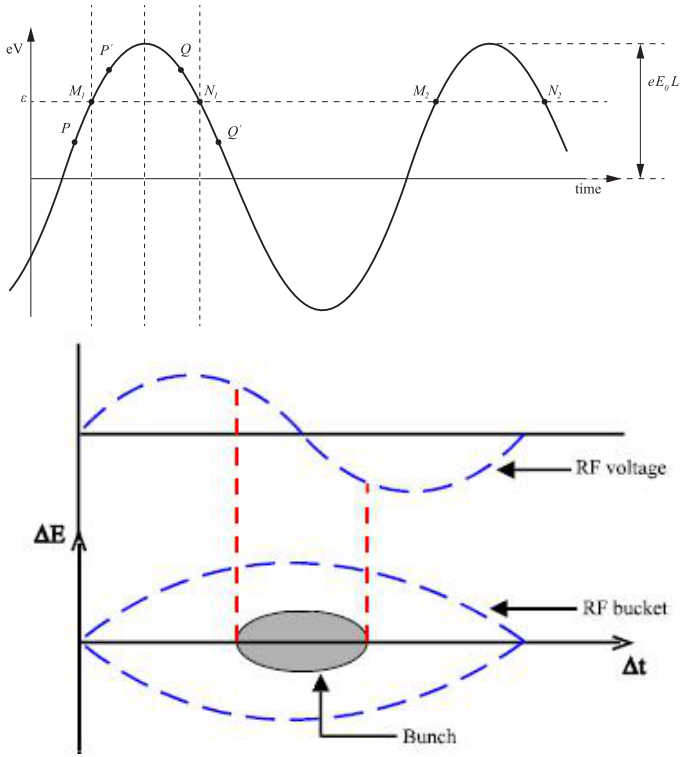


Phase Stability -1

- RF frequency $\omega = N\omega_{rev}$.
- e.g. LHC $f = 400$ MHz, number of RF buckets $N = 36,000$
- inject 2800 bunches into RF buckets
- Can potentially inject bunches of protons into every bucket (2.5 ns) but we use 25 ns spacing. Why?

Phase Stability -2

- What happens to particles not at the optimal phase wrt RF?
- Particle with nominal p injected at M_1 defines synchronous particle
- Particle arrives early at P $\rightarrow$ smaller energy gain
- Particle arrives late at P' $\rightarrow$ larger energy gain
- Particles oscillate in (E,t) about synchronous particle.



Phase Stability -3

- Two competing effects: (1) v increases with E; (2) R changes with p.
- $f = \frac{v}{2\pi R}$.
- $p = \frac{mv}{\sqrt{(1-v^2)}}$.
- Change in momentum $\rightarrow$ change in R.
- $\Delta R = D\frac{\Delta p}{p}$.
- Define $\eta_{\mathrm{rf}} = \frac{\Delta f/f}{\Delta p/p}$.
- $\eta_{\mathrm{rf}} = \frac{1}{\gamma^2} - \frac{D}{R_0}$
- Sign of η_{rf} flips as we increase p, need to flip phase of RF: transition.

Magnets

- Dipole magnets for bending beam.
- Quadrupoles for focusing beam (magnetic lenses).
- Warm iron dipole, maximum $\mathbf{B} = 2\mathrm{T}$.
- $p = qBR$ for $q = e$, $B = pR/0.3$ with p(GeV/c), B(T), R(m).
- Need larger B field for LHC $\rightarrow$ superconducting magnets.
- Use *NiTi* Type II superconductor cooled to 1.8K with superfluid He
- Increases maximum B field cf 4K and extremely good thermal conductor.

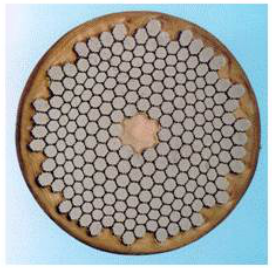
LHC Cable

- $6\mu m$ filaments embedded in Cu matrix (where does current flow?).
- 6300 filaments wound into strands.
- 36 strands make a cable
- Rutherford cable minimized residual magnetization. Used for all superconducting magnets, e.g. MRI.
- Approximate $\cos\phi$ distribution of current $\rightarrow$ dipole field at centre.

(a) filament



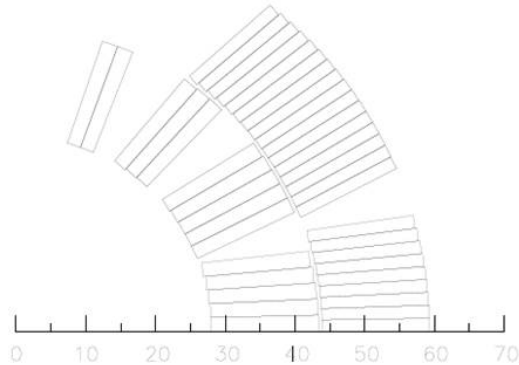
(b) strands



(c) cable

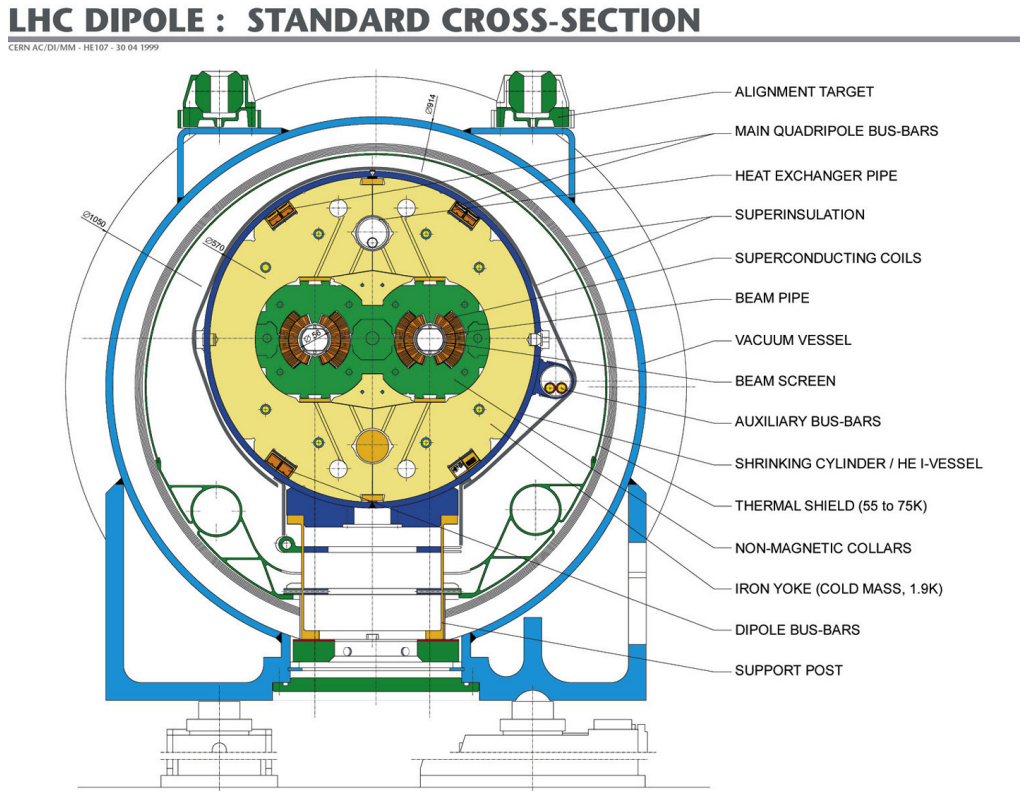


(d) dipole



LHC Magnet

- I= 11.3kA, B= 8.3T
- 2-in-1 design to fit in tunnel
- Supports force 400 tons/m.
- Energy stored in magnet 6.93 MJ.
- What happens if magnet warms up?

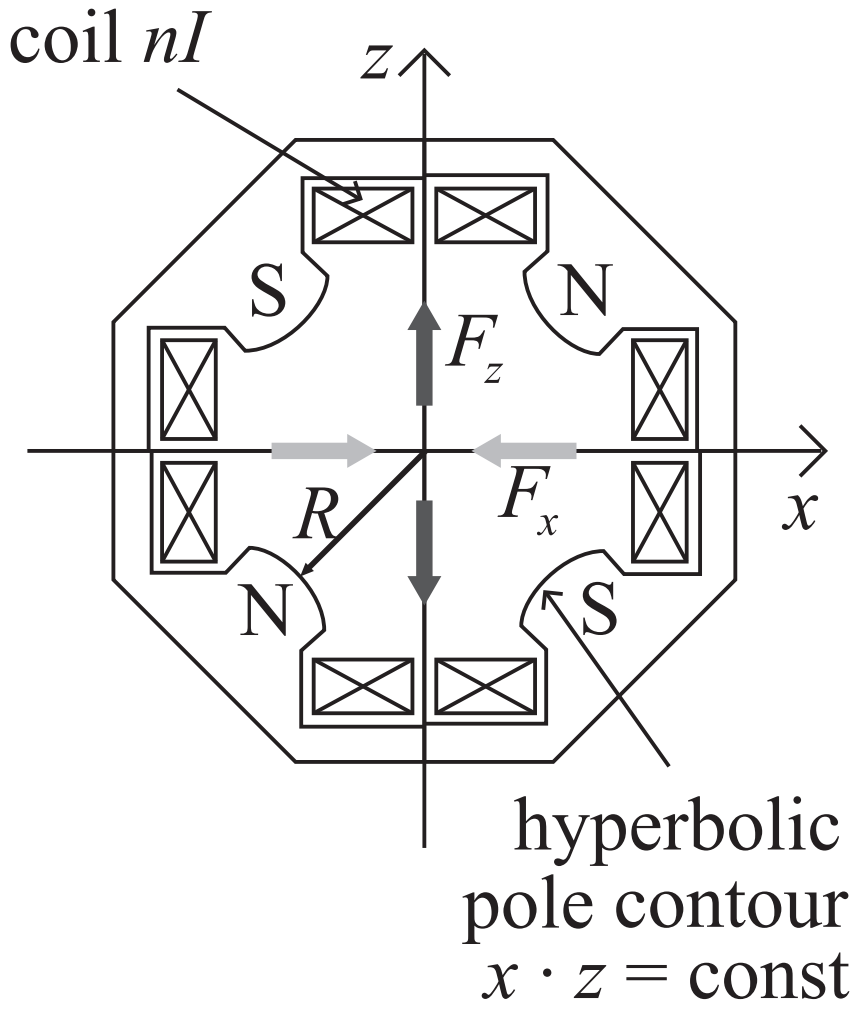


LHC Accident



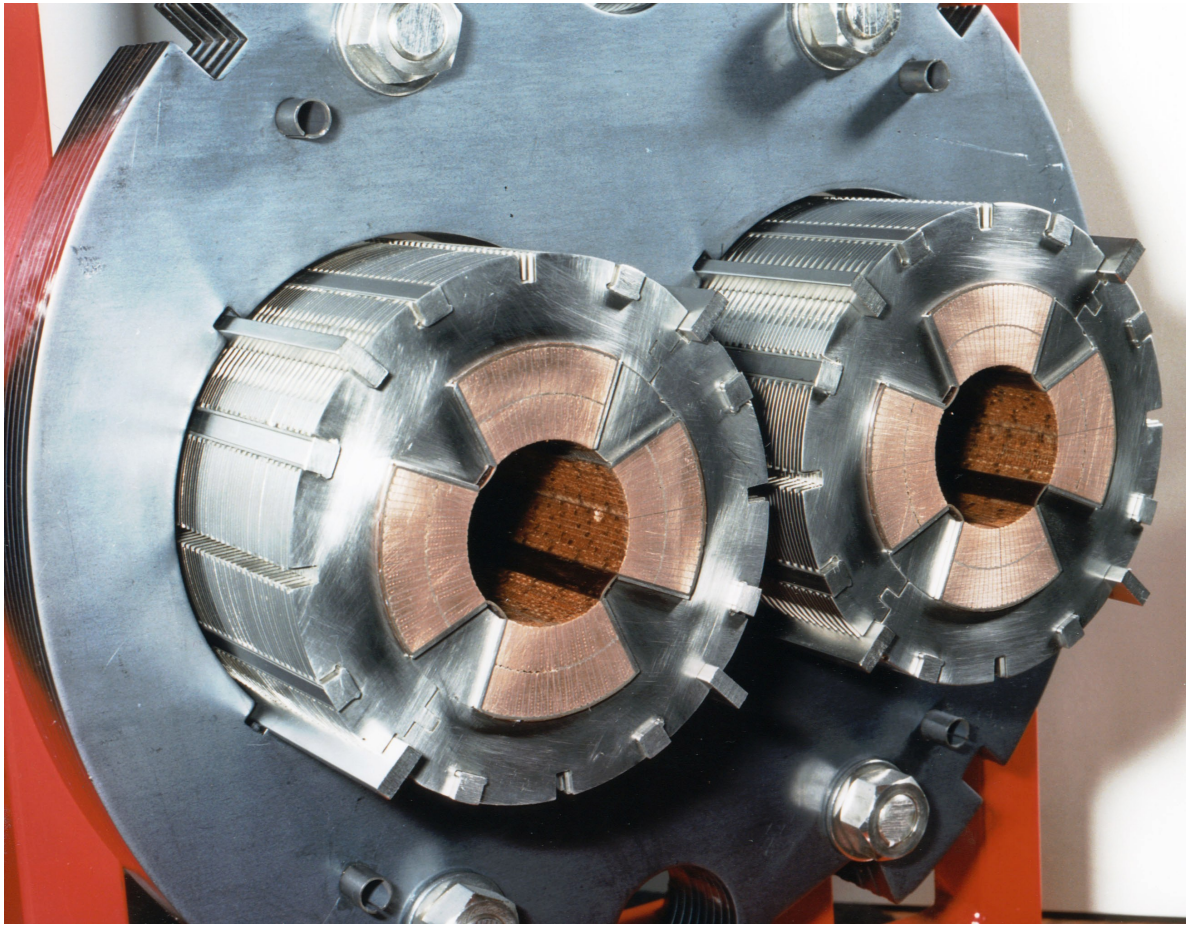
Quadrupoles

- $B_x = -gz$ and $B_z = -gx$
- $g = \frac{2\mu_0 nI}{R}$
- $F_x = qvB_z = -qvgx$ and
- $F_z = -qvB_x = qvgz$.
- focussing in x and defocussing in z



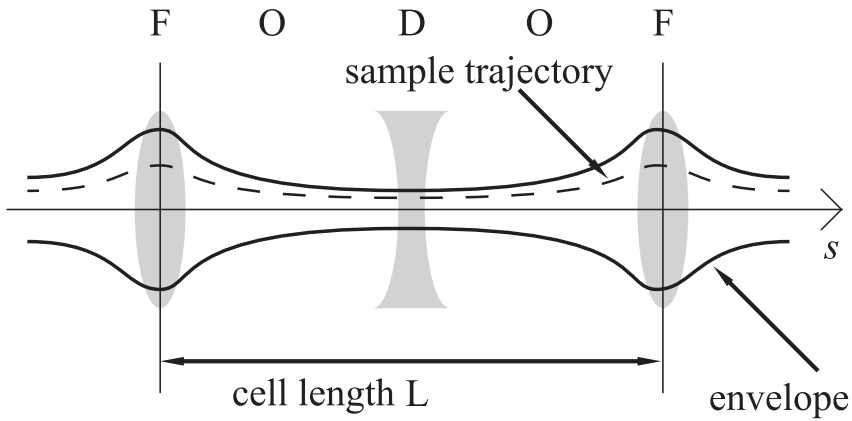
LHC Quadrupoles

- Field gradient $\frac{\mathrm{d}B}{\mathrm{d}x} = 233\ \mathrm{T/m}$.



Focusing

- Thin lens $1/f = qgl/p$
- Combine two thin lenses F and D (see problem set 4 Q8)
- $1/f = 1/f_1 + 1/f_2 + l/(f_1 f_2)$
- e.g $f_1 = -f_2 = f_Q$
- $\rightarrow f = f_Q^2/l$
- Strong focusing alternating series of FODO cells
- What should we place in the drift length l?



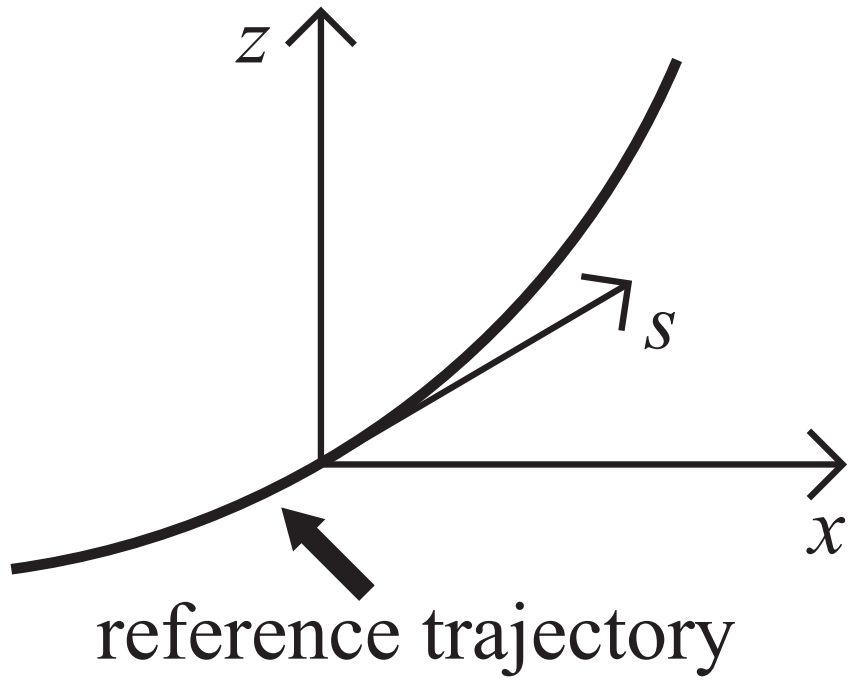
SPS

- Clearly see quads and dipoles



Beam Phase Space -1

- Magnetic force from focusing magnets
- Quadrupole strength: $k = \frac{1}{B\rho} g$
- At constant v, $p_z/p = z'$ transform equation of motion from (z, p_z) to (z, z').
- Hill's equation: $z'' + k(s)z = 0$
- Q1: What would happen if $k(s) = 0$?
- Q2: What if $k(s)$ constant?



Beam Phase Space -2

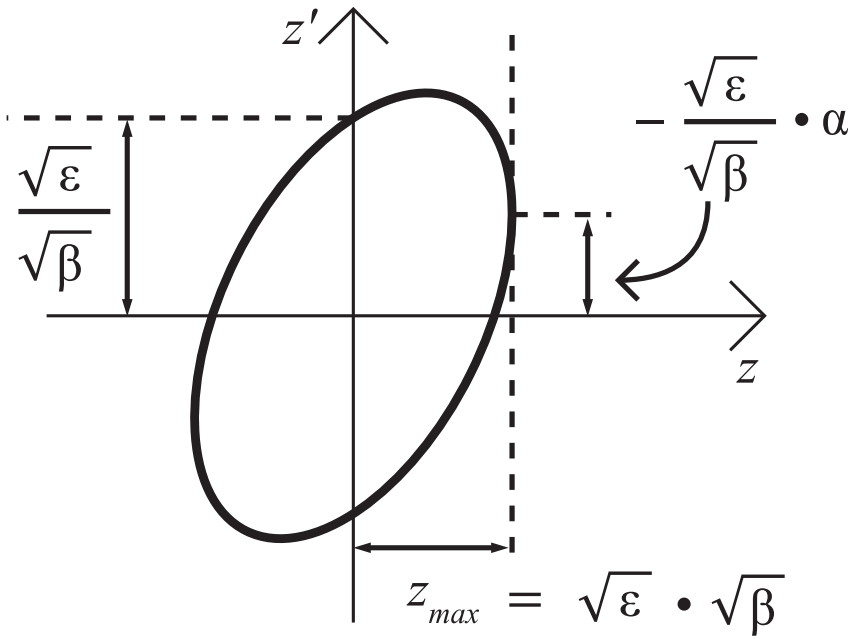
- Trial solution similar to SHM:
-

$$\begin{pmatrix} z \\ z' \end{pmatrix} = \begin{pmatrix} \sqrt{\epsilon}\sqrt{\beta(s)}\cos(\varphi(s)-\varphi_0) \\ -\frac{\sqrt{\epsilon}}{\sqrt{\beta(s)}}[\sin(\varphi(s)-\varphi_0)+\alpha(s)\cos(\varphi(s)-\varphi_0)] \end{pmatrix}. \tag{1}$$

- $\alpha(s) = -\beta'(s)/2$ and $\phi'(s) = 1/\beta(s)$; initial value ϵ and ϕ_0
- verify this: differentiate z(s) twice, sub $\rightarrow$ Hill's eqn, equate coefficients.
- $\beta(s)$ depends on strength of magnetic focusing.

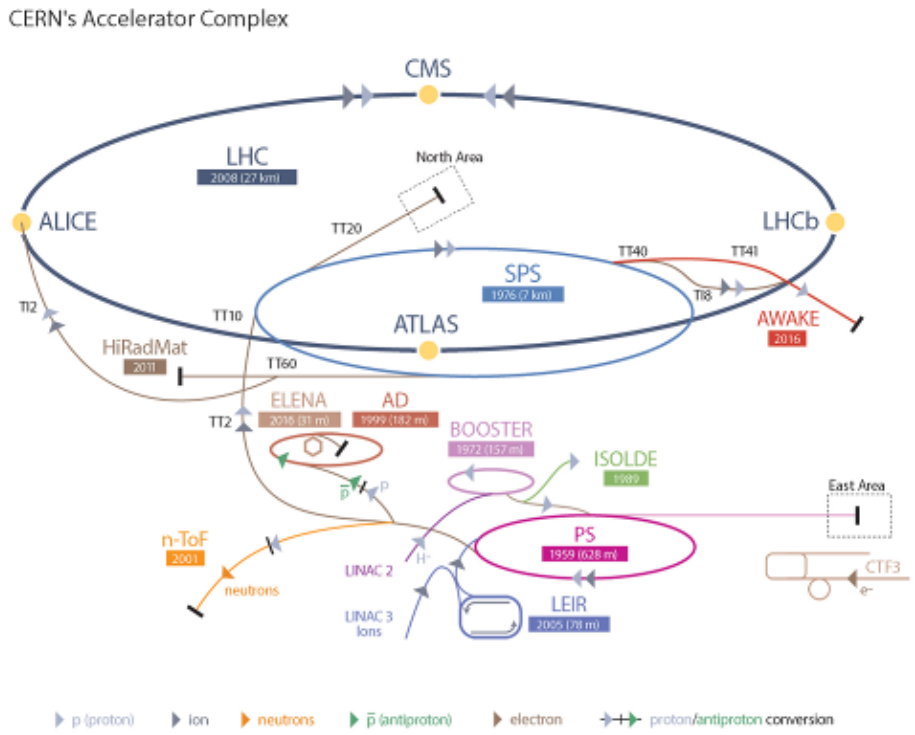
Beam Phase Space -3

- Particle follows ellipse in (z, z') space.
- Area $A = \pi\epsilon$ (see problem 5 Q1)
- At constant speed, area conserved.
- For particle bunch, define ϵ so that ellipse contains 90% beam.
- Focusing: decrease $z \rightarrow$ increase z'



Beam Phase Space -4

- Normalised emittance $\gamma\epsilon$ conserved $\rightarrow$ ϵ decreases as beam is accelerated.
- Q: Why do we need a long chain of accelerators?

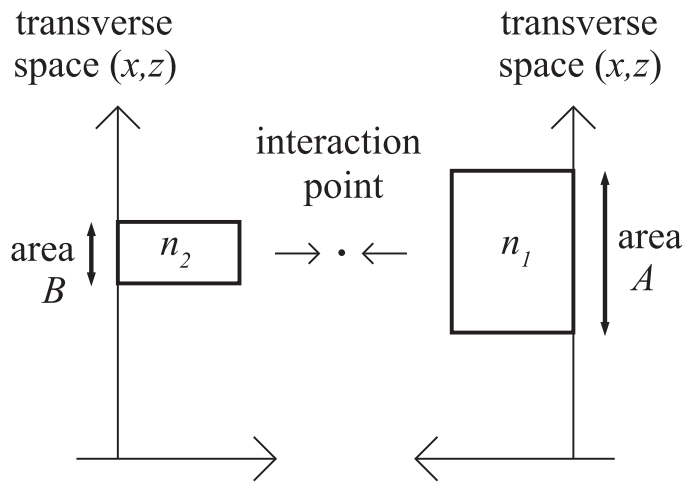


Q-values & Resonances

- Define tune shift $\int_s^{s+C} d\phi = 2\pi Q_z$
- similar definition for x-direction,
- What happens if Q_z is an integer?
- Avoid resonances $nQ_x + lQ_z = m$, where n, l and m are integers.

Luminosity-1

- Rate of reactions $R = \mathcal{L}\sigma$
- $\mathcal{L}$ as important as E for physics reach!
- Simple to get high $\mathcal{L}$ for fixed target geometry, just collide beam with block of matter.
- Highly non-trivial to get high $\mathcal{L}$ for a collider.
- Start with very simplified model.
- One LH beam particle, collision probability $p = (n_1/A)\sigma$
- n_2 particles in bunch in LH beam, b bunches and revolution frequency f.
- $\mathcal{L} = \frac{n_1 n_2}{A} bf.$

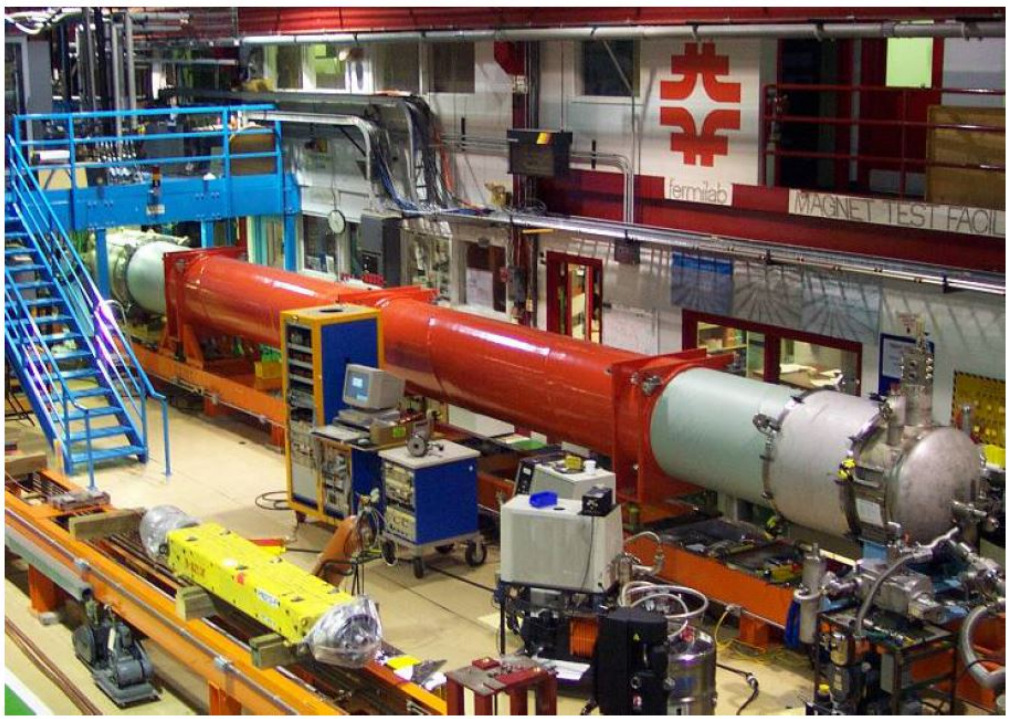


Luminosity-2

- More realistic: transverse density distributions ρ_1 and ρ_2
- $\mathcal{L} = bfn_1 n_2 \int_s \varrho_1 \varrho_2 dS$
- $I_1 = n_1 efb$
- $\mathcal{L} = \frac{I_1 I_2}{e^2 bf} \int_S \varrho_1 \varrho_2 dS$
- Assume Gaussian beam profiles, σ
- $\mathcal{L} = \frac{I_1 I_2}{e^2 bf} \frac{1}{2\pi(\sigma_1^2+\sigma_2^2)}$
- Assume $\epsilon_z = \epsilon_x$ and identical beams(simplicity)
- PS5 Q1 gives: $\sigma^2 = \epsilon\beta^\star$
- $\mathcal{L} = \frac{I_1 I_2}{4\pi e^2 bf} \frac{1}{\beta^* \epsilon}$.
- Q: How to optimise $\mathcal{L}$?

Low Beta Quadrupoles

- Quadrupoles placed either side of IP.
- $\beta^\star = 0.5\ \mathrm{m}$ cf $\beta = 100\ \mathrm{m}$ in arcs.
- Transverse beam size at IP $\sigma = 16\mu\mathrm{m}$
- What happens to β inside quadrupoles? Injection?



How to Measure Luminosity

- In e^+e^- pick known cross section and measure rate.
- Bhaba scattering $e^+e^- \to e^+e^-$ at small angles.
- Can't do this for hadron colliders
- Van der Meer Scan: see problem 5 Q3.

Synchrotron Radiation

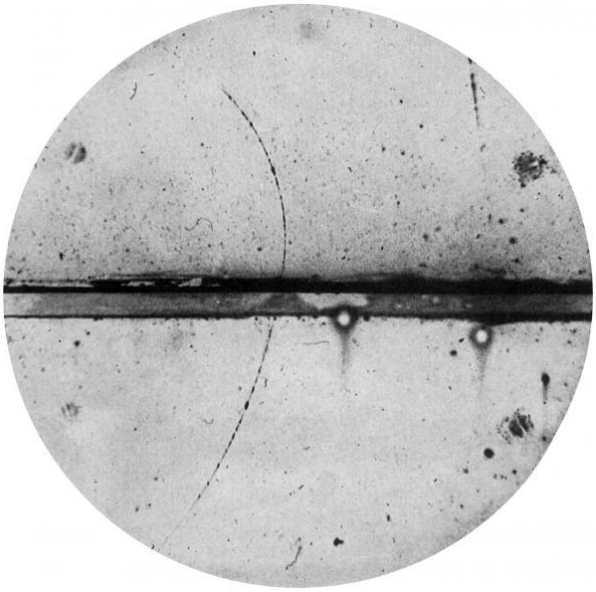
- Power radiated by particle mass m, Lorentz factor γ in ring of radius ρ
- $P = \frac{2 r_c m \gamma^4}{\rho^2}$
- Very important limitation for electron accelerators
- Need very large ρ expensive!
- Use linear accelerator to avoid this?
- Use synchrotron radiartion as very bright X-ray source, e.g. Diamond light source at RAL.

Accelerator Summary

- RF cavities for acceleration.
- Phase stability
- Dipole magnets for bending and quadrupole for focusing.
- superconducting magnets
- Strong focusing FODO cells
- Luminosity and how to optimise it.

Further reading accelerators

- Chapter 3 *Particle Physics in the LHC Era, OUP*.
- E. Wilson, *An introduction to Particle Accelerators*, OUP, 2001. A very good introduction to the subject.
- P. Bryant and K. Johnsen, *The Principles of Circular Accelerators and Storage Rings*, Cambridge University Press, 1993.
- K. Wilie, *The Physics of Particle Accelerators, an introduction*, Oxford, 2000. A more advanced introduction to the subject.



The Oxford MPhys Course

C4: Particle Physics

Problem Set Four

1. The resolution for an electromagnetic (EM) calorimeter is often parametrised in the form

$$\frac{\Delta E}{E} = \frac{a}{\sqrt{E}} \oplus c.$$

(a) Describe the origin of the two terms in this equation. A fuller model also includes a term of the form b/E due to electronic noise in the readout. Give a simple justification for the form of this term.

(b) How could an EM calorimeter be calibrated in-situ at the LHC (i.e. using data from the LHC collisions)?

(c) The resolution of a hadronic calorimeter is generally worse than that of an EM calorimeter. What is the physical cause of this effect?

(d) How could the energy scale for hadronic jets be calibrated in an LHC detector?

2. This question is about the basic principles of a silicon detector.

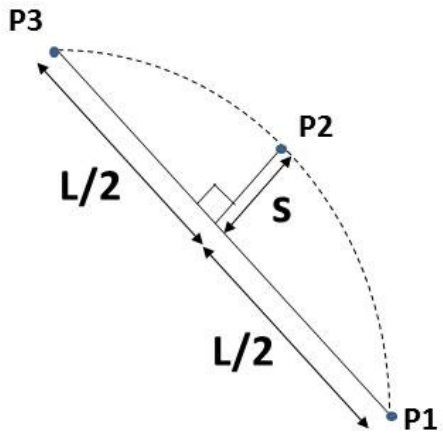
(a) Give a simple explanation of how a signal is generated in a silicon micro-strip detector when it is traversed by a fast charged particle. Why is it necessary to apply a large electric field across the detector?

©University of Oxford 2021

(b) Consider such a detector made with p-type strips in n-type bulk silicon. The thickness of the bulk is 300 μm, and the density of free carriers in the bulk is 10^{18} m^{-3}. By solving Poisson's equation in one-dimension, evaluate the magnitude of the bias voltage that should be applied to remove the free carriers from the bulk ("fully deplete" the detector).

(c) The detector is divided into detecting elements "strips" which are 100 μm wide . Assuming that a charged particle only creates a "hit" in one strip, determine the spatial resolution of the detector. (hint: first show that the Root Mean Square (RMS) of a uniform distribution between 0 and 1 is $1/\sqrt{12}$. Then assume that particles cross the detector uniformly.) Show that the radius of curvature R in the plane perpendicular to a uniform magnetic field B, for a relativistic charged particle with electric charge q and momentum in the same plane p_T, is given by $R = p_T/Bq$. (hint: first show that the magnitude of the momentum is constant and then use $F = \frac{\mathrm{d}p}{\mathrm{d}t}$).

(d) A simplified model for a silicon tracking detector assumes that the trajectories of charged particles are measured at three points (P1, P2 and P3) as shown in the figure.



Determine an expression for R in terms of L and the Sagitta s. You may assume that $L \gg s$. Assuming that the dominant contribution to the uncertainty comes from the measurement of the point P2, determine an expression for the resolution in $1/p_T$. Evaluate the resolution for a particle with unit charge, assuming, $B = 2$ T and $L = 0{\cdot}6$ m, using the spatial resolution you calculated above. Express you answer in units of $(\mathrm{GeV/c})^{-1}$. What is the effect of multiple scattering on the momentum resolution? Sketch a graph of resolution versus momentum to illustrate the contribution from multiple scattering and measurement precision.

3. Super-Kamiokande is a 50,000 tonne water Čerenkov detector, located in the Kamioka mine in Japan. It has a cylindrical geometry, with an inner wall covered with 11,200 photomultiplier tubes and an outer volume designed to veto entering cosmic ray muons.

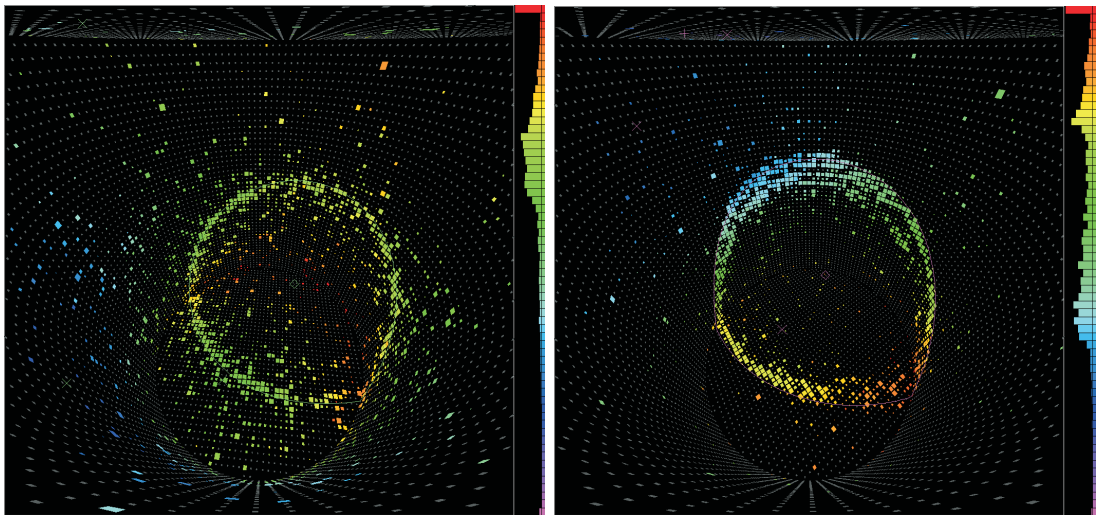
(a) What is Čerenkov radiation? Calculate the direction of Čerenkov radiation with respect to the direction of motion of fast charged particles in water, if the refractive index of water is 1.33.

(b) Calculate the threshold energy above which electrons and muons emit Čerenkov radiation in water. What consequences does this have for the measurement of

(i) *the solar neutrino flux?*

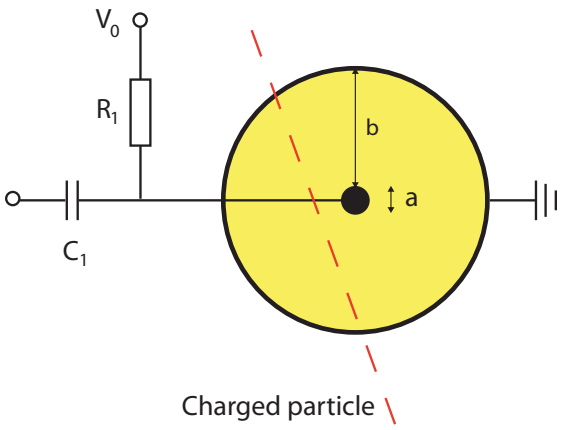
(ii) *the flavour ratio of atmospheric neutrinos?*

(c) The figures show two events recorded in the Super-Kamiokand detector. Which of the events is more likely to be a muon event and which an electron event? Explain your choices. Given that the time resolution of the photomultiplier tubes is $\mathcal{O}(1)$ ns, which additional property of muon events can aid their identification?



(d) How can particle energies be determined in a detector such as Super-Kamiokande? Why is it important to know the direction of the incident neutrino and how might that be determined from the event?

4. This question explores some of the basic electrostatics of proportional and drift chambers used for charged particle tracking. A proportional chamber is one in which the signal collected is proportional to the ionisation energy loss of the charged particle. A drift chamber exploits fast timing with a knowledge of the drift velocity in the chamber gas to improve the position resolution of interactions.



(a) Explain briefly the operating principle of this chamber.

(b) A particular chamber configuration has an anode wire radius of $a = 10$ μm (radius not diameter - diagram is not accurate!) and a potential on the wire of $V_0 = +3$ kV. The cathode is a cylinder of radius $b = 10$ mm at zero potential, concentric with the wire. Assume that the avalanching occurs at a distance $\delta = 1$ μm from the anode. Show, using Ramo's theorem, that the induced charge signal from a charge Q displaced through a distance dr in the chamber is,

$$dq = \frac{Q}{V_0}\frac{dV}{dr}dr,$$

Calculate the relative contributions to the charge pulse from the proportional chamber by electrons moving to the wire, and by the positive ions moving away. Show that the contribution of the electrons to the signal is negligible.

(c) The velocity of the ions is a linear function of the electric field, E, in the tube:

$$\frac{dr}{dt} = \mu E.$$

The coefficient μ is equal to 1.7 cm^2 s^{-1} V^{-1} for this chamber. Using this relation and the expression for dq given above, show that the signal current for the positive ions (as a function of time) is

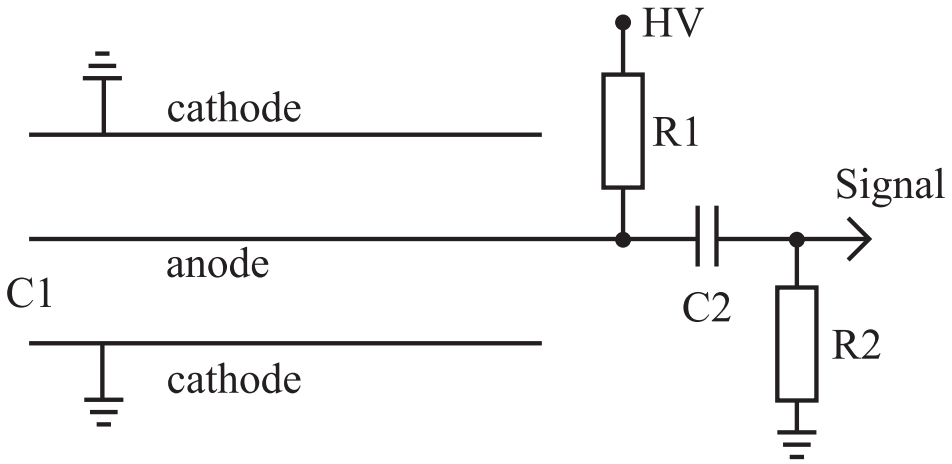
$$i(t) = \frac{QC}{4\pi\epsilon_0}\frac{1}{t + \frac{\pi\epsilon_0(a+\delta)^2}{\mu V_0 C}}$$

and find the total drift time (T) of the ions. Sketch the time development of the current pulse (from ions) due to the passage of this particle. Show further that the total charge (from the positive ions) collected (after T) is:

$$Q\frac{\ln\left(\frac{b}{a+\delta}\right)}{\ln\left(\frac{b}{a}\right)}.$$

(d) Calculate the time after which 'enough' charge is collected, i.e. when half of the total charge is available.

(e) **optional, not marked.** A blocking capacitor decouples the high voltage (connected to the wire via resistor R_1) from readout electronics. This introduces a time constant $\tau_1 = R_1 C_1$. What is a good choice for τ_1 and why? Further, the chamber has electric capacitance C_p. How does this affect your choice of capacitance C_1?
Hints: draw an equivalent circuit diagram with the signal current represented by a current source. Connect R_1 to ground (d.c. voltages are irrelevant for this type of signal analysis). Assume also that the input to the pre-amplifier is at ground potential. Calculate the signal current into the pre-amplifier, applying circuit analysis using complex impedances.
A further time constant $\tau_2 = R_2 C_2$ arises from the readout electronics. What value of τ_2 would you choose and why? Does mode of operation, drift or proportional, influence the design?



5. This question focusses on bremsstrahlung.

(a) Describe the main features of bremsstrahlung and pair production and indicate how they depend upon the atomic number of the material in which they occur.

(b) A very crude model of the initial development of an electromagnetic shower is that a high energy electron or positron of energy E_0 undergoes a bremsstrahlung process after a distance L (one radiation length) and loses half of its energy to a secondary photon, or that a high energy photon initiates a pair production process after travelling a distance L, splitting its energy equally between the two secondary particles. These processes continue until the photons and charged particles each have an energy less than the critical energy $E_C (\ll E_0)$, at which point the multiplication ceases. Develop this model and answer the following questions for both an incident electron of energy E_0 and for an incident photon of the same energy.

(i) How many photons plus charged particles will there be after N radiation lengths?

(ii) What is the energy of each particle in the shower after N radiation lengths?

(iii) In this model, the fraction of secondaries that are photons after N radiation lengths tends to a constant when N is large. What is this fraction?

(iv) What is the depth (in units of L) at which the number of particles in the shower is a maximum, and what is the number of particles at maximum?

(c) Compute the depth and the number of particles when multiplication ceases for a 4 GeV electron entering lead glass ($L = 2{\cdot}5$ cm, $E_C = 10$ MeV).

(d) The total energy deposited by the shower in the glass may be assessed by measuring the number of photons emitted as Čerenkov radiation. Indicate why this Čerenkov photon number is nearly proportional to the total shower energy. Assume all charged particles in the shower emit equal amounts of Čerenkov radiation.

6. Consider a magnetic flux density $\mathbf{B}$ of the form

$$\begin{aligned} B_x &= A_1\, y \\ B_y &= A_2\, x \\ B_z &= 0, \end{aligned}$$

where A_1 and A_2 are constants.

(a) Show that this field is fully compatible with Maxwell's equations if and only if $A_1 = A_2$.

(b) A particle is initially moving in the plane $y = 0$, parallel to the z-axis, at a position $x = x_0$. It enters a region of magnetic field specified by the equation above, which exists over a short length l and can be neglected outside this region. Assuming that the change in the particle's x-position, while it traverses the field region can be neglected, calculate the angular deflection the particle experiences by traversing this region. Hence show that a beam of particles with a range of values of x_0 will be brought to a focus and determine the focal length f. Show that in the y-direction, the quadrupole acts like a defocussing lens with a focal length of $-f$.

(c) We can represent the horizontal motion of a beam in an accelerator in terms of x, the distance from the central orbit, and x' the angle it makes to the horizontal. We can write the transfer function from one location to another as a matrix equation

$$\begin{pmatrix} x_2 \\ x_2' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_1' \end{pmatrix}$$

Write down the matrices for

(i) F, a thin focussing lens in the x-direction, focal length f;

(ii) O, a drift space of length L (no forces act on the particles in this region);

(iii) D, a a thin de-focussing lens in the x-direction, focal length $-f$.

(d) The overall transfer matrix for a combination of focussing quadrupole, drift space and a defocussing quadrupole is given by matrix multiplication. $T = ODOF$ (note that the 'F' matrix on the RHS represents the first element the beam sees). Take the focal length of the focussing quadrupole to be f and the defocussing quadrupole to be $-f$. Let the length of the drift space between the F and D quadrupoles be L and consider a drift length s_h for the second O matrix. Show that

$$T = \begin{pmatrix} 1 - L/f - s_h L/f^2 & L + s_h L/f + s_h \\ -L/f^2 & L/f + 1 \end{pmatrix}.$$

(e) For the overall system to be focussing we require that for any initial beam co-ordinate x we can bring it to $x = 0$ after some distance s_h, with s_h positive.

$$T \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ x' \end{pmatrix}.$$

Show that one can achieve $x_1 = 0$ for positive values of s_h (i.e. the combined system is focussing) if $L < f$. In the orthogonal plane, the sign of f is reversed. Show that the system is always focussing in this plane.

7. Synchrotrons have a periodic ring-shaped lattice of focusing and bending magnets. Specifying the position of a beam particle on the circumference of the accelerator by the distance s, the equations of motion in both horizontal and vertical directions about the ideal orbit have the form

$$\frac{d^2x}{ds^2} + K(s)x = 0, \quad \text{where } K(s+L) = K(s) \text{ is periodic.}$$

The solution is a quasi-periodic function of the form

$$x(s) = \sqrt{\varepsilon\beta(s)}\cos(\phi(s) - \phi_0), \quad \text{where } \varepsilon \text{ and } \phi_0 \text{ are constants.}$$

Both ε and $\beta(s)$ have dimensions of length. The phase $\phi(s)$ is related to the beta function $\beta(s)$ by $d\phi/ds = 1/\beta$.

(a) Explain the significance of the beta function.

(b) Show that x and $x' \equiv dx/ds$ satisfy,

$$\frac{x^2 + (\beta(s)x' + \alpha(s)x)^2}{\beta(s)} = \varepsilon, \quad \text{where } \alpha(s) = -\frac{1}{2}\frac{d\beta}{ds}.$$

(c) Reduce the expression to the standard form for an ellipse in $x - x'$ space and show that its area is $\pi\varepsilon$. This is the emittance of the beam (in one transverse direction). Explain its importance to the design and operation of an accelerator.

(d) Explain how the emittance is changed as the beam is accelerated. Hence explain why we need a chain of pre-accelerators to inject beam into a high energy accelerator such as the LHC.

©University of Oxford 2021

8. A question exploring accelerators.

(a) Explain the terms *luminosity* and *integrated luminosity*. Describe a method to measure the luminosity in an e^+e^- collider.

(b) A collider with two beams of unequal width has 3×10^{11} particles per bunch, two bunches in each beam and the frequency of revolution is 1 MHz. Particles are uniformly distributed in cylindrical bunches which move parallel to the cylinder axis. The radius of each bunch in the wider beam is 4×10^{-4} m. Calculate the luminosity.

(c) A symmetric collider has equal horizontal and vertical emittances and equal horizontal and vertical β functions. Both beams have identical parameters. Explain why and by how much the luminosity would change when all β functions become 3 times smaller at the interaction point.

(d) The energy radiated per turn, E_0, by synchrotron radiation for an electron of energy E moving in an orbit with radius of curvature ρ is given by $E_0 = \frac{4}{3}\pi \frac{r_e}{(m_e c^2)^3} \frac{E^4}{\rho}$, where $r_e = 2{\cdot}8 \times 10^{-15}$ m is the classical radius of the electron.

Two designs for a future electron-positron collider with a centre-of-mass energy of 500 GeV are being considered. Superconducting cavities with an accelerating gradient of 35 MV/m will be used.

First consider a circular collider with one 2·1 m-long superconducting cavity to accelerate the particles. How many turns do the particles need to reach the nominal energy (to make the calculation easier, ignore the energy loss due to synchrotron radiation)? Estimate the circumference of the collider so that the energy lost per turn due to synchrotron radiation is 10% of the beam energy at the nominal energy.

(e) Now consider linear collider where the particles are accelerated from 0 to 250 GeV by several 2·1 m-long cavities. How many accelerating modules would be required? How long would such an accelerator be, assuming that the particles collide head-on at the interaction point?

9. This question is about the measurement of luminosity at a pp collider such as the LHC using Van der Meer(VDM) scans. Consider a collider with a revolution frequency f_{rev} and n_b colliding bunches. Let the number of particles in beam 1(2) be $n_1(n_2)$ and let the normalised bunch density be $\rho_1(x, y)$. The luminosity is defined by the interaction rate for a process with a given cross section σ by $R = L\sigma$. In terms of the beam parameters, the luminosity is expressed by,

$$L = n_b f_{rev} n_1 n_2 \int \rho_1(x, y)\rho_2(x, y)dxdy.$$

(a) Explain the origin of this equation.

(b) Assume that the particle densities are uncorrelated in the x and y directions. The luminosity can then be written as

$$L = n_b f_{rev} n_1 n_2 \Omega_x \Omega_y,$$

where the beam overlap is defined in x (with an equivalent definition in y) as

$$\Omega_x = \int \rho_1(x)\rho_2(x)dx$$

Assume that the beams have Gaussian distributions with a common centre and an RMS of σ_{x1} (σ_{x2}). Show that

$$\Omega_x = \frac{1}{\sqrt{2\pi}(\sigma_{x1}^2 + \sigma_{x2}^2)^{1/2}}$$

Note that σ refers to the cross section and σ_{x1}, σ_{x2} refer to Gaussian widths of the beams.

(c) Let $R(x)$ be the interaction rate as a function of the separation of the beams in the x direction. $R(x) \propto \exp\frac{-x^2}{2(\sigma_{x1}^2+\sigma_{x2}^2)}$. The values of $(\sigma_{x1}^2 + \sigma_{x2}^2)$ can be determined from a VDM scan if the interaction rate for some particular process $R(x)$ is measured as the separation between the two beams x is varied. Assume that the fraction of bunch crossings which register a hit in a counter is p, show that the mean number of collisions which produce such hits is given by $\mu = -\ln(1-p)$.

(d) Why is it advantageous to use a small counter? What is the problem with this technique at very high luminosity? Suggest a possible detector technique that could be used for such a counter?

10. This question is also about Hill's equation which was considered in the earlier question.

$$\frac{d^2x}{ds^2} + K(s)x = 0, \quad \text{where } K(s+L) = K(s) \text{ is periodic.}$$

We can write the trial solution as,

$$x(s) = \sqrt{\varepsilon\beta(s)}\cos(\phi(s) - \phi_0), \quad \text{where } \varepsilon \text{ and } \phi_0 \text{ are constants.}$$

To simplify the algebra, let $u^2(s) = \beta$.

(a) Show that the trial solution is a valid solution to Hill's equation and hence show that

$$u'' - 1/u^3 + K(s)u = 0. \tag{1}$$

(b) How can we use these results to determine the Q-value? Discuss the significance of these results.