Stefaan Tavernier

Experimental Techniques in Nuclear and Particle Physics

Springer

Experimental Techniques in Nuclear
and Particle Physics

Stefaan Tavernier

Experimental Techniques in Nuclear and Particle Physics

Springer

Prof. Stefaan Tavernier
Vrije Universiteit Brussel
Fak. Wetenschappen
Dept. Natuurkunde (DNTK)
Pleinlaan 2
1050 Bruxelles
Belgium
stefaan.tavernier@vub.ac.be

Preface

I have been teaching courses on experimental techniques in nuclear and particle physics to master students in physics and in engineering for many years. This book grew out of the lecture notes I made for these students. The physics and engineering students have rather different expectations of what such a course should be like. I hope that I have nevertheless managed to write a book that can satisfy the needs of these different target audiences. The lectures themselves, of course, need to be adapted to the needs of each group of students. An engineering student will not question a statement like "the velocity of the electrons in atoms is ≈1% of the velocity of light", a physics student will. Regarding units, I have written factors h and c explicitly in all equations throughout the book. For physics students it would be preferable to use the convention that is common in physics and omit these constants in the equations, but that would probably be confusing for the engineering students.

Physics students tend to be more interested in theoretical physics courses. However, physics is an experimental science and physics students should understand how experiments work, and be able to make experiments work. As a post doc, I have never designed any electronics board, but many times I have had to find out why the board I have given did not do what it was supposed to do and fix the problem. This is an essential skill any experimental physicist should have. I hope this book will help the students in acquiring this skill and provide her or him with a sufficient basic knowledge on nuclear and particle detection techniques such that she or he is able to read, and understand, the scientific literature in this field.

Brussels, Belgium
Stefaan Tavernier

Contents

Chapter 1
Introduction

1.1 Documentation

The present textbook is an introduction to measurement techniques in nuclear and
particle physics and to particle accelerators. This subject is a part of the standard cur-
riculum in Physics, Nuclear Engineering and Medical Physics. The course should
preferably be taken together with a more theoretically oriented course on nuclear
physics or particle physics. References [1–3] are a few of the many good hand-
books providing this. However, the book can be used as a stand-alone textbook on
the experimental aspects of nuclear and particle physics. The only pre-requisite is
the general mathematics and physics background that is required for all students in
Physics and Engineering.

   The emphasis in this textbook is on the principles of operation and the basic
characteristics of measurement systems and accelerators. For a more detailed and
complete description of measurement procedures, the readers should consult, for
example reference [4], 'A Handbook of Radioactivity Measurements Procedures'.
Such books, however, assume that the reader already has a basic knowledge of the
principles of particle measurement, and this is what the present textbook aims to
provide.

   For preparing these lectures, and lecture notes, I have used material from many
different sources, and no attempt is made to give complete references. I found the
following books particularly useful:

– 'Radiation Detection and Measurement' by Glen Knoll [5]. This is a reference
  book about nuclear measurement techniques. It contains extensive references to
  the original literature and a wealth of useful information.
– 'Review of Particle Physics' by Amsler [6]. This document is oriented towards
  particle physics, but several parts of it are also very useful for researchers active
  in the field of nuclear science. In particular, the tables with numerical values are
  very convenient and useful. Moreover, this document is freely downloadable from
  the Web. Mainly the section 'Constants, Units, Atomic and Nuclear Properties',
  is of interest to students. Several parts of this last document are reproduced as
  annexes at the end of these lecture notes.

– 'Principles of Charged Particle Acceleration' by Stanley Humphries (Ref. [10] in Chap. 3) and 'An Introduction to Particle Accelerators' by Edmund Wilson (Ref. [3] in Chap. 3). Both books contain an excellent introduction to particle accelerators.
– 'Techniques for Nuclear and Particle Physics Experiments' by W.R. Leo [7].

Data on properties of isotopes and nuclear reactions can be found the National Nuclear Data Centre tables [8]. Other useful information can be found in the Physical Reference Data from the National Institute of Standards and Technology (NIST) [9] and in Kaye and Laby Tables of Physical and Chemical Constants [10]. These three documents are also freely downloadable from the Web.

At the end of each chapter there is a list of other material used in preparing that chapter, followed by a few references.

1.2 Units and Physical Constants

In nuclear and particle physics it is common to use units that are somewhat different from those that are standard elsewhere. The charge is expressed in number of proton charges, 1 proton charge $= 1.602 \times 10^{-19}$ C, and the electron has the same charge as the proton, but of opposite sign. Energy is usually expressed in 'electron-volt'. One electron-volt (eV) is the energy that a proton or electron acquires if it goes through a potential difference of 1 V and hence $1\,\text{eV} = 1.602 \times 10^{-19}$ J. To avoid using large numbers, one often uses

$$1\,\text{keV} = 10^3\,\text{eV}$$
$$1\,\text{MeV} = 10^6\,\text{eV}$$
$$1\,\text{GeV} = 10^9\,\text{eV}$$
$$1\,\text{TeV} = 10^{12}\,\text{eV}$$

A mass is expressed in kg, but in nuclear and particle physics, it is common to express the mass as the equivalent energy using the well-known relation $E = mc^2$. The mass of a proton is 1.672×10^{-27} kg, but I will usually write that the mass of a proton is 938.272 MeV/c^2. This simply means that $m_{\text{proton}}c^2 = 938.272$ MeV. In this relation, c represents the speed of light in vacuum, $c = 299\,792\,458$ m s^{-1}. Similarly, for the momentum P, the quantity cP has the dimension 'energy', and I will mention the momentum in units of MeV/c. In the physics literature, it is common practice to omit factors c and $\hbar$ in the equations, but in these lecture notes, I will always write these factors explicitly.

The mass of atoms and isotopes is usually expressed in 'unified atomic mass units'. By definition this unit is 1/12th of the mass of a ^{12}C atom. One unified atomic mass unit equals 931.494 MeV/c^2 or 1.660×10^{-27} kg.

For X-rays and gamma rays, there are a few simple and important relations:

$$E = h\nu = \hbar\omega, \quad \text{and} \quad \lambda\nu = \text{c}$$

In these equations, ν = frequency, λ = wavelength; $\omega = 2\pi\nu$ is the angular frequency and h is the Planck constant. The reduced Planck constant $\hbar$ is also often used:

$$\hbar = \frac{h}{2\pi} = 1.05410^{-34} \text{Js}$$

In nuclear and particle physics, the quantity hc or $\hbar c$ often enters into the calculations. In convenient units, these quantities are given by

$$\hbar c = 197.610^{-15} \text{MeV.m}$$
$$hc = 1.242\, eV.\mu\text{m}$$

The fine structure constant, α, and the classical electron radius, often enter into theoretical calculations. These quantities are given by

$$\text{fine structure constant: } \alpha = \frac{e^2}{4\pi\hbar c\varepsilon_0} \approx \frac{1}{137.035}$$

$$\text{classical electron radius: } r_0 = \frac{e^2}{m_0 c^2 4\pi\varepsilon_0} = 2.81810^{-15}\text{m,}$$

where m_0 is the electron mass; $m_0 c^2 = 511$ keV.

The fine structure constant is the squared charge of the electron combined with some other fundamental constants so as to give a *dimensionless number*. The fact that this number is much smaller than unity somehow says in an absolute, dimension-independent way that the charge of the electron is small. The name 'classical electron radius' is misleading, because it has nothing to do with the true dimension of the electron. The true dimension of the electron is not known but it is certainly smaller than 10^{-18} m.

Annex 1 lists most of the numerical constants that are useful in the contexts of nuclear and particle physics.

1.3 Special Relativity

Whenever objects travel at a speed that is a significant fraction of the speed of light, it is necessary to use relativistic formulas for kinetic variables such as speed, energy and momentum. This is the case if the kinetic energy of a particle is a sizeable fraction of, or larger than, the rest energy mc^2. In nuclear physics, alpha particles and nuclear fragments are always slow compared to the speed of light, and non-relativistic equations are often sufficient. However, most of the time electrons have a velocity close to the speed of light, and it is essential to use the correct relativistic equations. In high-energy particle accelerators, all the particles move at a velocity close to the velocity of light, and relativistic equations must be used. I have therefore added a section to remind the reader of the main elements of special relativity.

Fig. 1.1 The observer S′
moves with a velocity 'v'
relative to the observer S.
Each observer has his own
reference system, {x, y, z}
and {x′, y′, z′}
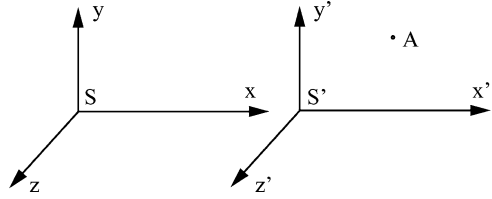
Consider the situation illustrated in Fig. 1.1, two observers, S and S′ are moving
with a velocity 'v' relative to each other. In non-relativistic physics, the relation
between the position coordinates x, y, and z and the time coordinate t of an object
located at point 'A', as seen by the two observers, is given by

$$\begin{cases} x' = x - vt \\ y' = y \\ z' = z \\ t' = t \end{cases}$$

In non-relativistic physics, the kinetic energy '$E_{kinetic}$' and the momentum 'P' of
a particle of mass 'm' with a velocity 'v' is given by

$$E_{kinetic} = \frac{1}{2}mv^2 \quad P = mv$$

In special relativity theory, the above transformation relations are replaced by the
Lorentz transformation:

$$\begin{cases} x' = \dfrac{x - v.t}{\sqrt{1 - (v/c)^2}} \\ y' = y \\ z' = z \\ t' = \dfrac{t - v.x/c^2}{\sqrt{1 - (v/c)^2}} \end{cases}$$

The Lorentz transformation can be written in a more elegant way with the help
of the parameters β and γ as

$$\begin{cases} x' = \gamma(x - \beta ct) \quad \beta = \dfrac{v}{c} \\ y' = y \\ z' = z \qquad\qquad \gamma = \dfrac{1}{\sqrt{1 - (v/c)^2}} \\ ct' = \gamma(ct - \beta x) \end{cases}$$

An essential property of the Lorentz transformation is that it guarantees that the speed of light is the same for all observers. This transformation has a number of surprising consequences: one is that the time is no longer an absolute time, but depends on the observer.

The total energy E and the momentum P $\{E, P_x c, P_y c, P_z c\}$ of any object obey the same Lorentz transformation as the time and position coordinates $\{ct, x, y, z\}$. We therefore have

$$\begin{cases} P'_x c = \gamma (P_x c - \beta E) \\ P'_y c = P_y c \\ P'_z c = P_z c \\ E' = \gamma (E - \beta P_x c) \end{cases}$$

An important consequence of the Lorentz transformation for the energy and the momentum is the relation

$$m_0^2 c^4 = E^2 - \vec{P}^2 c^2 \tag{1.1}$$

In Eq. 1.1, $\vec{P}^2$ stands for $P_x^2 + P_y^2 + P_z^2$. This equation shows, among other things, that the energy of a particle at rest is $E = m_0 c^2$.

To obtain the correct relativistic expressions for the energy and the momentum of a particle as a function of its velocity, let us consider a particle at rest in the system S. The energy and the momentum of this particle are given by

$$\begin{cases} P_x c = 0 \\ P_x c = 0 \\ P_x c = 0 \\ E = m_0 c^2 \end{cases}$$

Consider the frame S' moving relative to S with a velocity v in the opposite direction of the x-axis. The energy and the momentum of this particle in S' are given by

$$\begin{cases} P_x c = \beta \gamma \, m_0 c^2 \\ P_x c = 0 \\ P_x c = 0 \\ E = \gamma \, m_0 c^2 \end{cases}$$

Clearly the particle is moving with velocity v in the system S'. We conclude that the energy and the momentum of a particle with velocity v are given by

$$p = \beta\gamma\, m_0 c^2$$

$$E = \gamma\, m_0 c^2$$

For relativistic particles the dependence of the kinetic energy and the momentum on the velocity is therefore is given by

$$E = E_{\text{kinetic}} + m_0 c^2 = \frac{m_0 c^2}{\sqrt{1 - (v/c)^2}} \tag{1.2}$$

$$P = \frac{v m_0 c^2}{\sqrt{1 - (v/c)^2}} \tag{1.3}$$

From Eq. (1.2), we immediately derive the relation between the velocity 'v' of the particle and its kinetic energy:

$$\frac{v}{c} = \sqrt{1 - \left(\frac{m_0 c^2}{E_{kinetic} + m_0 c^2}\right)^2} \tag{1.4}$$

Using the above equation one finds that the velocity of an electron with a kinetic energy of 1 MeV is 95% of the speed of light, while the velocity of a proton with a kinetic energy of 1 MeV is only 5% of the speed of light.

From Eq. (1.1), we obtain the correct relativistic relations between kinetic energy and momentum:

$$E_{kin} = \sqrt{P^2 c^2 + m_0^2 c^4} - m_0 c^2 \tag{1.5}$$

$$Pc = \sqrt{2 m_0 c^2 E_{kin} + E_{kin}^2} \tag{1.6}$$

If the energy of the particle is much larger than its rest mass energy, these equations, to a good approximation, simplify to $E_{\text{kinetic}} \approx E \approx Pc$.

Another important consequence of the Lorentz transformation is that the duration of time intervals is no longer the same for all observers. Consider an object at rest and located at the origin in reference system S. Suppose the object is emitting a light flash at time t_1 and another light flash at time t_2. The time difference between these two light flashes as seen in the system S is $t_1 - t_2$. In the system S', this time interval is given by

$$t_2' - t_1' = \frac{t_2 - t_1}{\sqrt{1 - v^2/c^2}} = \gamma(t_2 - t_1)$$

The same time interval is longer in the system S'. This is completely general. All events happening with an object in motion will be slower than that with the same

object at rest. In particular, the decay time of object travelling at a larger velocity will be longer by a factor γ than the decay time for the same object at rest.

1.4 Probability and Statistics

Radioactive decays, or interactions created at particle accelerators, are statistical phenomena. There are a few concepts borrowed from probability and statistics that we will use again and again during these lectures. Therefore, I have added a brief paragraph to remind the students of the most important of these concepts. Many more concepts of probability and statistics are essential tools in the analysis of experimental data in nuclear and particle physics. However, these concepts are not used in the present lecture notes and are therefore not mentioned here. Reference [11] is one of many good textbooks on this subject. The sections on probability and statistics in reference [6] also contain very useful methods for the statistical analysis of data.

Gaussian distribution. Consider a continuous statistical variable x with probability density function $f(x)$. The quantity $f(a)dx$ represents the probability that x takes a value between a and $(a + dx)$.

The Gaussian or normal probability density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{1.7}$$

The Gaussian distribution, Eq. (1.7), has the following properties:

(1)
$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

(2) The average value of x, often written as <x> or as $\bar{x}$, is given by

$$\bar{x} = < x > = \int_{-\infty}^{+\infty} x f(x)dx = \mu$$

(3) For a statistical variable x, with probability density function $f(x)$, the dispersion or root-mean-square (r.m.s.) deviation, usually written as σ, is given by

$$\sigma^2 = \left\langle (x - \bar{x})^2 \right\rangle = \int (x - \bar{x})^2 f(x)\,dx$$

For a Gaussian distribution, the r.m.s. dispersion is the parameter σ occurring in Eq. (1.7). Equation (1.7) therefore represents a Gaussian distribution with average value μ and dispersion σ.

Fig. 1.2 Gaussian or normal
distribution. FWHM stands
for 'full width at half
maximum'. The probability
that the value of x is within
one standard deviation σ of
the average value is 68% and
the probability that x is within
two standard deviations is
95%

Gaussian distribution
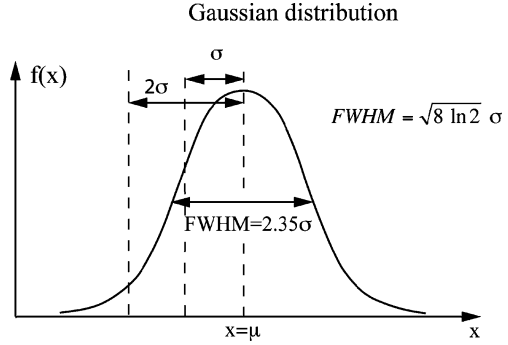
A Gaussian distribution has the well-known bell shape shown in Fig. 1.2. The width of such a distribution is often characterised by its 'full width at half maximum', usually abbreviated as FWHM. From the probability density distribution Eq. (1.7), one immediately finds that the FWHM of a Gaussian distribution is related to the dispersion σ by FWHM $= \sigma \sqrt{8 \ln 2} = 2.355\, \sigma$.

Assume that we have a large number of statistical variables x_i, with arbitrary probability density functions $f_i(x_i)$. Consider the new variable y defined as

$$y = \sum_i x_i$$

It can be shown that, under very general conditions, and regardless of what the distributions $f_i(x_i)$ are, the probability distribution of y is a Gaussian distribution. This is the central limit theorem, and it explains why Gaussian distributions are so common in many experimental situations. Imagine that you are measuring some physical quantity. The value you measure will not be the true value because the measurement will be affected by a measurement error. Usually, there are a large number of different effects, all adding up to the measurement error. It is, therefore, not surprising that a measurement error often has a Gaussian distribution.

Poisson distribution. Let us now consider a statistical variable n that can only take integer values. $P(k)$ represents the probability to observe the value $n = k$.

For a Poisson distribution these probabilities $P(n)$ are given by

$$P(n) = \frac{\lambda^n}{n!} e^{-\lambda}$$

It is easy to show that the Poisson distribution has the following properties:

(1)
$$P(0) = e^{-\lambda}$$

(2)
$$\sum_{n=0,\infty} P(n) = 1$$

(3)
$$< n >= \sum_{n=0,\infty} nP(n) = \lambda$$

(4) The dispersion or r.m.s., usually written as σ, is given by

$$\sigma^2 =< (n- < n >)^2 >= \sum_{n=1,\infty} (n- < n >)^2 P(n) = \lambda$$

It can also be shown that, for large values of λ, the Poisson distribution approaches a Gaussian distribution. This approximation is very good as soon as λ is larger than a few 10! Therefore, the FWHM of a Poisson distribution as a function of the average value λ is given by FWHM $= 2.355\sqrt{\lambda}$.

The number of occurrences of a particular event A will have a Poisson distribution if

– a large number of primary events, or primary situations, can give rise to the occurrence of A, and each of these primary events has only a small chance to give rise to the occurrence of A and
– there is no correlation between the primary events.

Assume that we have some amount of radioactive material and are observing the decay of this material. Let us further assume that the half-life of the material is long compared to the duration of the observation. This is clearly a situation where the above conditions are met. The probability that the decay of one particular atom is observed is very small, but there are a large number of atoms, and all have a small, but finite, probability to give rise to an observed decay. In addition, the decays are independent of one another. In this case, the observed number of decays will have a Poisson distribution.

Dispersion of a sum of two statistical variables. Let x and y be two statistical variables. Consider the statistical variable $z = x + y$. Then, we have the following important relations

$$\bar{z} = \bar{x} + \bar{y}$$
$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2$$

These relations hold for any probability distributions for the variables x and y.

1.5 The Structure of Matter at the Microscopic Scale

All matter is made up of atoms, and atoms have a size of the order of 10^{-10} m. We have known for about a hundred years that atoms are composed of a nucleus surrounded by a cloud of electrons. To the best of our knowledge, the electrons are truly elementary particles. If they have a dimension at all, that dimension is less than 10^{-18} m. However, the nucleus is a complex object. The dimension of the nucleus is of the order of 10^{-15} m. It is composed of protons and neutrons. These protons and

neutrons, in turn, are composed of quarks. To the best of our knowledge, the quarks are also truly elementary particles. If they have a dimension at all, that dimension is less than 10^{-18} m.

In ordinary matter there are two types of quarks, called up-quarks and down-quarks. A proton is made of two up-quarks and one down-quark, and a neutron is made of two down-quarks and one up-quark. The quarks have a charge that is a fraction of the proton charge: up-quarks have a charge of +2/3, down-quarks have a charge of −1/3. Hence all ordinary matter is made of three basic components: electrons, up-quarks and down-quarks. This very simple picture is not sufficient to describe the reality. First, there is a very enigmatic particle called the neutrino. It is electrically neutral and has a very small probability to interact. Furthermore, for all these particles there are corresponding antiparticles. Finally, this basic set of four components (up-quark, down-quark, electron and electron neutrino) is repeated two times with heavier versions of each particle. There is the electron and two heavy electrons, namely the muon and the tau-lepton. The muon has a mass of 105.65 MeV/c^2, about 200 times larger than the electron, and the tau-lepton has a mass of 1,777 MeV/c^2, about two times the mass of a proton! If I say that a muon is a heavy electron, this means that a muon is in all respects the same as the electron except for its mass. The muon decays with a lifetime of 2.2 μs into an electron and two neutrinos. The fact that muon decays does not make it less fundamental than the electron; the muon decays because this is energetically possible.

The forces between those elementary particles are

– gravitational force
– electromagnetic force
– weak force and
– strong colour force.

The gravitational force and the electromagnetic force are quite familiar from observations in the macroscopic world, while the weak force and the strong colour force only manifest themselves at the subatomic scale. This is because the last two forces are short-range forces.

The gravitational force is the overwhelming force at the macroscopic scale, but at the nuclear scale it is essentially unobservable. The ratio of the gravitational force to the electromagnetic force between an electron and a proton is 10^{+40}! At the atomic or nuclear scale, one can safely ignore the gravitational force.

The interaction between the fundamental components of matter is well described by the equations of relativistic quantum field theory. Mathematically, these equations are very complex, but one can get an intuitive feeling of what is going on by thinking of an interaction between particles as being due to the exchange of quanta of force. A collision between two charged particles, say two electrons, can be seen as due to the exchange of a quantum of electromagnetic force, as illustrated in Fig. 1.3.

This quantum of electromagnetic force is a virtual photon. It is not a real particle and can exist only for a very short time. Unlike a real photon, the mass of a virtual

Fig. 1.3 A collision between two electrons is due to the exchange of one or more virtual photons. A photon can be seen as a quantum of electromagnetic force

electron

photon

electron

electron

photon

anti up-quark

positron

(a)

up-quark

electron

photon

positron

photon

(b)
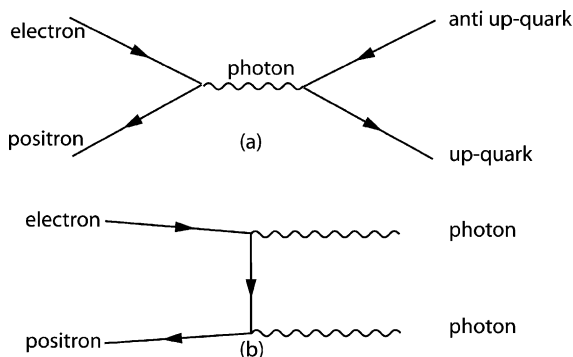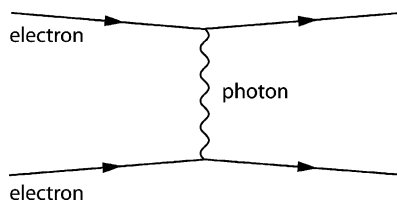
Fig. 1.4 Example of electromagnetic interactions of charged particles. In these diagrams the direction of time is from left to right. An arrow pointing against the direction of time, i.e. to the left, represents an antiparticle. (**a**) An electron and a positron annihilate each other and materialise again as a quark-anti-quark pair. (**b**) An electron and a positron annihilate each other into two gamma rays

photon can be different from zero and can even be negative. However, in quantum field theory there are a number of other possible processes. Figure 1.4(a) shows a process where an electron and an anti-electron (called a positron) meet and annihilate into a virtual photon. This cannot be a real, massless, photon because that would violate energy and momentum conservation laws. This virtual photon can materialise in any charged particle–antiparticle pairs. Figure 1.4(b) shows the annihilation of an electron–positron pair into two gamma rays. In this process, the photons are real particles, but the electron connecting the two gamma mission points is a virtual electron. The charge of a particle describes the strength of its coupling to the electromagnetic field. Photons couple only to charged particles. Photons themselves carry no charge and are massless.

In these few examples we see that the distinction between a 'particle' and a 'force' is disappearing at the subatomic level. The electromagnetic force manifests itself as quanta, called photons, while a particle, such as an electron manifests force-like or wave-like behaviour.

The most prominent force at the nuclear scale is the strong colour force. This force is also described by equations that are very similar to those describing the electromagnetic interaction. The force quanta of this strong force are 'gluons', and similar to photons these are massless particles. There are several very important differences between the strong colour force and the electromagnetic force. The strong

colour charges, the equivalent of the electrical charge for the strong force, appear in three kinds. These three charges were whimsically called 'red', 'green' and 'blue', and the force is therefore called the colour force. Only quarks have a colour charge and therefore only quarks feel the strong force. The antiparticles of quarks, called anti-quarks, have charges 'anti-red', 'anti-green' and 'anti-blue', and these charges are different from 'red', 'green' and 'blue'. Unlike the electromagnetic force, where the quanta of electromagnetic force carry no charge, gluons themselves have a colour charge. It can be shown that all these give rise to a force between colour-charged objects that does not decrease with distance. It takes an infinite amount of energy to separate two quarks!

As a result, all particles with colour are permanently locked up in colour-neutral systems. Quarks or gluons cannot exist as free particles. Only colour-neutral combinations of quarks can exist. Just as an electron and a positron together have a total charge zero, a quark and its anti-quark too have a total colour charge zero, and a suitable colour-combination of a quark and its anti-quark can form a colour-neutral system.

It is possible to show that the only possible combinations of quarks that are colour-neutral are a quark and an anti-quark, three quarks or three anti-quarks. Of course, all combinations or multiples of these can also be colour-neutral. The only quark combinations that exist in nature are a combination of a quark and an anti-quark, of three quarks or of three anti-quarks! All particles that are composed of quarks are called hadrons. A very large number of different hadrons exist. However, most of these are extremely short-lived and have a lifetime of only $\approx 10^{-23}$ s. This lifetime is so short that such particles cannot travel a macroscopic distance before they decay. Even at a very high energy, the maximum distance they will travel is less than the size of an atom. Therefore, such particles can never be observed directly. Their existence is inferred from the observation of their decay products. A few hadrons have a much longer lifetime, in the range 10^{-8} to 10^{-16} s. This lifetime is also very short, but at a velocity close to the velocity of light, these particles can travel distances in the range from microns to hundreds of meters. Table 1.1 lists some among the most common hadrons with a long lifetime, together with their main properties.

The most familiar quark–anti-quark combination is the π meson. It exists as a π^+ (up-quark + anti-down-quark) and π^- (down-quark + anti-up-quark) or as π^0 (up-quark + anti-up-quark and down-quark + anti-down-quark). All the particles consisting of a quark and an anti-quark are called mesons. The combinations of three quarks and three anti-quarks are called baryons and anti-baryons, respectively.

The most familiar combinations of three quarks are the protons and neutrons, but many more exist. Protons and neutrons, and more generally all hadrons, are colour-neutral systems; therefore, there is no strong force between these particles except at very short distances. This is similar to the electromagnetic force between neutral atoms or molecules. There is no electromagnetic force between two neutral atoms except at very short distances. However, between two neutral molecules that are touching, there is a force called the van der Waals force. It is an electromagnetic force. Similarly, between two nucleons that are some distance apart there is no force. As they come closer together, first there is a force that is attractive, and as they come

Table 1.1 Some hadrons and their main properties. In this table the antiparticle of particle x is written as $\bar{x}$

Name	Quark content	Lifetime [s]	Mass $\times c^2$ [MeV]	Main decay modes
π^+	$\{u, \bar{d}\}$	$2.6\ 10^{-8}$	139.6	$\mu^+ \nu_\mu$
π^-	$\{d, \bar{u}\}$	$2.6\ 10^{-8}$	139.6	$\mu^- \bar{\nu}_\mu$
π^0	$\{u, \bar{u}\}, \{d, \bar{d}\}$	$1.2\ 10^{-17}$	135.0	2γ
K^+	$\{u, \bar{s}\}$	$1.2\ 10^{-8}$	473.7	$(\mu^+ \nu_\mu), (2\pi), (3\pi)$
K^-	$\{s, \bar{u}\}$	$1.2\ 10^{-8}$	473.7	$(\mu^-, \bar{\nu}_\mu), (2\pi), (3\pi)$
K_l^o	$\{s\,\bar{d}\}, \{d\,\bar{s}\}$	$5.2\ 10^{-8}$	497.7	3π
K_s^0	$\{s, \bar{d}\}, \{d, \bar{s}\}$	$8.9\ 10^{-11}$	497.7	2π
proton	$\{u, u, d\}$	$>10^{32}$ year	938.3	
neutron	$\{d, d, u\}$	898	939.6	$p\,e\,\bar{\nu}_e$
Λ^0	$\{d, s, u\}$	$2.63\ 10^{-10}$	1115.7	$(n\,\pi^0), (p\,\pi^-)$
$\bar{p}$	$\{\bar{u}, \bar{u}, \bar{d}\}$	$>10^{32}$ year	938.3	
$\bar{n}$	$\{\bar{d}, \bar{d}, \bar{u}\}$	898	939.6	$\bar{p}\,e^+\nu_e$
$\bar{\Lambda}^0$	$\{\bar{d}, \bar{s}, \bar{u}\}$	$2.63\ 10^{-10}$	1115.7	$(\bar{n}\,\pi^0), (\bar{p}\,\pi^+)$

even closer the force becomes strongly repulsive. Thus, protons, neutrons and all other hadrons appear to have a size, and that size is about 10^{-15} m.

Another force that is observed at the subatomic scale is the 'weak force'. This force causes many decay processes in hadrons and nuclei. Mathematically, it is also described by equations that are similar to the electromagnetic force. The main difference comes from the fact that the quanta of the weak force are massive objects: the W^+, W^- and the Z bosons. They have a mass that is approximately 90 times the proton mass.

All elementary particles have a weak charge. Figure 1.5(a) and (b) illustrate a few reactions that are mediated by the weak force. Figure 1.5(b) represents the 'decay' of a down-quark into an up-quark, an electron and a neutrino. This reaction can never be observed since there are no free quarks, but this is the process that explains the neutron decay (see Fig. 1.5(c)). This process can be understood as follows: the down-quark in the neutron emits a virtual, charged, W^- boson. The charge needs to be conserved, and that is possible if the down-quark (charge $= -1/3$) changes into an up-quark (charge $= +2/3$). The mass of a W boson is 90 times the proton or neutron mass, so this process seems impossible. Remember though that in quantum mechanics a process that is energetically impossible is nevertheless possible, provided this unphysical state lasts only a short time. This is the well-known 'tunnelling effect'. A particle can cross an energy barrier that it does not have enough energy to cross, provided the barrier is very thin, so the particle is not long in the unphysical state of negative energy.

The uncertainty of the energy and the duration of this state are related by $\hbar \leq \Delta E \Delta t$. If the uncertainty of the energy needs to be 90 times the proton mass, then the distance the virtual W boson can travel is given by

Fig. 1.5 Examples of some reactions mediated by the exchange of quanta of the weak force. (**a**) Annihilation of a down-quark and an anti-up-quark into an electron and an anti-neutrino. (**b**) 'Decay' of a down-quark into an up-quark, an electron and an anti-neutrino. (**c**) The phenomena depicted in (**a**) and (**b**) cannot exist as such because free quarks cannot exist. However, the decay of a neutron into a proton, an electron and a neutrino is due to the process (**b**)
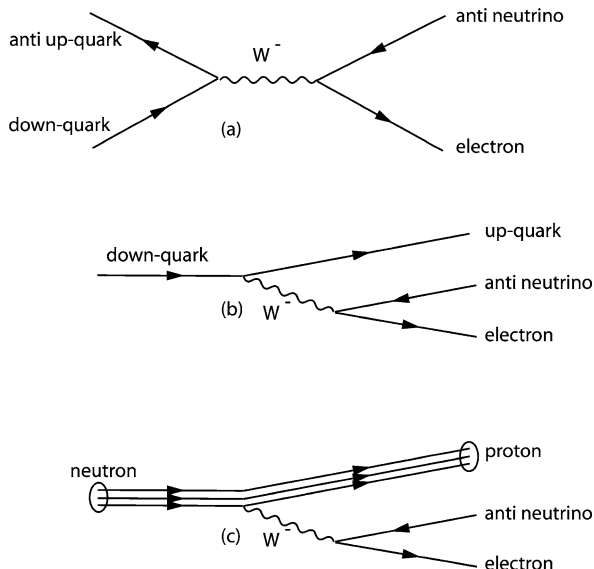
$$c\Delta t \approx \frac{\hbar c}{90\,\text{GeV}} \approx 10^{-18}\text{m}$$

Hence, the process illustrated in Fig. 1.5(c) is possible, provided it all happens in a distance of the order of 10^{-18} m. The 'weak force' is not really weak, it has a short range!

The complete list of all fundamental building blocks of nature is given in Table 1.2, and the quanta corresponding to the fundamental forces of nature are listed in Table 1.3. These tables summarise all that is presently known about the fundamental building blocks of nature and about the fundamental forces. There are three 'up-quark like' quarks: the up-quark, the c-quark and the t-quark; there are three 'down-quark like' quarks: the down-quark, the s-quark and the b-quark; there are three leptons: the electron, the muon and the τ-lepton and there are three neutrinos: the electron neutrino, the muon neutrino and the τ-neutrino. The force quanta

Table 1.2 Fundamental building blocks in nature. All these particles have spin = $\frac{1}{2}$. For each of these particles there is a corresponding antiparticle

Particle	Electric charge	Colour charge	Weak charge
Up-quark, c-quark, t-quark	2/3	Yes	Yes
Down-quark, s-quark, b-quark	−1/3	Yes	Yes
Electron, muon, τ-lepton	−1	No	Yes
Neutrino (ν_e, ν_μ, ν_τ)	0	No	Yes

Table 1.3 Quanta corresponding to the fundamental forces in nature. All these quanta have a spin = 1

Name	Mass [GeV/c^2]	Electric charge	Colour charge
Photon	0	No	No
W$^+$, W$^-$	80.4	Yes	No
Z	91.2	No	No
Gluon	0	No	Yes

are the photon, the W- and Z-boson and the gluons. However, many of these particles cannot be observed as real particles. The quarks and gluons are permanently locked in colour-neutral systems such as protons and neutrons, and the W- and Z-bosons are so short-lived that they can never be observed directly. Only their decay products can be observed. Only the proton, the neutron in a nucleus, the electron and the neutrino are stable. The neutrinos interact so weakly with other particles that they are largely decoupled from the world we observe. Hence, ordinary matter only consists of protons, neutrons and electrons. All the other particles only exist if they have been created in some very high-energy collision. Soon after their creation the particles decay into one of the few stable particles. Such particles have also been created in huge numbers in the beginning of the Universe.

Table 1.4 lists the particles that are stable, or at last can exist long enough to travel truly macroscopic distances. For all these particles, except the photons and some neutral mesons, there are corresponding antiparticles. For the rest of the present lecture we will only consider the particles listed in Table 1.4.

At this point it should be explained that quantum mechanics implies that exploring small dimensions needs high energies. The study of the structure of matter at the subatomic scale requires the use of high-energy probes, and in reactions between small objects, particles of high energy are produced. Indeed, in quantum mechanics a wave is associated with every particle, and the wavelength λ is given by

$$\lambda = \frac{h}{P} = \frac{hc}{Pc} = \frac{1237[\text{MeV}]10^{-15}[\text{m}]}{Pc} \tag{1.8}$$

If the momentum Pc of a particle is 1237 MeV, the corresponding wavelength is 10^{-15} m. To probe the structure of objects much smaller than 10^{-15} m, we need

Table 1.4 List of the most common directly observable particles

Particle	Mass	Lifetime	Charge	Main interactions
Electron	0.511 MeV	Stable	-1	Electromagnetic
Muon	105.7 MeV	$2.2\ 10^{-6}$ s	-1	Electromagnetic
Hadron	see Table 1.1	see Table 1.1	0, +1, -1	Strong
Photon	0	Stable	0	Electromagnetic
Nuclei	1–240 times the proton mass	Many stable	1–92	Strong
Neutrino	$<\approx2$ eV	Stable	0	Weak

particles with energy much larger than 1237 MeV. If the energy of an object is much larger than its mass, then energy E and momentum Pc become the same.

If a particle is enclosed in a one-dimensional box of dimensions 'a', the energy levels of this particle are given by

$$E_n = \frac{n^2\pi^2(\hbar c)^2}{2mc^2a^2} \tag{1.9}$$

The above result is found by solving the Schrödinger's equation in the very simple case of an infinitely deep potential well with vertical slopes. We see that the spacing between the energy levels in a small object will be very large. Qualitatively this remains true whatever potential confines objects to a small volume. In atoms, the energy levels of the electrons are in the range 1 eV to a few keV. In nuclei, the energy levels spacing is in the range 10 keV to 10 MeV. The energy level differences in hadrons are of the order of 1 GeV.

If the particles we regard as fundamental today, quarks, electrons etc., are in fact composed of some as yet unknown more fundamental objects, we know that the corresponding size has a scale smaller than 10^{-18} m. Therefore, the corresponding energy levels and particle masses are expected to be of the order of 1 TeV or larger. If this is the case, the particles we know today are only the lowest energy levels of these objects, and a whole spectrum of particles that are much more massive than the particles we know today, exists.

All these explain why it needs very high-energy accelerators to study very small particles. It is the motivation for the vast worldwide effort to build very high-energy particle accelerators.

1.6 Nuclei and Nuclear Decay

All the familiar matter surrounding us is made up of atoms. The atoms are made up of a nucleus containing most of the mass, and of electrons. The nucleus is a bound state of protons and neutrons held together by the strong nuclear force. Neutrons and protons are collectively referred to as 'nucleons'. The charge of the nucleus is equal to the number of protons. The mass of the nucleus is about 1% smaller than the sum of the masses of the constituent neutrons and protons. The difference is due to the binding energy of the nucleus. Figure 1.6 shows the binding energy per nucleon for all stable elements. Notice that the binding energy per nucleon increases with the mass of the nucleus, reaches a maximum for iron and nickel and then decreases again. The helium nucleus is unusual: it has an exceptionally large binding energy.

Like the electron structure of the atom, the nucleus also has a number of discrete energy levels. The energy difference between the nuclear energy levels is typically of the order of 1 MeV, and the transitions between such states give rise to particles with energies in the 1 MeV range.
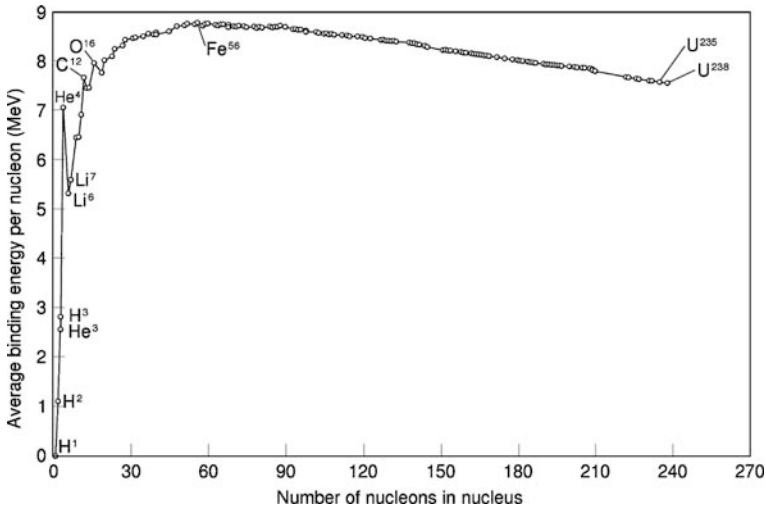
Fig. 1.6 Binding energy per nucleus for all stable isotopes

The charge of the nucleus determines the chemical properties of the atom. The presence of neutrons in the nucleus is essential to make the nucleus stable. A nucleus needs an approximately equal number of protons and neutrons to be stable. High-mass nuclei need more neutrons than protons to be stable. Nuclei of the same charge but with a different number of neutrons are called isotopes. Only a few of all the possible isotopes are stable. For example iron has four stable isotopes: $^{54}_{26}Fe, ^{56}_{26}Fe, ^{57}_{26}Fe, ^{58}_{26}Fe$. The notation $^N_Z A$ stands for a nucleus A containing Z protons and in total N nucleons. The indication of the number of protons in the nucleus is redundant since the element symbol already gives the charge of the nucleus; this number is therefore often omitted. Besides the stable isotopes, there are many unstable isotopes with a larger or a smaller number of neutrons. These decay in one or more steps until a stable configuration is reached.

1.6.1 The Beta Decay

An unstable nucleus can sometimes reach a lower energy state by changing a proton into a neutron or vice versa. To conserve the charge in this transformation process, an electron or a positron must be emitted. The corresponding reactions on free protons or neutrons are

$$n \rightarrow p + e^- + \bar{\nu}_e$$
$$p \rightarrow n + e^+ + \nu_e$$
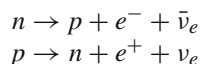
The symbols ν_e and $\bar{\nu}_e$ stand for the electron neutrino and the electron antineutrino, respectively. These reactions are mediated by the exchange of W bosons

as illustrated in Fig. 1.3. Because the decay is mediated by the weak force, the corresponding lifetime can be quite long. On free particles, only the decay of neutrons to protons is possible since the mass of the protons is less than the mass of the neutrons. Inside a nucleus, each of the two reactions is possible if the mass of the new nucleus plus the mass of the electron or positron is smaller than the mass of the original nucleus. Written as the decay of a nucleus the process is

$$
{}^{N}_{Z}A \rightarrow {}^{N}_{Z+1}A + e^{-} + \bar{\nu}_{e}
$$
$$
{}^{N}_{Z}A \rightarrow {}^{N}_{Z-1}A + e^{+} + \nu_{e}
$$
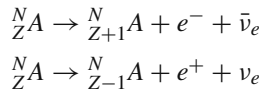
In nuclear physics these decays are called β^{-} and β^{+} decays. The neutrino or antineutrino produced in the β decay is almost never observed, but it takes away part of the energy liberated in the reaction. The mass difference between the initial nucleus and the sum of the masses of the final-state particles appears as kinetic energy of the particles in the final state. The energy corresponding to this mass difference is usually denoted by Q. The mass of the nucleon in the final state is much larger than the mass of the electron or the neutrino and also much larger than the mass difference. As a result the kinetic energy of the final-state nucleus is very small, usually only a few keV. To first approximation all the energy is shared between the electron and the neutrino only. For a proof of this statement see the solution to Exercise 7. Since there are three particles in the final state the momenta of the final state particles are not determined uniquely. The electron and the neutrino can have a kinetic energy varying between zero and the maximum allowed energy. The mass of the neutrino is extremely small, less than 2 eV; therefore the maximum energy of the electron is equal to the energy Q. In tables with nuclear decays, the corresponding maximum energy of the electron is usually listed as $E_{\max}$.

1.6.2 The Alpha Decay

A very heavy nucleus has another possibility to reach a more stable configuration. It can just fall apart into two lighter nuclei. This will be energetically favourable since for very heavy nuclei the binding energy per nucleon decreases with increasing mass of the nucleus. For a few very high mass isotopes this can happen by splitting the nucleus into two more or less equal parts. For nearly all other isotopes this always happens through the emission of a helium nucleus, also called alpha particle. The corresponding reaction is

$$
{}^{N}_{Z}A \rightarrow {}^{N-4}_{Z-2}A + \alpha
$$
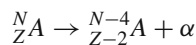
Because the transition is mediated by the strong force, the long decay time of some alpha emitters is surprising. The explanation is that the decay requires the alpha particle to tunnel through a potential barrier of the nucleus. If the final-state nucleus in the alpha emission is in a well-defined nuclear level, energy and

momentum conservation fixes the kinetic energy of the alpha particle. Of course a given isotope can decay by alpha emission to several distinct nuclear levels, and the corresponding alpha particles have different energies.

1.6.3 The Gamma Decay

After a β^- or a β^+ decay, or after alpha particle emission, the nucleus is often not in its ground state, but in some excited nuclear state. Transitions between the excited levels and the ground state can give rise to the emission of gamma rays. Most transitions giving rise to the emission of a gamma ray are extremely fast. However, for some transitions the direct decay mechanism is forbidden, and the corresponding decay is much slower. A nucleus that is trapped in one of these metastable states is called an isomer, and this is denoted by a letter m after the mass number, e.g. ^{60m}Co.

1.6.4 Electron Capture and Internal Conversion

A nucleus also has a number of other possibilities to decay. There is a finite probability that an electron from the cloud of electrons surrounding the atom is present inside the nucleus. Consider a β^+ decay. The nucleus can reach the same nuclear final state and conserve the charge by capturing an orbital electron of the atom in the reaction

$$p + e^- \rightarrow n + \nu_e$$
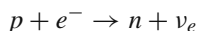
This process is called 'electron capture'. This reaction occurs mainly in heavy nuclei where the nucleus is larger and the electron orbits are smaller. In most cases the captured electron is the K-shell electron, but L-shell electron capture also occurs. After the electron capture the empty level left in the electronic structure of the atom is filled by an outer electron, and the excess energy is liberated by the emission of an X-ray or by the emission of an Auger electron. The relative importance of β^+ decay and electron capture also depends on the difference in energy between the two nuclear levels. If this difference is less than 511 keV, the β^+ decay is impossible and only electron capture occurs.

Another process involving the orbital electrons is the 'internal conversion'. It is a different decay mechanism for transitions that usually emit gamma rays. In this process the nuclear excitation energy is directly transferred to the atomic electron rather than to a gamma ray. Unlike the electron produced in a β^- decay, the electron produced by the internal conversion process always has the same energy. The electron that is most likely to be involved in the internal conversion process is the K-shell electron, but the electrons in other orbitals may also receive the conversion energy. An isotope decaying by the internal conversion process will therefore exhibit

a group of electron energies, the differences in energy being equal to the differences in the binding energies between the electronic orbitals.

1.6.5 The Radioactive Decay Law

The probability that a nucleus decays in some small time interval dt is given by

$$P = \lambda \, dt$$

where λ is a constant characteristic of the decay. The important fact is that this probability is independent of how long the nucleus is already waiting to decay. Assume there are N_0 nuclei at time $t = 0$, and let us denote the number of nuclei at any time after $t = 0$ by $N(t)$. We have

$$\frac{dN(t)}{dt} = -N(t) \, \lambda dt$$

The solution of this differential equation with the boundary condition $N(0) = N_0$ is

$$N(t) = N_0 \, e^{-\lambda t}$$

The normalised probability density function for the observation of a decay after a time t is

$$f(t) = \lambda e^{-\lambda t} dt$$

The average decay time of the isotope is given by

$$\tau = \langle t \rangle = \int\limits_0^\infty t \, \lambda e^{-\lambda t} \, dt = \frac{1}{\lambda}$$

The decay law can therefore be written as

$$N(t) = N_0 \, e^{-\frac{t}{\tau}}$$

In nuclear physics it is customary to characterise a decay by its 'half-life' $T_{1/2}$, rather than by the average decay time τ. The half-life is defined as the time it takes for half of the original nuclei to decay. We therefore have

$$\tfrac{1}{2} = e^{-\frac{T_{1/2}}{\tau}}$$
$$T_{1/2} = \tau \ln 2$$

Until now we have only considered the decay of one isotope into a stable final state. In many cases the decay process under consideration is part of a chain of decays, and the number of each type of isotope increases because new isotopes are added by the decay of some parent isotope and decreases because of its own decay. In that case one should consider the abundance of all the isotopes involved in the chain and a much more complicated expression is obtained.

1.6.6 The Nuclear Level Diagram

A very useful tool for understanding the nuclear decay mechanism is the nuclear level diagram illustrated in Fig. 1.7. In this diagram the *x*-direction represents the charge of a nucleus, and the *y*-direction the energy of the nuclear levels. A given isotope has a fixed number of protons and neutrons. The different energy levels of this isotope are represented as short horizontal lines in the level diagram. The gamma emission process and the electron conversion process correspond to a transition between two levels that are situated one above the other in the level diagram, and this emission is represented by a vertical arrow pointing downwards. A β^- or a β^+ decay is represented by arrows pointing in the down-right or down-left direction.

Annex 6 lists a number of isotopes with decay modes that make them useful as radioactive sources for a variety of applications in nuclear and particle physics. For the nuclei undergoing a β^+ decay, the positron comes to rest after a short range in matter and annihilates with an electron into two gamma rays of 511 keV emitted back to back.
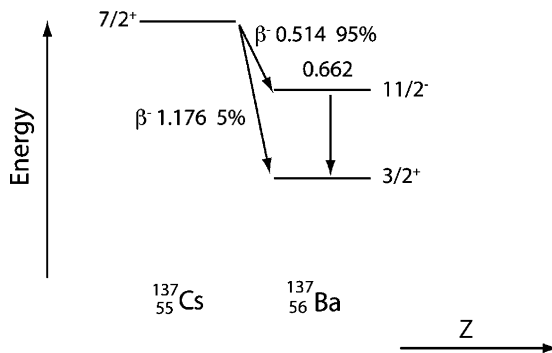
Fig. 1.7   Nuclear level diagram for the ^{137}Cs decay. The numbers to the side of the level (e.g. 3/2$^+$) represent the spin parity of the level. The number above the line representing the level is the energy in MeV relative to the ground state. The numbers next to the symbol β^- represent the end point energy of the electron and the fractional probability of the transition. There are two competing processes for the transition from the excited ^{137}Ba level to the ground state. About 90% of the transitions give rise to a gamma of 662 keV, while about 10% proceed through internal conversion. Therefore only 85% of the ^{137}Cs decays give rise to a gamma ray of 662 keV

1.7 Exercises

1. Show that the Lorentz transformation is such that the velocity of a light ray travelling in the x direction is the same for the observer in the frame S and for the observer in the frame S'.
2. What is the mean free path before decay for a charged pion with a kinetic energy of 1 GeV?
3. Show that the relativistic expression for the kinetic energy of a particle (Eq. 1.2) reduces to the non-relativistic expressions if the velocity of the particle is small compared to the velocity of light.
4. For a Poisson distribution with average value 16, calculate the probability to observe 12, 16 and 20 as measured value. Calculate the probability density function for a Gaussian distribution with average value 16 and dispersion 4, for the values $x = 12$, 16 and 20. Compare the results.
5. Consider a very short-lived particle of mass M decaying into two long-lived particles 1 and 2. Assume you can measure accurately the energies and momenta of the two long-lived particles. How will you calculate the mass of the short-lived particle from the known energies and momenta of the two long-lived objects?
6. Calculate the order of magnitude of the energy levels in atoms and in nuclei using the 'particle in a box' approximation, Eq. (1.9). Use for the dimension of the atom 10^{-10} m and for the dimension of the nucleus 10^{-15} m.
7. Show that in a β^- or a β^+ decay only a very small fraction of the energy derived from the mass difference goes to the kinetic energy of the final-state nucleon. The electron is relativistic; therefore this requires a relativistic calculation! Hint: the 3-body problem can be reduced to a 2-body problem by considering the electron–neutrino system as one object with a mass of a few MeV.

References

1. B. Povh, K. Rith, C. Scholz and F. Zetsche, Particles and nuclei, an introduction to the physical concepts, Springer (2006).
2. F. Halzen and A.D. Martin, Quarks and leptons: an introductory course in modern particle physics, John Wiley & Sons (1984).
3. W.E. Burcham and M. Jobes, Nuclear and particle physics, Longman Scientific&Techical (1994).
4. A Handbook of Radioactivity Measurements Procedures, National council of radiation protection and measurement, Bethesda, Maryland, 2nd edition (1985).
5. G.F. Knoll, Radiation detection and measurement, 3rd edition, John Wiley & Sons (2000).
6. C. Amsler et al., Review of particle physics', Phys. Lett. B 667, 1 (2008), This document is downloadable from the WEB at: http://pdg.lbl.gov/pdg.html.
7. W.R. Leo, Techniques for nuclear and particle physics experiments, Springer (1992).
8. Nuclear Data Center tables, Brookhaven national laboratory, available from the WEB at http://www.nndc.bnl.gov/.
9. National Institute of Standards and Technology, Physical reference data, http://physics.nist.gov/PhysRefData/contents.html
10. Kaye and Laby, Tables of Physical and Chemical constants, National physical laboratory, available on the web at http://www.kayelaby.npl.co.uk/
11. W.T. Eadie, D. Drijard, F.E. James, M. Roos and B. Sadulet, Statistical methods in experimental physics, 3rd edition, North Holland (1988).

Chapter 2
Interactions of Particles in Matter

The aim of this chapter is to introduce the reader to the different ways subatomic particles interact with matter. For a more in depth discussion of the subject, for references to the original literature, and for a derivation of many of the formula quoted in this chapter, see Ref. [4]. Reference [6, 9] of Chap. 1 are useful web resources containing extensive numerical data on the interactions of subatomic particles in matter.

2.1 Cross Section and Mean Free Path

If one of the particles in Table 1.4 travels in any piece of material, it will have a certain probability to interact with the nuclei or with the electrons present in that material. In a very thin slice of matter, this probability is obviously proportional to the thickness of the slice and to the number of potential target particles per unit volume in the material. Furthermore, it will depend on the nature of the interaction. That intrinsic part of the probability is expressed with the help of the quantity 'cross section'. The cross section is the convenient quantity to discuss the interactions of particles in matter. If a particle crosses perpendicularly through an infinitesimally thin slice of matter, the probability to interact and the cross section σ are related by Eq. (2.1). This equation is the definition of the cross section and it is illustrated in Fig. 2.1.

$$dW = dx \, N \, \sigma \qquad (2.1)$$

In this equation, dW is the probability to undergo an interaction of a certain type, dx is the thickness of a very thin section of the material and N is the number of scattering centres per unit volume. The cross section has the dimensions of a surface. In nuclear and particle physics, the commonly used units for the cross section are the barn and cm^2, with 1 barn $= 10^{-24}$ cm^2. It is easy to see that, in classical mechanics, the cross section for the collision of a point particle with a hard sphere is just be the surface of a section through the middle of the sphere. This explains the name 'cross section'.

If a beam of particles enters a slab of material, the number of affected particles in the beam will increase due to the collisions of these beam particles with the nuclei

dx

N scatter centres
per unit volume

particle trajectory
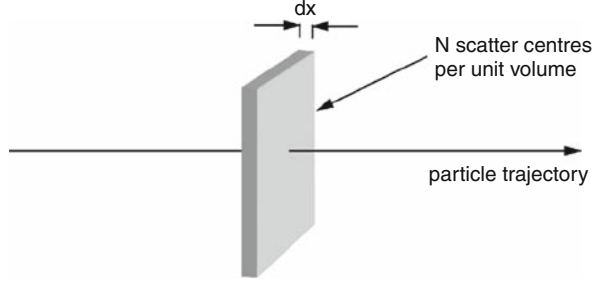
or electrons present in the material. To describe this mathematically, let us define
$P(x)$ as the probability that a particle has interacted after travelling a distance x in
the medium. Obviously, we have $P(0) = 0$. From the definition of the cross section
(Eq. 2.1), we know that $P(x + \Delta x)$ and $P(x)$ are related by

$$P(x + \Delta x) = P(x) + [1 - P(x)] N\sigma \, \Delta x$$

$$\frac{P(x + \Delta x) - P(x)}{\Delta x} = [1 - P(x)] N\sigma$$

In this expression Δx represents some small distance in the x direction. Taking
the limit $\Delta x \to 0$, we obtain that $P(x)$ satisfies the following differential equation:

$$\frac{dP(x)}{dx} = [1 - P(x)] N\sigma$$

$$\frac{d[1 - P(x)]}{dx} = -[1 - P(x)] N\sigma$$

The solution of this differential equation, with the boundary condition
$[1 - P(0)] = 1$, is

$$[1 - P(x)] = e^{-xN\sigma}$$

The probability density function for the interaction of a particle after a travelling
distance x in the medium is given by

$$W(x) = [1 - P(x)] N\sigma = e^{-xN\sigma} N\sigma$$

Therefore, the mean free path λ of a particle before the first collision is given by

$$\lambda = \int_0^\infty W(x) x \, dx = \int_0^\infty e^{-xN\sigma} xN\sigma \, dx$$

$$= \frac{1}{N\sigma} \int_0^\infty e^{-x} x \, dx = \frac{1}{N\sigma}$$

If the material contains two different types of scattering centres, X and Y, the above discussion generalises to

$$\lambda = \frac{1}{N_x \sigma_x + N_y \sigma_y}$$

and

$$\frac{1}{\lambda} = \frac{1}{\lambda_X} + \frac{1}{\lambda_Y}$$

$$\lambda_X = \frac{1}{N_X \, \sigma_X}; \lambda_Y = \frac{1}{N_Y \, \sigma_Y} \tag{2.2}$$

N_X and N_Y are the number of scattering centres of each type per unit volume.

If we consider collisions on the nuclei of atoms, N represents the number of atoms per unit volume. The relative atomic weight A_r of an element is defined as the average weight of the atoms divided by 1/12th of the weight of carbon. 'A_r' gram of an element contains N_A scattering centres, where N_A is the number of Avogadro. One gram of the material contains N_A/A_r atoms, and one cubic metre contains $\rho N_A/A_r$ atoms. We thus have

$$N = \frac{\rho \, N_A}{A_r}$$

A particle can have different ways to interact. For example a proton can scatter elastically from a nucleus, or it can scatter and bring the nucleus in an excited state. The cross section corresponding to a particular type of interaction is called a partial cross section, and the sum of all partial cross sections is the total cross section. One can also consider the partial cross section where the proton is scattered in a particular direction. This is called a differential cross section and this is usually written as $d\sigma/d\Omega$, where $d\Omega = \sin\theta d\theta d\varphi$. The total cross section is then given by

$$\sigma_{tot} = \int \frac{d\sigma}{d\Omega} d\Omega$$

2.2 Energy Loss of a Charged Particle due to Its Interaction with the Electrons

When a charged particle penetrates in matter, it will interact with the electrons and nuclei present in the material through the electromagnetic force. If the charged particle is a proton, an alpha particle or any other charged hadron (discussed in Chap. 1), it can also undergo a nuclear interaction and this will be discussed in Sect. 2.5. In the present section we ignore this possibility. If the particle has 1 MeV or more as energy, as is typical in nuclear phenomena, the energy is large compared to the binding energy of the electrons in the atom. To a first approximation, matter can be

seen as a mixture of free electrons and nuclei at rest. The charged particle will feel the electromagnetic fields of the electrons and the nuclei and in this way undergo elastic collisions with these objects.

The interactions with the electrons and with the nuclei present in matter will give rise to very different effects. Let us assume for the sake of definiteness that the charged particle is a proton. If the proton collides with a nucleus, it will transfer some of its energy to the nucleus and its direction will be changed. The proton is much lighter than most nuclei and the collision with a nucleus will cause little energy loss. It is easy to show, using non-relativistic kinematics and energy–momentum conservation, that the maximum energy transfer in the elastic collision of a proton of mass 'm' with nucleus of mass 'M' is given by (see Sect. 7.3, Eq. 7.1):

$$\Delta E_{\max} = \frac{1}{2}mv^2 \left(\frac{4\,mM}{(m+M)^2} \right)$$

If the mass of the proton m is much smaller than the mass of the nucleus M, we therefore have

$$\Delta E_{\max} \approx \frac{1}{2}mv^2 \left(4\frac{m}{M} \right) \qquad (m << M)$$

In the limit that the mass of the nucleus goes to infinity, no energy transfer is possible.

In a collision with a nucleus the proton will lose little energy, but its direction can be changed completely; it can even bounce backwards. In collisions with electrons, on the other hand, a large amount of energy can be transferred to the electrons, but the direction of the proton can only be slightly changed. Indeed, there is a maximum possible kinematical angle of deviation in such collisions. It needs a relativistic calculation to derive this angle. As a result, most of the energy loss of the proton is due to the collisions with the electrons, and most of the change of direction is due to the collisions with the nuclei.

A proton, and more generally any charged particle, penetrating in matter leaves behind a trail of excited atoms and free electrons that have acquired some energy in the collision. The energy distribution of these electrons is

$$\frac{dn}{dE} \propto \frac{1}{E^2}$$

Most of these electrons have only received a very small amount of energy. However, some of the electrons acquire sufficient energy to travel macroscopic distances in matter. These high-energy electrons are sometimes called δ-electrons. These have sufficient energy themselves to excite or ionise atoms in the medium. This type of energy loss due to the interaction of the charged particle with electrons is often referred to as 'energy loss due to ionisation'. This is strictly speaking not correct since many atoms are only brought to an excites state, not ionised. Figure 2.2 illustrates the passage of a charged particle in matter and shows some of the ionisation electrons.
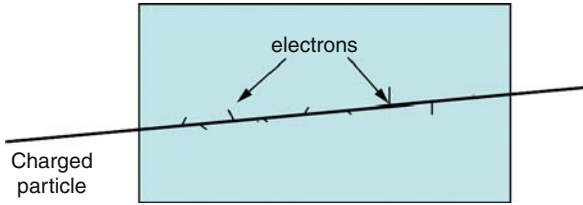
Fig. 2.2 A charged particle penetrates in matter. It loses energy by transferring a small amount of energy to each of a large number of electrons along its trajectory. Some of these electrons have enough energy to travel a macroscopic distance, and also cause further ionisation along their trajectory

When discussing the biological effects of radiation the term 'Linear Energy Transfer' (LET) is often used to refer to the energy loss of charged particles. The linear energy transfer is defined as the amount of energy transferred, per unit track length, to the *immediate vicinity* of the trajectory of the charged particle. For heavy- and low-velocity particles, the energy loss per unit track length and the LET are the same. For light and fast particles, however, the two quantities differ considerably. Part of the energy loss of an electron of several MeV is used to eject energetic δ-electrons from the atoms in the medium. These energetic electrons do not deposit their energy in the immediate vicinity of the track and therefore do not contribute to the LET.

The energy loss of a high-energy charged particle in matter due to its interactions with the electrons present in the matter is given by the Bethe-Bloch equation:

$$\frac{dE}{dx} = \rho \frac{Z_{\text{nucl}}}{A_{\text{r}}} (0.307 \, \text{MeVcm}^2/\text{g}) \frac{Z^2}{\beta^2} \left[\frac{1}{2} \ln \left(\frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} \right) - \beta^2 - \frac{\delta(\beta)}{2} \right] \tag{2.3}$$

See for example Ref. [4] for the derivation of this equation. The symbols used in the above equation are defined below:

dE/dx = energy loss of particle per unit length
Z = charge of the particle divided by the proton charge
c = velocity of light
$\beta\gamma$ = relativistic parameters as defined in Sect. 1.3
ρ = density of the material
Z_{nucl} = dimensionless charge of the nuclei
A_{r} = relative atomic weight
I = mean excitation energy in eV. Parameter usually determined experimentally. It is typically around (10 eV times Z_{nucl})
$T_{\max}$ = maximum energy transfer to the electron. For all incoming particles except the electron itself this is to a good approximation given by $\approx 2 \, m_e c^2 \beta^2 \gamma^2$. For electrons $T_{\max}$ is the energy of the incoming electron.
$\delta\beta$ = density-dependent term that attenuates the logarithmic rise of the cross section at very high energy. See (Ref. [6] in Chap. 1) for a discussion of this term.

Fig. 2.3 Energy loss in air
vs. the kinetic energy for
some charged particles.
Figure calculated using
Eq. (2.3)
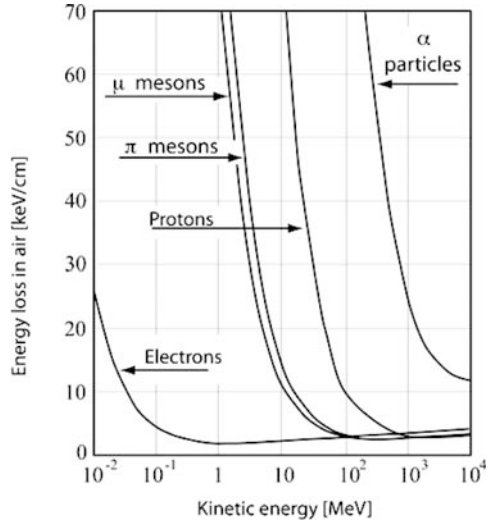
For the purpose of a qualitative discussion the Bethe–Bloch equation can be approximated as

$$\frac{dE}{dx} \approx \rho \, (2 \, \text{MeVcm}^2/\text{g}) \frac{Z^2}{\beta^2} \qquad (2.4)$$

If the density is expressed in g/cm^3, the energy loss is in units MeV/cm. In the literature, the term 'energy loss' sometimes refers to the loss divided by the density. In the latter case, the energy loss has the units MeV cm^2/g. For electrons with energy of more than 100 keV, the velocity is close to the velocity of light ($\beta \approx 1$), and the energy loss is about 2 MeV/cm multiplied by the density of the medium.

For all particles, the energy loss decreases with increasing energy and eventually reaches a constant, energy-independent value. That value is approximately the same for all particles of unit charge (see Fig. 2.3).

For alpha particles the velocity is usually much less than the velocity of light, and the energy loss is much larger. However, the Bethe–Bloch equation is valid only if the velocity of the particle is much larger than the velocity of the electrons in the atoms, and for alpha particles, this condition is usually not satisfied. The velocity of electrons in atomic orbits is of the order of 1% of the velocity of light. For particle velocities that are small compared to the typical electron velocities in the atoms, the energy loss increases with the energy and reaches a maximum when the particle velocity is equal to the typical electron velocity. After this maximum, the energy loss decreases according to the Bethe–Bloch equation. This behaviour is illustrated in Figs. 2.4 and 2.13.

Since particles lose energy when travelling in a medium, they will eventually have lost all their kinetic energy and come to rest. The distance travelled by the

Fig. 2.4 Energy loss of alpha particles divided by the density as a function of the alpha particle energy in different materials. Figure adapted from [1]
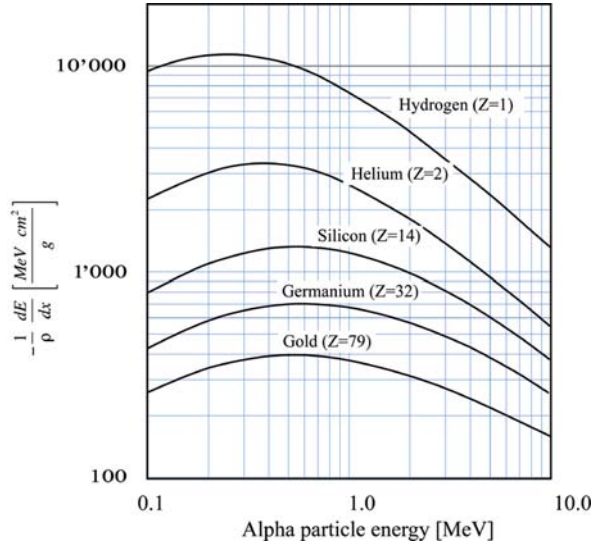
particles is referred to as the range. As the particle penetrates in the medium, its energy loss per unit length will change. The energy loss of a particle as a function of its distance of penetration is illustrated in Fig. 2.5. The energy loss increases towards the end of the range. Close to the end it reaches a maximum and then abruptly drops to zero. This maximum of the energy loss of charged particles close to the end of their range is referred to in the literature as the 'Bragg peak', and the variation of the energy loss with the residual energy as the 'Bragg curve'. However, all the particles with a given kinetic energy do not have exactly the same range. This is due to the statistical nature of the energy loss process. There are fluctuations on the range called range straggling.
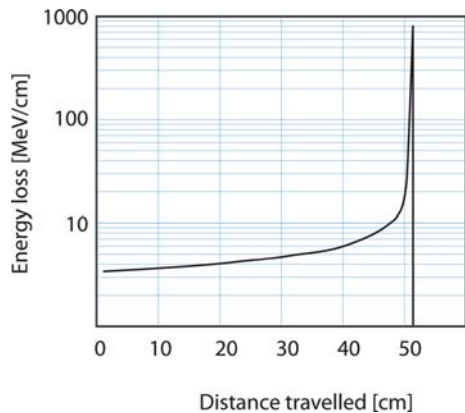
Fig. 2.5 Energy loss of a proton of 300 MeV along its trajectory in water. The energy loss increases towards the end of the range, reaches a maximum and rather abruptly drops to zero just before the particle stops. The data for this figure were obtained from Ref. [9] in Chap. 1

Fig. 2.6 Range of protons and alpha particles in silicon as a function of their kinetic energy. The data for this figure were obtained from Ref. [9] in Chap. 1
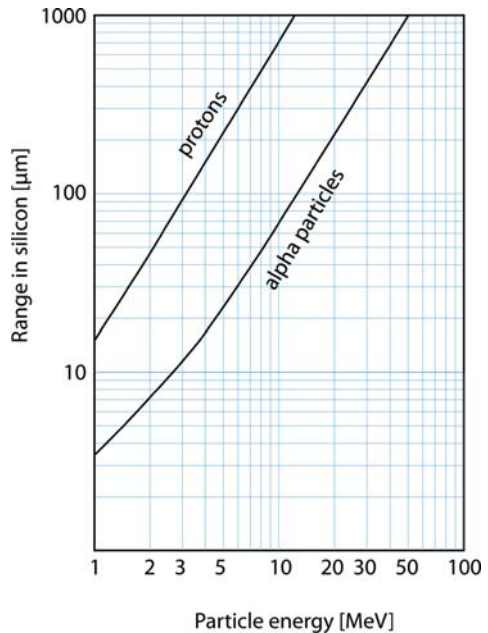
Heavy nuclear fragments produced by nuclear fission are also energetic charged particles but behave somewhat differently from alpha particles. Nuclear fragments tend to pick up electrons as they travel in the medium. Therefore they behave as particles with a charge that is smaller than the charge of the fragment itself. As they slow down, the fragments pick up more and more electrons, and the energy loss decreases rather than increases. For alpha particles, this electron pick-up only occurs at the very end of the range.
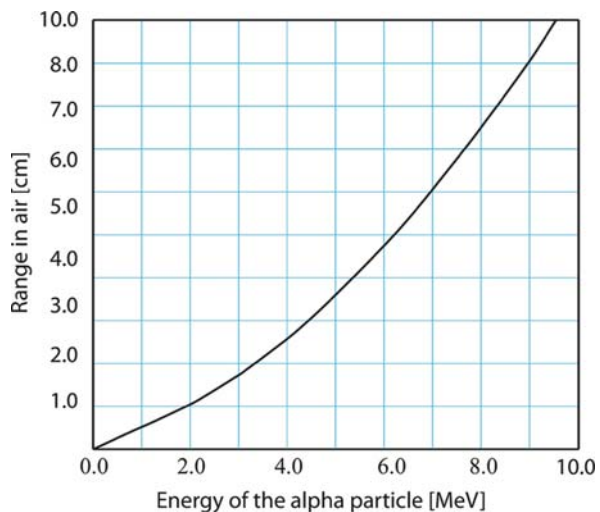
Fig. 2.7 Range–energy plot for alpha particles in dry air at 20°C and standard pressure. The data for this figure were obtained from Ref. [9] in Chap. 1

$$LET = \frac{1}{\rho}\frac{dE}{dx}$$

Figures 2.6 and 2.7 illustrate the range of particles in air and silicon.

2.3 Other Electromagnetic Interactions of Charged Particles

Multiple scattering. The collisions of charged particles with the nuclei will cause the charged particle to change direction. This is illustrated in Fig. 2.8. Such erratic changes in the direction of a particle along its trajectory is called direction straggling or multiple scattering. For small angles of deviation, this change in angle is more or less Gaussian and the root mean square (r.m.s.) direction deviation of a particle traversing a thickness L of material is given by

$$\sqrt{\langle \Theta^2 \rangle} = \frac{Z}{Pc\beta}(20 \text{ MeV})\sqrt{\frac{L}{X_0}}$$

$$\frac{1}{X_0} \approx 4\alpha \, r_0^2 \frac{\rho N_A}{A_r} Z_{\text{nucl}}(1 + Z_{\text{nucl}}) \ln\left(\frac{183}{\sqrt[3]{Z_{\text{nucl}}}}\right)$$

(2.5)

In this equation X_0 represents the radiation length. This is a quantity that characterises how charged particles or gamma rays interact in a material. It depends on the density and the charge of the nucleus. The simple analytical expression for the radiation length given in Eq. (2.5) is only an approximation. A more exact but much more complicated expression is given in (Ref. [6] in Chap. 1).

The definitions of the symbols used in the multiple scattering formula and in the expression for the radiation length are given below. The other symbols have the same meaning as in Eq. (2.3).

Θ = scattering angle relative to the incoming particle in radians
P = momentum of the incoming particle
X_0 = radiation length of the material
N_A = Avogadro's number
α = fine structure constant ($\alpha \approx 1/137$)
r_0 = classical electron radius ($2.82 \ 10^{-15}$ m)

Notice that Θ represents the angle in space. The symbol Θ_p represents the angle projected on a plane containing the direction of the incoming particle. These two quantities are related by

$$\sqrt{\langle \Theta_p^2 \rangle} = \frac{1}{\sqrt{2}}\sqrt{\langle \Theta^2 \rangle}$$

Table 2.1 lists the radiation length for some common materials.

Fig. 2.8 A charged particle
traversing a slice of matter
will change direction due to
multiple scattering on atomic
nuclei
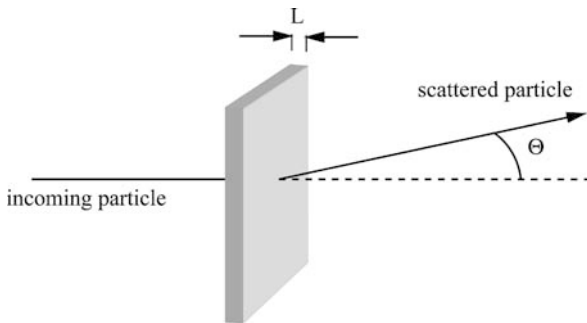
Table 2.1 Radiation length X_0 for some common materials

Material	Radiation length X_0
Air	304 m
Water	36 cm
Shielding concrete	10.7 cm
Nylon	36.7 cm
Aluminium (Al)	8.9 cm
Silicon (Si)	9.36 cm
Iron (Fe)	1.76 cm
Lead (Pb)	0.56 cm
Uranium (U)	0.32 cm

At nuclear energies, the momentum of the particles is of the order of $Pc \approx 1$ MeV,
and particles will, on average, scatter over a very large angle in one radiation length.
After one radiation length the information about the original direction is essentially
lost. However, alpha particles or protons of a few MeV have a range that is only a
very small fraction of the radiation length. Hence, they will stop before they have
scattered over a large angle. Electrons, on the other hand, can penetrate to a sig-
nificant depth in the material, and electrons will therefore be strongly affected by
multiple scattering. Figure 2.9 shows typical trajectories for an electron, a proton
and an alpha particle of 10 MeV in silicon.

The path of an electron in matter can be several centimetres, but the distance
travelled according to a straight line is usually much shorter than the actual length of
the trajectory. Electrons do not have a well-defined range. The number of electrons

Fig. 2.9 A typical trajectory
for an electron, a proton and
an alpha particle of 10 MeV
in silicon. The electron
trajectory is drawn on a scale
10 times smaller than the
trajectory of the proton and
the alpha particle
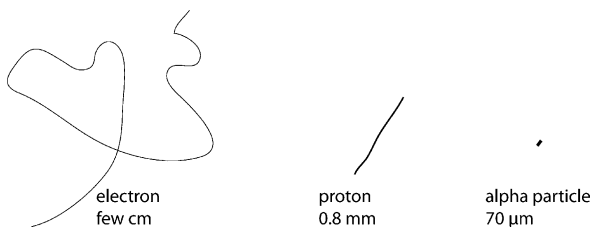
that can penetrate through a slice of material will decrease more or less linearly with the thickness of the slice of material.

Cherenkov effect. The Cherenkov effect is a light emission effect that occurs whenever a charged particle travels in a medium faster than the speed of light in that medium. In a medium with optical index of refraction 'n', the velocity of light is c/n. Typical values for the refractive index in liquids or solids are around 1.5, and the velocity of light in these materials is about 66% the speed of light. The Cherenkov effect is somewhat similar to the bow wave that accompanies a speedboat in water, or the 'supersonic bang' of a plane going at a speed faster than the speed of sound.

This effect is illustrated in Fig. 2.10. This phenomenon is easily understood by following the Huygens' principle used to explain optical and acoustical phenomena. If a charged particle travels in a medium, the electric field of the charged particle will polarise the medium. After the particle has passed, the medium returns to its original unpolarised state. This change of polarisation condition in the medium represents an electromagnetic perturbation that will propagate in space at the speed of light. The left-hand side of Fig. 2.10 shows the case where the particle travels at a speed lower than the speed of light in the medium. The small electromagnetic perturbations caused by the polarisation and depolarisation of the medium propagate faster than the particles. At any point in space far away from the particle's trajectory, these perturbations arrive randomly and annihilate each other. The right-hand side of Fig. 2.10 shows the case where the particle travels at a speed faster than the speed of light in the medium. The small electromagnetic perturbations caused by the polarisation and depolarisation of the medium propagate less rapidly than the particles. All the elementary perturbations unite together in one wavefront. The phases between all these elementary perturbations are not randomly distributed. They add up together to produce a finite perturbation. This perturbation represents a wave
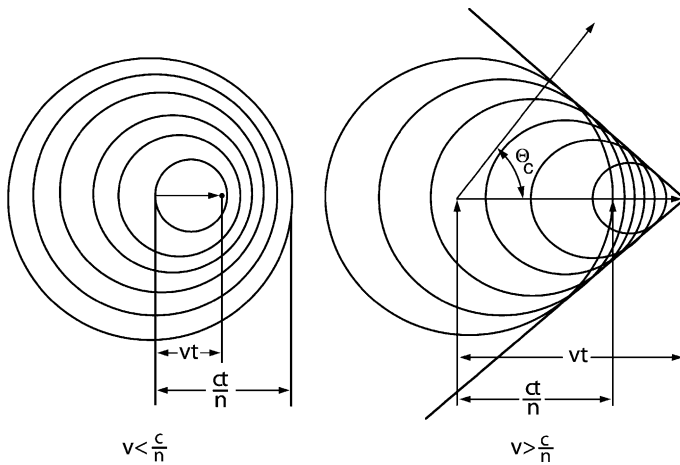
Fig. 2.10 (*Left*) A particle is travelling at a speed lower than the speed of light in the medium. (*Right*) A particle is travelling at a speed greater than the speed of light in the medium

travelling in the direction fixed by the speed of the particle and the speed of light in the medium.

From the geometry of the problem, we can easily derive the value of the angle between the particle and the wave. To find this angle, consider the right-angled triangle shown in the left-hand side of Fig. 2.10. Two sides of this triangle are of length ct/n and vt, respectively. We therefore have

$$\cos(\theta_c) = \frac{(c/n)t}{v\,t} = \frac{c}{nv} \tag{2.6}$$

The Cherenkov effect thus consists of the emission of optical photons in the direction given by Eq. (2.6). A similar situation prevails when an airplane is flying at supersonic speed. It is accompanied by a loud acoustical 'bang' that propagates in a direction given by a similar equation.

The intensity of the Cherenkov effect can be calculated from first principles by solving the Maxwell equations with the proper boundary conditions. The result of this calculation is

$$\frac{d^2E}{d\hbar\omega.dx} = \hbar\omega\frac{Z^2\alpha}{\hbar c}\left[1 - \frac{c^2}{n^2v^2}\right] \quad v > \frac{c}{n}$$
$$\frac{d^2E}{d\hbar\omega.dx} = 0 \qquad\qquad v < \frac{c}{n}$$

In the above equation, the notation is as follows:

Z = charge of the particle in units 'proton charge'
E = energy emitted in the form of optical photons
n = optical refractive index
c = velocity of light in vacuum
v = velocity of the particle
$\hbar\omega$ = energy of the emitted photon
α = fine structure constant (1/137)
$\hbar c$ = numerical constant of value $197\ 10^{-9}$ eV m

Dividing Eq. (2.7) by $\hbar\omega$ gives the number of Cherenkov photons produced per-photon-energy interval and per-unit-length. A high-energy electron produces about 220 photons/cm in water ($n = 1.33$) and about 30/m in air, in the visible part of the spectrum.

From Eq. (1.4), we derive that a charged particle will emit Cherenkov radiation if the kinetic energy exceeds the threshold value given by

$$E_{\text{threshold}} = mc^2\left(\sqrt{\frac{n^2}{n^2-1}} - 1\right) \tag{2.8}$$

The threshold for the Cherenkov effect of electrons in water is 264 keV. For protons the threshold is 486 MeV. At nuclear energies, only electrons can acquire
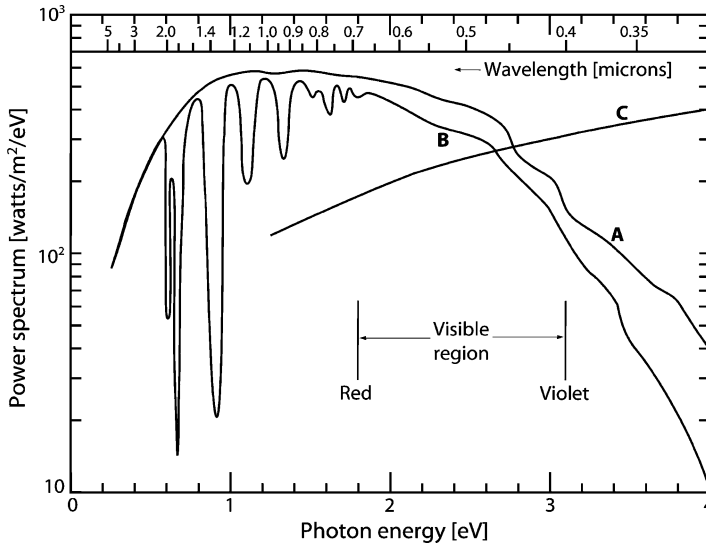
Fig. 2.11 (**A**) Energy spectrum of sunlight above the atmosphere, (**B**) Energy spectrum of sunlight at sea level with the Sun at its zenith. (**C**) Energy spectrum of Cherenkov emission. Figure reproduced from [1], with permission

a speed that exceeds the speed of light in a medium, and therefore emit Cherenkov radiation.

Comparing the energy spectrum of Cherenkov light with the spectrum of solar light, we see that the Cherenkov radiation contains more energy in the blue part of the spectrum; therefore, Cherenkov radiation appears as blue light (see Fig. 2.11). Cherenkov radiation is causing the characteristic blue glow in the water surrounding the core of a water pool reactor, as illustrated in Fig. 2.12.

The Cherenkov effect only represents a small loss of energy compared to the energy loss due to ionisation considered before. It is nevertheless an interesting effect because it depends only on the velocity of the particle. If one knows the energy or the momentum of a particle by other means, measuring the Cherenkov effect allows knowing the mass, and therefore the nature of that particle.

Transition radiation. This is a weak effect somewhat similar to the Cherenkov effect. It is also due to the polarisation of the medium by the charged particle. It depends on the plasma frequency in the material. The plasma frequency is usually expressed as a quantity with dimension energy, and it is given by the equation

$$\hbar\omega_{\mathrm{p}} = \sqrt{4\pi N_e r_0^3 m_e c^2 / \alpha}$$

In this equation N_e is the number of electrons per unit volume in the material, r_0 the classical electron radius and α the fine structure constant. The plasma frequency $\hbar\omega_{\mathrm{p}}$ is about 30 eV for materials with density 1.
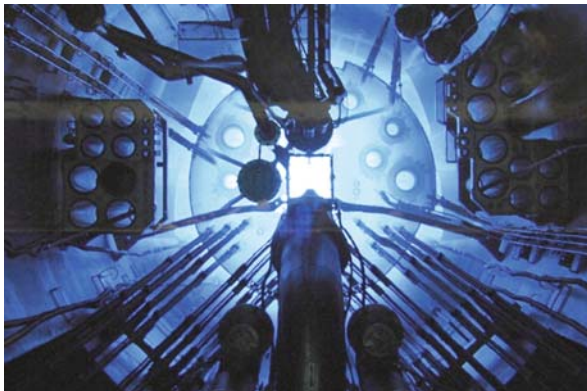
Fig. 2.12 The Cherenkov effect is causing the blue glow in the water surrounding the core of a water pool reactor (here the OPAL reactor, Australian Nuclear Science and Technology Organisation). Gamma rays originating from the reactor core convert to electron–positron pairs in the water. Many of these electrons or positrons travel at speeds exceeding the speed of light in water and thus produce Cherenkov radiation. Photograph courtesy of the Australian Nuclear Science and Technology Organisation

Whenever a particle of charge Z traverses a boundary between vacuum and some material, a small amount of energy is emitted as energetic photons. The power spectrum of these photons is logarithmically divergent at low energy and decreases rapidly for $\hbar\omega > \gamma\hbar\omega_p$, where γ is the relativistic γ factor of the particle. About half of the energy is emitted in the form of photons with energy in the range $\gamma\hbar\omega_p/10$ to $\gamma\hbar\omega_p$. The typical energy of the emitted photons is given by

$$E_{\text{typical}} = \gamma\hbar\omega_p/4$$

The average number N_γ of photons with energy larger than $\gamma\hbar\omega_p/10$ is

$$N_\gamma \approx 0.8\,\alpha\,Z^2 \approx 0.59\%Z^2$$

The total energy emitted by this effect when a charged particle traverses a boundary between vacuum and a medium is given by

$$E = \alpha\,Z^2\,\gamma\hbar\omega_p/3$$

We see that the probability to emit energetic transition radiation photons, and the total amount of energy emitted by this effect, is indeed quite small. But if a charged particle penetrates through a stack with a large number of thin foils, or through material with a foam-like structure, there can be a large number of transitions, and the effect becomes significant. All the above equations are valid only if the thickness of the foils, and the thickness of the gaps between the foils, is larger than a 'formation length' given by $\gamma c/\omega_p$. For $\gamma = 1,000$ this 'formation length' is about 10 μm.

The energy emitted is proportional to γ, and for particles with $\gamma = 1,000$ the energy of the photons is in the soft X-ray region. This effect is interesting because it can be used for the identification of particles. It is most useful for the identification of electrons.

Also for thin foils the total amount of energy radiated by the transition radiation effect is much smaller than the amount of energy emitted by the bremsstrahlung effect discussed below. But the bremsstrahlung power spectrum is almost constant up to the energy of the radiating particle, and therefore much harder. In the range of the typical transition radiation energy, the number of photons emitted by transition radiations is much larger than the number of photons emitted by bremsstrahlung.

Bremsstrahlung. Any charged particle undergoing acceleration will emit electromagnetic radiation. If a high-energy charged particle deviates from its trajectory due to a collision with a nucleus, this collision is necessarily accompanied by electromagnetic radiation. The emission is strongly peaked in the direction of flight of the charged particles.

The intensity of the radiation emitted can be calculated from first principles using quantum electrodynamics. In the case of particles other than electrons or positrons this emission is negligible, except at very high energy. For electrons or positrons, the amount of radiation emitted is also governed by the quantity 'radiation length' X_0 introduced in Eq. (2.3). The average energy loss due to bremsstrahlung by an electron of energy E, in a thickness of matter dx, is given by

$$\frac{dE}{dx} = -\frac{E}{X_0} \tag{2.9}$$

In a thin foil the photons have a $1/E$ energy spectrum, and photons with energy of up to the total energy of the charged particle do occur. The power spectrum of the radiation is therefore a constant extending up to the energy of the radiating particle. For very high energy ($E > 1$ TeV) this power spectrum becomes peaked towards high energy. The emission of photons is a stochastic process, Eq. (2.9) giving only the average energy radiated.

We notice that the energy loss due to bremsstrahlung is proportional to the energy of the charged particle. For electrons, the energy loss due to bremsstrahlung exceeds the energy loss due to ionisation above some critical energy E_c. This critical energy depends on the nuclear charge of the atoms in the medium and is approximately given by $E_c = [800 \text{ MeV}]/(Z + 1.2)$. For a particle of mass M, other than an electron, bremsstrahlung is suppressed by a factor $(m_{\text{electron}}/M)^2$. Therefore, for all particles other than electrons or positrons, bremsstrahlung is negligible at energies below 1 TeV.

The average energy of an electron that is losing energy according to Eq. (2.8) is given by

$$E(x) = E_0 \, e^{-\frac{x}{X_0}}$$

In the above equation, x represents the distance travelled in the medium. In one radiation length, an electron of more than 10 MeV loses about half of its energy in the form of bremsstrahlung.

Overview of the electromagnetic interactions of charged particles. The electromagnetic interactions of charged particles with a kinetic energy in the range 100 keV to a few 10 MeV are summarised below.

Electrons: Electrons lose energy by exciting and ionising atoms along their trajectory. Per centimetre, electrons will lose about 2 MeV multiplied by the density. Electrons typically travel several centimetres before losing all their energy. The trajectories of electrons are erratically twisted due to multiple scattering. They will also lose a significant fraction of their energy by bremsstrahlung, particularly at higher energies. If the energy exceeds 264 keV, electrons show Cherenkov radiation in water.

Positrons: Positrons behave in exactly the same way as electrons except that, after coming to rest, a positron will annihilate with electrons that are always present. This annihilation gives rise to a pair of back-to-back gamma rays of 511 keV.

Alpha particles: The energy loss of alpha particles is much larger than that of electrons. It is of the order of 1000 MeV/cm times the density of the medium. As a result, alpha particles travel only tens of microns in solids and a few centimetres in gases. The trajectory of alpha particles is approximately straight.

Protons: Protons ionise much more than electrons but less than alpha particles. The range in solids is of the order of 1 mm. The trajectory of protons is approximately straight .

Nuclear fragments: Nuclear fragments show extremely high ionisation, and therefore the range of such nuclear fragments is typically only a few microns long.

For charged particles with a much larger energy than 10 MeV, the range before the particles have lost all their energy will be much greater. The energy loss will be of the order of 2 MeV/cm times the density of the medium for $Z = 1$ particles. A particle of 1 GeV will travel several metres in a solid before it has lost all its energy in excitation and ionisation of the atoms on its trajectory.

The energy loss of muons as a function of the momentum is illustrated in Fig. 2.13. The plot covers the whole range, from $E_{\text{kinetic}} \approx 1$ eV to 100 TeV. Remember that, if the energy of the particle is much larger than the mass, the energy and the momentum become the same. Since a muon is immune to the strong colour force, it will almost never undergo a nuclear interaction. A high-energy muon can travel several kilometres in solid matter before losing all its energy. If the energy of the muon is above 1 TeV (10^{12} eV) it will lose most of its energy through bremsstrahlung. If the energy is between a few 100 MeV and 1 TeV, it will lose energy at a rate of ≈ 2 MeV/cm times the density of the medium.

Around $P \approx 1$ MeV/c, the velocity of the muon becomes comparable to the velocity of electrons in atoms, and the Bethe–Bloch equation no longer holds. Starting from zero momentum, the energy loss will first increase, reach a maximum around $P = 1$ MeV/c, and then decrease again.
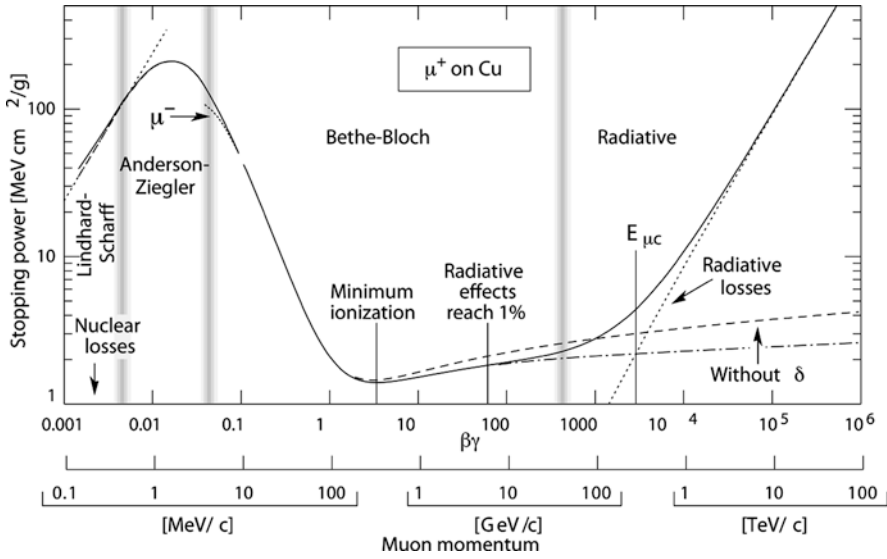
Fig. 2.13 Energy loss of a muon in copper between 100 keV and 100 TeV. Figure reproduced from Ref. [6] in Chap. 1, with permission

2.4 Interactions of X-Rays and Gamma Rays in Matter

X-rays and gamma rays are both high-energy photons. In the energy range 1–100 keV, these photons are usually called X-rays and above 100 keV they are usually called gamma rays. Some authors use the term 'gamma rays' to refer to any photon of nuclear origin, regardless of its energy. In these notes, I often use the term 'gamma ray' for any photon of energy larger than 1 keV. In the next few sections, the interactions of gamma rays with matter are discussed.

Photoelectric effect. If a charged particle penetrates in matter, it will interact with all electrons and nuclei on its trajectory. The energy and momentum exchanged in most of these interactions are very small, but together, these give rise to the different processes discussed in the previous chapter. When a photon penetrates in matter, nothing happens until the photon undergoes one interaction on one single atom. Gamma rays can interact with matter in many different ways, but the only three interaction mechanisms that are important for nuclear measurements are the photoelectric effect, the Compton effect and the electron–positron pair creation.

In the photoelectric absorption process, a photon undergoes an interaction with an atom and the photon completely disappears. The energy of the photon is used to increase the energy of one of the electrons in the atom. This electron can either be raised to a higher level within the atom or can become a free photoelectron. If the energy of gamma rays is sufficiently large, the electron most likely to intervene in the photoelectric effect is the most tightly bound or K-shell electron. The photoelectron then appears with an energy given by

$$E_{\text{kinetic}} = \hbar\omega - E_{\text{binding}}$$

In this equation, 'E_{binding}' represents the binding energy of the electron. In addition to a photoelectron, the interaction also creates a vacancy in one of the energy levels of the atom. This vacancy is quickly filled through rearrangement of the electrons; the excess energy being emitted as one or more X-rays. These X-rays are usually absorbed close to the original site of the interaction through photoelectric absorption involving less tightly bound electrons. Sometimes the excess energy is dissipated as an Auger electron instead of X-rays. In the Auger process, an electron from the outer shell falls into the deep vacancy, and another electron from the outer shell is expelled from the atom and takes up the excess energy.

The photoelectric effect is the dominant mode of interaction of the gamma rays of energy less than 100 keV. The energy dependence of the cross section is very approximately given by

$$\sigma \approx \text{Const} \frac{Z^n}{E_\gamma^{3.5}}$$

In this equation, Z represents the charge of the nucleus and E the energy of the X-ray. The coefficient 'n' varies between 4 and 5 over the energy range of interest. The photoelectric cross section is a steeply decreasing function of energy (see Fig. 2.17). Every time the photon energy crosses the threshold corresponding to the binding energy of a deeper layer of electrons, the cross section suddenly increases. Such jumps in the cross section are clearly visible in Figs. 2.17 and 2.18.

Compton scattering. Compton scattering is the elastic collision between a photon and an electron. This process is illustrated in Fig. 2.14. This is a process that can only be understood from the point of view of quantum mechanics.

A photon is a particle with energy $\hbar\omega$. From Eq. (1.1), we know that the photon has an impulse momentum $\hbar\omega/c$. Energy and momentum conservation constrain the energy and the direction of the final state photon. Using energy and momentum conservation, it is straightforward to show that the following relation holds (see Exercise 2):

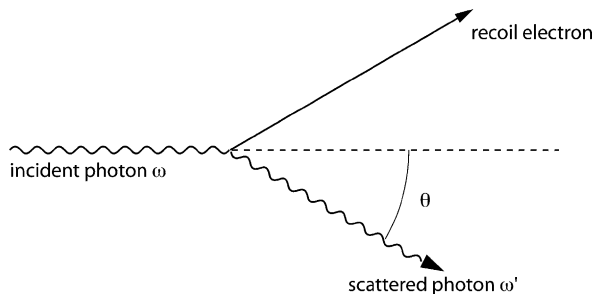$$\hbar\omega' = \frac{\hbar\omega}{\left(1 + \dfrac{\hbar\omega}{m_e c^2}(1 - \cos\theta)\right)} \tag{2.10}$$

Fig. 2.14 Illustration of Compton scattering and definition of the scattering angle θ

By using straightforward energy conservation, we ignore the fact that the electrons are not free particles but are bound in the atoms, and this will cause deviations from the simple expression above.

The value of the Compton scattering cross section for photon collisions on free electrons can only be derived from a true relativistic and quantum mechanical calculation. It is known as the Nishina–Klein formula (see Ref. [4] and references therein).

$$\frac{d\sigma}{d\Omega} = \frac{r_0^2}{2}\left(\frac{\hbar\omega'}{\hbar\omega}\right)^2\left(\frac{\hbar\omega}{\hbar\omega'} + \frac{\hbar\omega'}{\hbar\omega} - \sin^2\theta\right)\tag{2.11}$$

Equation (2.11) gives the differential cross section for the Compton scattering into a solid angle $d\Omega$. Integration over all angles gives the total cross section σ. The result of the integration is given in Ref. [4]. For energies either much larger or much smaller than the electron mass, a simple and compact expression for the total cross section is obtained.

$$\sigma = \frac{8\pi}{3}r_0^2 \qquad\qquad \hbar\omega << m_ec^2$$

$$\sigma = r_0^2\,\pi\,\frac{m_ec^2}{\hbar\omega}\left[\ln\left(\frac{2\hbar\omega}{m_ec^2}\right) + \frac{1}{2}\right] \hbar\omega >> m_ec^2$$

In these formulas, r_0 represents the classical electron radius introduced in Sect. 1.2. We see that for photon energies below the mass of the electron, the Compton cross section is independent of energy, and for photon energies above the electron mass, the cross section decreases as $(\text{energy})^{-1}$.

The Nishina–Klein formula only applies to scattering of gamma rays from free electrons. If the photon energy is much larger than the binding energy of electrons in atoms, the effects due to this binding are small.

If the gamma energy is small, there is a large probability that the recoil electron remains bound in the atom after the collision. The atom as a whole takes up the energy and the momentum transferred to the electron. In this case the interaction is called coherent Compton scattering or Rayleigh scattering. If the Compton interaction ejects the electron from the atom, the interaction is called incoherent Compton scattering.

The angular distribution of Compton scattering described by Eq. (2.11) is illustrated in Fig. 2.15. For photon energies much below the electron mass, the scattering is rather isotropic and back-scattering is about as likely as scattering in the forward direction. If the photon energy is much larger than the electron mass, the scattering is peaked into the forward direction.

Pair production. If the energy of the photon is at least two times larger than the mass of an electron, the energy of the photon can be used to create an electron and positron pair. This process is illustrated in Fig. 2.16. However, this reaction is not possible in empty space. In fact, energy and momentum cannot be conserved in this process. To see this, just imagine that the reaction gamma → electron + positron

Fig. 2.15 A polar plot of the cross section for Compton scattering. The curves show the magnitude of the differential cross section as a function of the scattering angle for different values of the incident photon energy. Figure calculated with Eq. (2.11).
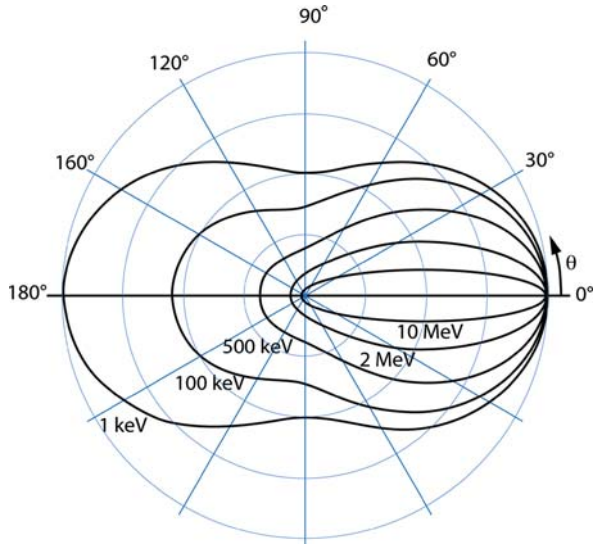
could take place. In that case it would be possible to go to the centre of mass system of the final state electron–positron pair. In that system, the sum of the impulse moments of the electron and the positron is zero. Therefore, the original gamma ray should have zero momentum. This is impossible. However, if the reaction gamma → electron + positron happens in the strong electric field of the nucleus, the nucleus can take up momentum, and in this way the energy and momentum can be conserved and the reaction becomes possible.

The cross section for pair production is given in Ref. [4]. It rises quickly from the threshold value of $2\, m_e$ to a constant value at high energy, as is illustrated in Fig. 2.17. The high-energy limit of the cross section is given by

$$\sigma = \frac{7}{9}\, 4\alpha r_0^2\, Z_{nucl}\, (Z_{nucl} + 1)\, \ln\left(\frac{183}{3\sqrt{Z_{nucl}}}\right)$$

The expression above reminds us of the quantity 'radiation length' X_0 introduced in Sect. 2.3. From the above expression of the cross section it immediately follows that, if a beam of high-energy photons penetrates in a medium, the number of unconverted gamma rays will decrease according to
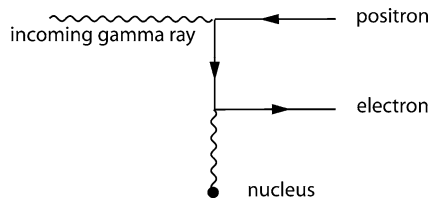
Fig. 2.16 An electron–positron pair can only be created if a certain amount of momentum can be exchanged with the nucleus

Fig. 2.17 Photon total cross sections as a function of energy in carbon and lead, showing the contributions of different processes: $\sigma_{p.e}$ = Atomic photoelectric effect (electron ejection, photon absorption); $\sigma_{Rayleigh}$ = Coherent scattering (Rayleigh scattering/atom neither ionised nor excited); $\sigma_{Compton}$ = Incoherent scattering (Compton scattering off an electron); k_{nuc} = Pair production, nuclear field; k_e = Pair production, electron field; $\sigma_{g.d.r}$ = Photonuclear interactions, most notably the Giant Dipole Resonance. Figure reproduced from Ref. [6] in Chap. 1, with permission
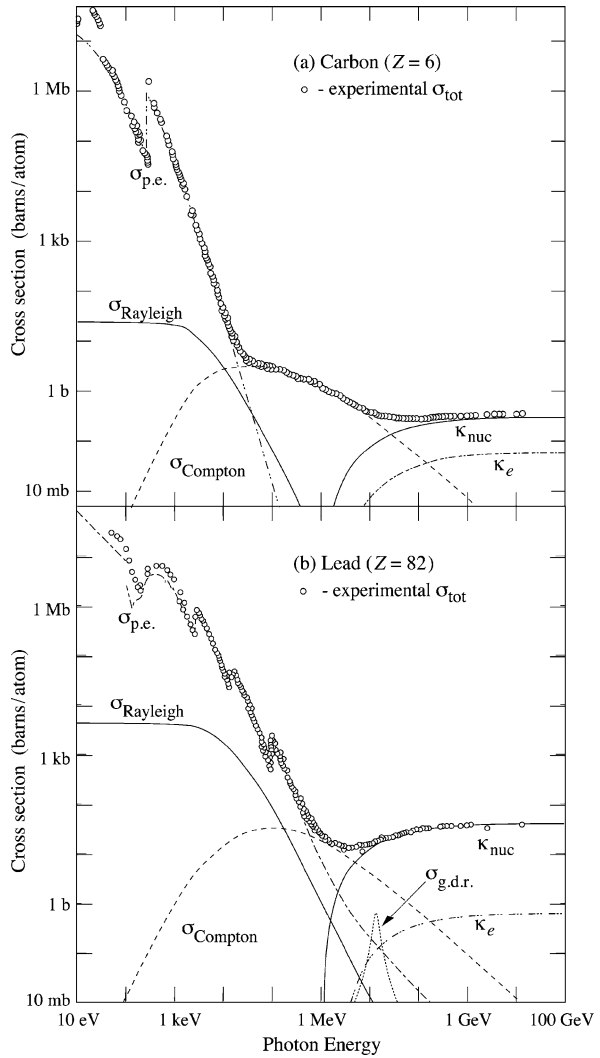
$$e^{-\frac{7}{9}\frac{dx}{X_0}}$$

Therefore, in one radiation length, a high-energy gamma ray has a 54% chance of converting into an electron–positron pair.

Overview of the interactions of gamma rays. It is convenient to consider three energy ranges when discussing gamma interactions:

– In the range 1–100 keV, the interactions are dominated by the photoelectric effect. The mean free path in water varies from about one micron at 1 keV to several

centimetres at 100 keV. The cross section strongly depends on the charge of the
nucleus.
– In the range 100 keV–1 MeV, the Compton scattering process dominates the cross
section. The mean free path rises slowly and is about 10 cm in water at 500 keV.
– At energies above 1 MeV, the pair creation process dominates the cross section.
The mean free path of gamma rays of very high energy is equal to (9/7) times the
radiation length.

If a beam of photons with intensity I_0 enters matter, the intensity $I(x)$ will satisfy
the following equation (see Sect. 2.1):

$$\frac{dI(x)}{dx} = -I(x)N\sigma$$

The cross section σ is the sum of the cross section for the photoelectric effect,
the Compton effect and the pair creation effect. It is customary to write this in terms
of the linear attenuation coefficient μ defined as $\mu = N\sigma$. Obviously, the total lin-
ear attenuation coefficient is the sum of the linear attenuation coefficients for the
photoelectric effect, the Compton effect and the pair creation effect

$$\frac{dI(x)}{dx} = -I(x)\mu$$

The intensity of the unscattered beam and the photon mean free path are therefore
given by

$$I(x) = I_0\,e^{-x\mu}, \qquad \lambda = \frac{1}{\mu}$$

One should be aware of the fact that the Compton scattering does not remove the
photons. They just lose some energy and change direction. In the literature, and in
particular in numerical tables, the following quantities are often used.

$\dfrac{\mu}{\rho}$: photon mass attenuation coefficient

$\dfrac{\rho}{\mu} = \lambda\rho$: photon mass attenuation length

Figure 2.18 shows the photon mass attenuation length of several materials as a
function of the photon energy.

Gamma rays with energies much above 1 MeV will cause electromagnetic show-
ers. On average in about one radiation length the original gamma ray gives rise to
an electron–positron pair. This electron and positron will create a large number of
secondary gamma rays by bremsstrahlung. In one radiation length, an electron or a
positron will radiate about half of its energy in this way. Many of these secondary
gamma rays will again create electron–positron pairs, and these will again undergo
bremsstrahlung and so on. If the energy of the initial gamma ray is large enough,
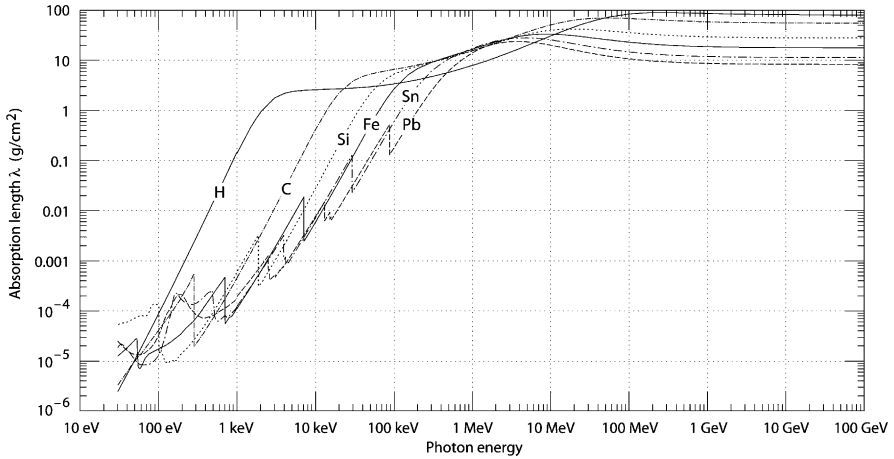
Fig. 2.18 The photon mass attenuation length $\lambda\rho = 1/(\mu/\rho)$ for various elemental absorbers as a function of photon energy. The intensity I remaining after traversal of thickness t (in mass/unit area) is given by $I = I_0 \exp(-t/\lambda)$. The accuracy is a few percent. For a chemical compound or mixture, $1/\lambda_{eff} \approx \Sigma_{elements} w_Z/\lambda_Z$, where w_Z is the proportion by weight of the element with atomic number Z. Figure reproduced from Ref. [6] in Chap. 1, with permission

the number of particles in the shower will grow exponentially. But at each step the average energy of the particles in the shower decreases, and fewer of the secondary gamma rays have sufficient energy to produce electron–positron pairs. The number of the particles in the shower will reach a maximum and start decreasing; eventually all electrons, positrons and gamma rays are absorbed or stopped.

2.5 Interactions of Particles in Matter due to the Strong Force

A proton or a neutron has an apparent size of slightly more than 10^{-13} cm, and the cross section for the collision on another proton or a neutron is therefore expected to be $\approx 4 \times 10^{-26}$ cm^2. A nucleus with atomic number A has a diameter that is $(A)^{1/3}$ times the proton diameter and a geometrical cross section that is $(A)^{2/3}$ times that of a proton. The cross section for the interaction of a proton on a nucleus of atomic number A is therefore expected to be

$$\sigma \approx 4 \times 10^{-26} (A)^{2/3} \text{ cm}^2$$

The mean free path for protons in material with atomic number A is therefore

$$\lambda = \frac{1}{N\sigma} \approx \frac{A^{1/3}}{\rho} \frac{1}{N_A \, 4 \times 10^{-26}} \approx \frac{A^{1/3}}{\rho} \, 35 \, \text{g/cm}^2$$

In this last equation, $N_A = 6.022 \times 10^{23}$ represents the Avogadro number, and we used the fact that the number of scattering centres per unit volume N is given by $N = \rho(N_A/A)$. The mean free path for protons calculated above is sometimes called the 'hadronic interaction length'. The calculation above gives a value for the numerical coefficient slightly larger than 35 g/cm^2. The value 35 g/cm^2 is by convention taken in the definition of the hadronic interaction length.

However, this very simplistic argument cannot be correct. Quantum mechanics is essential for the understanding of phenomena at this dimension scale. Indeed, in quantum mechanics a wave is associated with every particle, and the wavelength is given by $\lambda = h/P$. If this wavelength is small compared to the size of the nucleus, we should indeed expect the simplistic conclusion above to be a fair approximation. However, a proton with a kinetic energy of 10 MeV has a momentum $cP = 137$ MeV, and the quantum mechanical wavelength associated with such a proton is $\approx 10 \times 10^{-15}$ m (remember $hc = 2\pi .197 \times 10^{-15}$ MeV m). This is comparable to or larger than the size of a nucleus, and quantum mechanics is expected to be very important. At energies above ≈ 1 GeV, the wavelength associated with the proton becomes small compared to the size of a nucleus, and the simplistic result is, indeed, more or less correct. Particles with a kinetic energy above ≈ 1 GeV are usually referred to as 'high-energy particles'.

At nuclear energies, the proton–nucleus cross section will deviate substantially from the result derived above because of the effects of quantum mechanics. In addition, at low energy the electrostatic repulsion between the positive proton and the positive charge of the nucleus will prevent the proton and the nucleus from approaching each other sufficiently for a nuclear interaction to occur. This electrostatic repulsion strongly suppresses nuclear interactions at energies below a few 100 keV (see Exercise 3). The effect of this electrostatic repulsion is negligible if the energy of the proton is much larger than 100 keV.

At high energy all hadrons, on average, undergo a nuclear interaction after a distance approximately equal to the hadronic interaction length. This mean free path is in the range 10–100 cm in solids. A very high-energy proton will lose a few MeV per cm due to ionisation in a solid, and the range of the proton due to the energy loss will be larger than the hadronic interaction length. The proton will most of the time undergo a nuclear interaction before it has lost all its energy in ionisation. In such a nuclear interaction the target nucleus will be broken up. The nuclear fragments produced in this way are usually very unstable, and return to a stable condition in several steps. One particular case that needs to be mentioned is the collision of a high-energy proton with a very heavy nucleus. A very heavy nucleus has many more neutrons than protons. For example, lead has 82 protons and ≈ 125 neutrons. Nuclei with atomic charge up to about 20 have approximately equal numbers of protons and neutrons. For larger atomic charges, the neutron excess slowly increases with increasing nuclear mass. If a very heavy nucleus is broken up in a collision with a high-energy proton, the fragments will quickly expel their excess neutrons and a large number of secondary neutrons are produced. A proton of 1 GeV will, on average, produce ≈ 25 neutrons in a heavy target such as lead. This process of neutrons production is called spallation, and it is an efficient way to produce neutrons.

In addition to breaking up the nucleus, the high-energy protons will also undergo a violent collision with one or more protons or neutrons in the nucleus. In this collision a number of additional hadrons is produced. At a few GeV of energy, only a handful of secondary hadrons are produced, and this number increases slowly with energy. Typically 90% of the secondary particles produced are pions, with approximately equal numbers of π^+, π^- and π^0. Also other hadrons can be produced, but the probability of production decreases rapidly with increasing mass. If the energy of the primary proton is large enough, these pions and other hadrons will also have sufficient energy to produce further nuclear interactions, and an avalanche of hadrons is produced.

Photonuclear interactions. The term photonuclear interaction refers to the strong interaction of a gamma ray with a nucleus. It may come as a surprise that gamma rays can undergo strong interactions, since it is the very nature of gamma rays to be insensitive to the strong force. The explanation is that the strong interactions of gamma rays are an indirect effect. Indeed, a gamma ray in matter will create particle–antiparticle pairs for any elementary charged particle that exists. As for the case of the production of electron–positron pairs, this reaction is impossible because of energy and momentum conservation. However, if the produced particles only need to live a short time and exchange some energy and momentum with a nucleus, these reactions become possible. One can see strong interactions of a gamma ray as first the production of a quark–antiquark pair, and subsequently the interaction of this virtual quark pair with the nuclei in matter. The quark–antiquark pair is sensitive to the strong colour force and interacts with matter in the same way as any other hadron. The strong interactions of gamma rays are very similar to the strong interactions of any other hadron, except that the cross sections are a factor ≈ 100 lower. Below about 1 GeV, the cross section shows strong resonant behaviour. Above about 1 GeV, the cross section is fairly energy independent. Below 10 MeV, the photonuclear cross sections are extremely small because of the mismatch between the energy needed to create the virtual quark–antiquark pair and the energy available in the gamma rays.

Neutron interactions. Similar to the photon, the neutron lacks an electric charge, and therefore it is not subject to Coulomb interactions with electrons and nuclei in matter. A neutron will penetrate in matter until it undergoes a strong interaction with a nucleus. Because of the marked difference in the behaviour of the neutrons, it is customary to refer to neutrons as 'high-energy neutrons', 'fast neutrons' and 'slow neutrons', depending on their energy.

High-energy neutron means a neutron with energy larger than 1 GeV. As far as strong interactions are concerned, a high-energy neutron will behave in a very similar way as a high-energy proton, and the mean free path is of the order of the hadronic interaction length.

The neutrons produced in a nuclear reactor typically have between 100 keV and 10 MeV of energy. These are called fast neutrons. At these energies, the neutron–nucleus cross sections are often very different from the cross sections at higher energies. In addition, neutron–nucleus cross sections show a very strong energy dependence. Many neutron interactions are characterised by resonances, i.e. the

cross section for a certain type of interaction shows a pronounced peak at a particular energy. At the resonance, the cross section can be many orders of magnitude larger than at slightly higher or lower energy. This is illustrated in Fig. 3.1 showing the energy dependence of the neutron capture cross section in uranium and plutonium.

For fast neutrons, the most probable way of interacting is by elastic scattering on the nuclei of the medium. The energy loss of neutrons is mainly due to elastic scattering, but neutrons can also interact in other ways with nuclei:

(1) Inelastic scattering: the nucleus is left in an excited state, which later decays by gamma emission or some other forms of radiation. This process will only become significant for neutrons with more than 1 MeV of energy.
(2) Radiative neutron capture: the nucleus absorbs the neutron and finds itself in an excited state, which decays by gamma emission.
(3) Neutron capture followed by emission of a charged particle or followed by fission.

Fast neutrons will undergo elastic collisions and lose their kinetic energy until the energy is equal to the thermal energy of surrounding matter. The thermal energy is equal to 3/2 kT, with $kT \approx 25$ meV at room temperature. In the hot environment of a nuclear reactor, thermal energy will, of course, correspond to the temperature in the reactor core.

'Slow neutrons' usually refers to neutrons with energy less than 0.5 eV.

For slow neutrons the most probable interactions are elastic scattering and neutron capture. The elastic scattering will reduce the energy of slow neutrons further
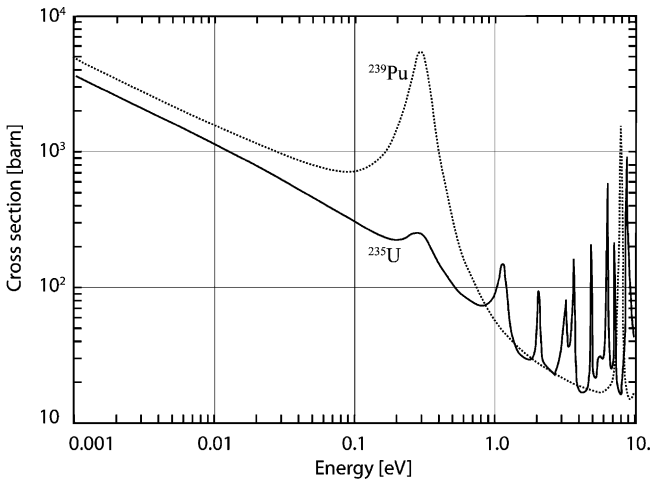
Fig. 2.19 Fission cross section of ^{235}U and ^{239}Pu as a function of energy. Both cross sections become very large at thermal energies. The cross sections show characteristic resonance peaks at certain energies. The data for this figure were obtained from Ref. [8] in Chap. 1

until they have on average the thermal energy (3/2 *kT*). For many isotopes the capture cross section is inversely proportional to the speed of the neutron, and becomes very large for thermal neutrons. This is, for example, the case for ^{235}U and ^{239}Pu as illustrated in Fig. 2.19.

2.6 Neutrino Interactions

Figure 2.20 shows the different ways electron neutrinos can interact in matter. In this figure, the interactions are shown as interactions on protons and neutrons; the underlying quark diagrams can be found in analogy with Fig. 1.5. These reactions can occur on free neutrons and protons, or on neutrons and protons bound in nuclei. It is straightforward to find the list of possible reactions by requiring conservation of the electric charge at the vertices, and by requiring that an electron neutrino can only turn into an electron, and an electron anti-neutrino can only turn into a positron. Similar diagrams can be drawn for muon neutrinos and for τ-neutrinos. The only difference is that the muon neutrino gives rise to a muon, and a τ-neutrino to a τ-lepton. In addition, neutrinos and anti-neutrinos can also scatter on the electrons present in matter. The reactions mediated by the exchange of W-bosons are called 'charged current' interactions, and the reactions mediated by the exchange of Z-bosons are called neutral current interactions.

The probability that a neutrino will interact with matter is extremely small. A neutrino has no electric charge and is not sensitive to strong interactions; it is only sensitive to weak interactions. A neutrino penetrating in matter will only interact when it comes within a distance of 10^{-18} m of one of the quarks present in the neutrons or protons inside the nuclei. Forgetting quantum mechanics, we would be tempted to say that the cross section for neutrino collisions in matter is of the order of 10^{-36} m^2, and the corresponding mean free path in matter is of the

Fig. 2.20 Diagrams for the possible interactions of electron neutrinos in matter. Similar diagrams exist for muon neutrinos and τ-neutrinos. These diagrams are not proper quantum field theory diagrams and the arrows represent the direction of motion of the particles

order of several 1000 km. This argument is completely wrong, because it is essential to use quantum mechanics and the correct description of neutrino interactions. However, also when all this is taken into account, neutrinos indeed have a very small probability of interacting.

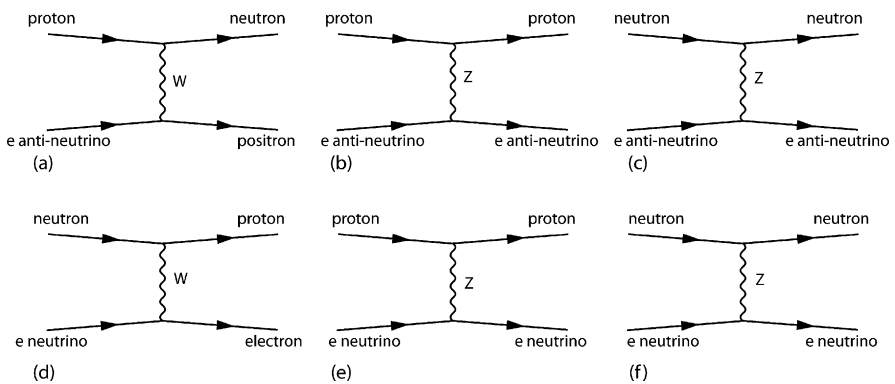At high energy, all neutrino cross sections are of similar magnitude and increase linearly with the energy of the neutrino. For neutrinos of $\approx 10^{11}$ GeV the cross section is close to the naive expectation above, but the cross section continues to rise for higher energy. At this extremely high energy, the cross section is no longer increasing linearly with the energy of the neutrino.

The daughter nuclei produced in a fission reaction are neutron rich and undergo several β^- decay processes before becoming stable. In a β^- decay a neutron in the nucleus becomes a proton through the reaction

$$n \rightarrow p + \text{anti-neutrino} + \text{electron}$$
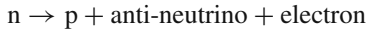
In a nuclear reactor, on average about six anti-neutrinos are produced per fission. Because of the nuclear effects, the energy spectrum of the anti-neutrinos is not the same as the energy spectrum from the decay of a free neutron. The energy spectrum extends from zero to ≈ 8 MeV, and the average energy is ≈ 1.5 MeV. The instant energy released in the fission of one ^{235}U isotope is 187 MeV, neutrino energy not included. An additional 9 MeV is carried away by the neutrinos. The total electron anti-neutrino flux near a nuclear reactor of 3 GW thermal power is therefore $\approx 1.5 \times 10^{21}$ per s. The exact number of anti-neutrinos and their energy spectrum depends on the relative importance of the different fissile isotopes in the reactor.

These anti-neutrinos can be observed through the inverse reaction shown in Fig. 2.20(a)

$$\text{anti-neutrino} + p \rightarrow \text{neutron} + \text{positron} \tag{2.12}$$
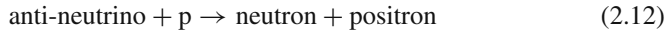
On a free proton this reaction is only possible if the energy of the anti-neutrino exceeds 1.804 MeV, because the mass of a neutron plus the mass of a positron is larger than the mass of a proton by this amount. Well above the energy threshold the cross section increases linearly with the neutrino energy E_v, and is given by

$$\sigma = 6.7 \times 10^{-42} \times E_v [\text{MeV}] \, \text{cm}^2.$$

For anti-neutrinos produced in a nuclear reactor, the cross section for reaction (2.12) on free protons is smaller because the energy is only just above the threshold. On average, for anti-neutrinos produced in a reactor, the cross section is $\sigma \approx 10^{-43}$ cm^2. The mean free path of such anti-neutrinos in normal solid matter is therefore of the order of one light year! Nevertheless, at a distance of ≈ 25 m from the core of a 3 GW reactor there will be about 500 such anti-neutrino interactions/hour in 1 m^3 of water. It is therefore quite possible to observe neutrino interactions. The difficulty is in distinguishing the anti-neutrino interactions from the much more abundant events caused by several other processes such as cosmic

rays and natural radioactivity in the detector or in the surrounding materials. The observation of reaction (2.12) at the Savannah river nuclear reactor by Reines and Cowan in 1959 [3] provided the first direct evidence for the existence of the neutrino. Such detectors could be used to monitor the fissile isotope inventory in a reactor core. Reaction (2.12) can also occur on a proton bound in a nucleus. In that case, the positron is expelled from the nucleus and the charge of the nucleus is lowered by one unit. The new nucleus thus formed is usually unstable.

The Sun produces energy by the fusion of hydrogen into helium. This requires the conversion of protons into neutrons in the reaction:

$$\text{proton} \rightarrow \text{neutron} + \text{positron} + \text{electron neutrino}.$$
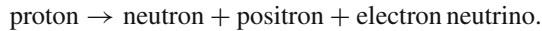
This reaction cannot occur if the neutron remains a free particle. However, if the neutron becomes part of a helium nucleus, the reaction is indeed possible. The Sun is therefore a source of electron neutrinos, not anti-neutrinos. The neutrinos from the Sun have also been observed. Because the neutrino flux from the Sun is much smaller than the neutrino flux close to a nuclear reactor, detecting these neutrinos requires very large detectors and a very careful control of the backgrounds.

High-energy particle accelerators also are intense sources of neutrinos. With accelerators not only electron neutrinos, but also muon neutrinos and τ neutrinos can be produced.

There is a subtle quantum mechanical effect causing the different flavours of neutrinos to change as a function of the distance travelled. For example an electron neutrino, after travelling some distance, can become a muon neutrino and later can again become an electron neutrino. This phenomenon is called neutrino oscillations. The effect is related to the mass of the neutrinos. A discussion of this effect is beyond the scope of the present book.

2.7 Illustrations of the Interactions of Particles

Several of the effects discussed in the present chapter can be illustrated with the help of bubble chamber pictures. A bubble chamber is an instrument that allows visualising the trajectory of charged particles. This technique has been extensively used until approximately 1980 for studying the properties of subatomic particles at high energies. The principle is as follows. A liquid is brought to a temperature above its boiling point by a sudden drop of pressure with the help of a mechanical piston. The superheated liquid immediately begins to boil and the bubbles preferentially form at points where the temperature is slightly higher. Bubbles will therefore form along the tracks of charged particles because energy is deposited along the track, and this energy is causing local heating of the liquid. The bubbles will eventually grow to quite large dimensions, but after a few milliseconds the bubbles will only be a few tenth of a millimetre in diameter. If a picture is taken a few milliseconds after the passage of the track, the trajectories of charged particles become visible as a trail of small bubbles. In a bubble chamber, there usually is a magnetic field
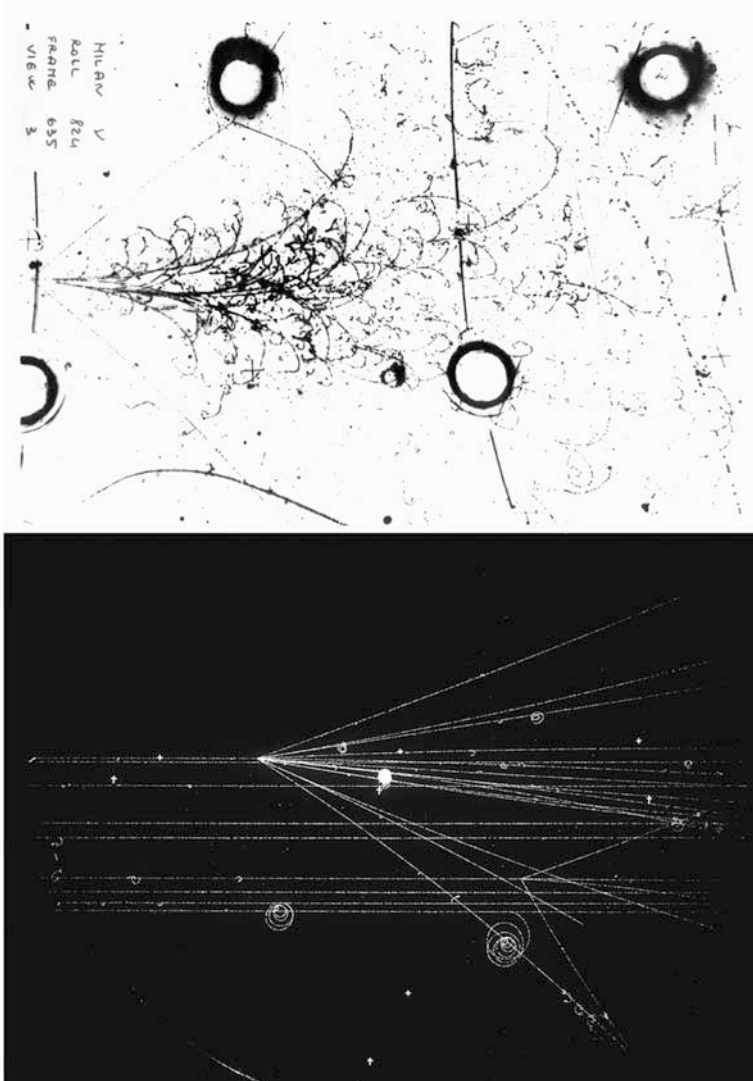
Fig. 2.21 (*Top*) A neutrino coming from the left interacts in the Gargamelle Bubble chamber filled with freon and produces two hadrons and a high-energy electron. The electron gives rise to an electromagnetic shower. (*Bottom*) Nine protons of 24 GeV enter the 30 cm hydrogen bubble chamber of CERN from the left. Two of the incoming protons interact with a hydrogen nucleus and produce secondary particles. The tightly spiralling tracks are electrons ejected from the hydrogen atoms by a high-energy particle. Images copyright CERN

perpendicular to the plane of the picture, and therefore all tracks bend clockwise to counter-clockwise depending on the charge.

The lower image in Fig. 2.21 shows a picture obtained with a bubble chamber filled with liquid hydrogen at a temperature of about 27 K. In liquid hydrogen the radiation length is ≈9 m and the hadronic interaction length is ≈6 m. The picture

shows a beam of protons with 24 GeV of energy entering the chamber from the left. Two of the protons collide with a hydrogen nucleus and produce a number of secondary particles, mostly pions. Because of the long radiation length in hydrogen, the tracks have very little multiple scattering and the trajectories are therefore nearly circles. Because of the energy loss, the curvature of the circles becomes smaller as the particle proceeds in the liquid. On the picture we notice several tracks originating on a high-energy track and completely curled up. These are electrons ejected for an atom by the high-energy track. The electron loses its energy and the radius of curvature becomes smaller and smaller until the electron stops

The upper image in Fig. 2.20 is a picture obtained with a bubble chamber filled with liquid freon. Freon is a general name for carbon-fluor compounds; in this case the freon used was CF_3Br. In this type of freon the radiation length is about 30 cm. The image corresponds to about 2 m in real space. In this picture a neutrino is coming from the left and interacting with a nucleus close to the left edge of the picture. In this interaction two hadrons are produced and one high-energy electron. The electron gives rise to a shower of electron–positron pairs and secondary gamma rays. Because of the much shorter radiation length compared to hydrogen, the tracks show much more multiple scattering in freon than in hydrogen. Therefore, the electron trajectories are erratically twisted and only vaguely resemble spirals.

2.8 Exercises

(1) Calculate the approximate mean free path of a high-energy neutron in dry air. Air is 80% ^{14}N and 20% ^{16}O by volume.
(2) Derive equation (2.10) using energy–momentum conservation.
(3) Consider two protons with the same kinetic energy and travelling on a head-on collision trajectory. The protons repel each other by the Coulomb force. At the point of closest approach the distance between protons is 2×10^{-10} m. What is the energy of each of these protons?
(4) Derive the expression for the threshold energy for the Cherenkov effect (Eq. 2.8) starting from $E = \gamma m_0 c^2$.
(5) The diameter of atoms is of the order of the Bohr radius and is given by $a = \dfrac{4\pi \varepsilon_0 \hbar}{m_e e^2}$. Use the Heisenberg uncertainty relation to argue that the velocity of the electrons in atoms is of the order $v \approx c\alpha$, where α is the fine structure constant.
$$\alpha = \frac{e^2}{4\pi \, \hbar c \, \varepsilon_0} \approx \frac{1}{137}.$$

References

1. C.F. Williamson, J.P. Boujot and J. Picard, CEA-R3042 (1966).
2. J.D. Jackson, Classical electrodynamics, John Wily & Sons (1975)
3. F. Reines and C. Cowan, Phys. Rev. 113, 273 (1959).
4. C. Leroy and P. Rancoita, Principles of radiation interaction in matter and detection, World Scientific (2004).

Chapter 3
Natural and Man-Made Sources of Radiation

3.1 Natural Sources of Radiation

In 1895, Henri Becquerel discovered that some uranium salts emit penetrating radiation that can be made visible with ordinary photographic emulsions. This discovery was the beginning of nuclear science. Today we know that most natural radioactivity is due to a few very long-lived unstable isotopes that were formed 4.5×10^9 years ago. At that time the Earth was formed, and presumably, a very large number of different radioactive isotopes were produced. Those with a shorter half-life have all decayed and only those with a very long half-life still exist today. The isotopes contributing most to the natural radioactivity are uranium-235, uranium-238 and thorium-237. The half-life of each of these isotopes is listed in Table 3.1. After a long decay chain with a succession of alpha and beta emissions, these isotopes eventually decay to one of the stable isotopes of lead. There are many other isotopes with a very long lifetime occurring naturally; the most important of these is potassium-40. Several short-lived isotopes also occur naturally; these are either decay products of long-lived isotopes or are produced by cosmic rays. It is worth mentioning radon (^{222}Rn), which is an alpha emitter with a half-life of 3.8 days. This isotope is produced in the decay of uranium. Although in general it is very rare, it can occasionally be found in high concentrations inside buildings or in thermal springs and can represent a health hazard. Some degree of radioactivity is present in all materials. The amount of activity present in natural materials varies by orders of magnitude. A value of up to 74,000 Bq/kg is usually considered "not radioactive".

Table 3.1 Some naturally occurring radioactive isotopes

Isotope	Half-life [years]
Thorium-232	1.405×10^{10}
Uranium-235	7.04×10^8
Uranium-238	4.468×10^9
Potassium-40	1.25×10^9
Carbon-14	5715

The best known example of the cosmogenesis is carbon-14. This isotope is constantly produced in the atmosphere by cosmic rays and is therefore present in approximately constant concentration in air. Any living organism takes its carbon from the air, either directly or indirectly. Any living biological material therefore contains the same fraction of ^{14}C. After the organism has died, it no longer absorbs carbon from the air and the fraction of ^{14}C starts to decrease. Measuring the ^{14}C concentration is, therefore, a very powerful method for dating archaeological samples.

Besides these naturally occurring radioactive isotopes, there are many artificial radioactive isotopes. These are produced in nuclear reactors or with particle accelerators. Often the production of such isotopes is an undesirable side effect, but some of these radioactive isotopes have useful applications and are made on purpose. Annexe 6 lists some commonly used radioactive isotopes.

Another important natural source of radiation are the 'cosmic rays'. In 1912, Victor Hess carried electrometers (see Sect. 4.1) to an altitude of 5300 m in a balloon flight. He found that the ionisation rate increased approximately four-fold over the rate at ground level. He concluded that this was caused by radiation from outer space. Today we know that this radiation primarily consists of positively charged nuclei. Of the primary charged particles in cosmic rays, ≈90% are protons, ≈9% are helium nuclei and ≈1% are electrons. This radiation spans an enormous energy range, from ≈1 GeV [10^9 eV] up to ≈10^{20} eV. The flux of primary cosmic ray particles decreases approximately like $E^{-2.7}$. The flux for particles with energy exceeding 10^{17} eV is ≈1/(km^2 h).

The Sun is an intense source of energetic particles, but the energy of these particles rarely exceeds ≈1 GeV. These particles are accelerated in plasma waves in the corona of the Sun and the intensity fluctuates considerably depending on the solar activity. Occasionally, there are bursts of activity, and the intensity of the radiation increases by several orders of magnitude during the bursts. The Earth's atmosphere will completely stop any radiation with energy less than 1 GeV and this radiation is therefore harmless for people living on the surface of the Earth, as most of us do. This stream of particles is deflected by the Earth's magnetic field towards the poles, where it causes eerie phenomena such as the Northern Lights.

The term 'cosmic rays' usually only refers to particles with a primary energy above 1 GeV. The Earth's atmosphere corresponds to about 10 times the hadronic interaction length and 30 times the radiation length. A high-energy particle coming from outer space will always interact somewhere at high altitude in the atmosphere. In the collision, a large number of secondary particles will be produced, mainly protons, neutrons and pions. This is illustrated in Fig. 3.1. The secondary protons and neutrons will again interact, producing new secondary particles of lower energy and so on. Eventually, the energy is so low that particles are stopped by ionisation of the air molecules. The result is that protons and neutrons very rarely reach the Earth's surface. However, charged pions have a lifetime of 2.6×10^{-8} s. The average distance travelled by a high-energy pion before it decays is given by
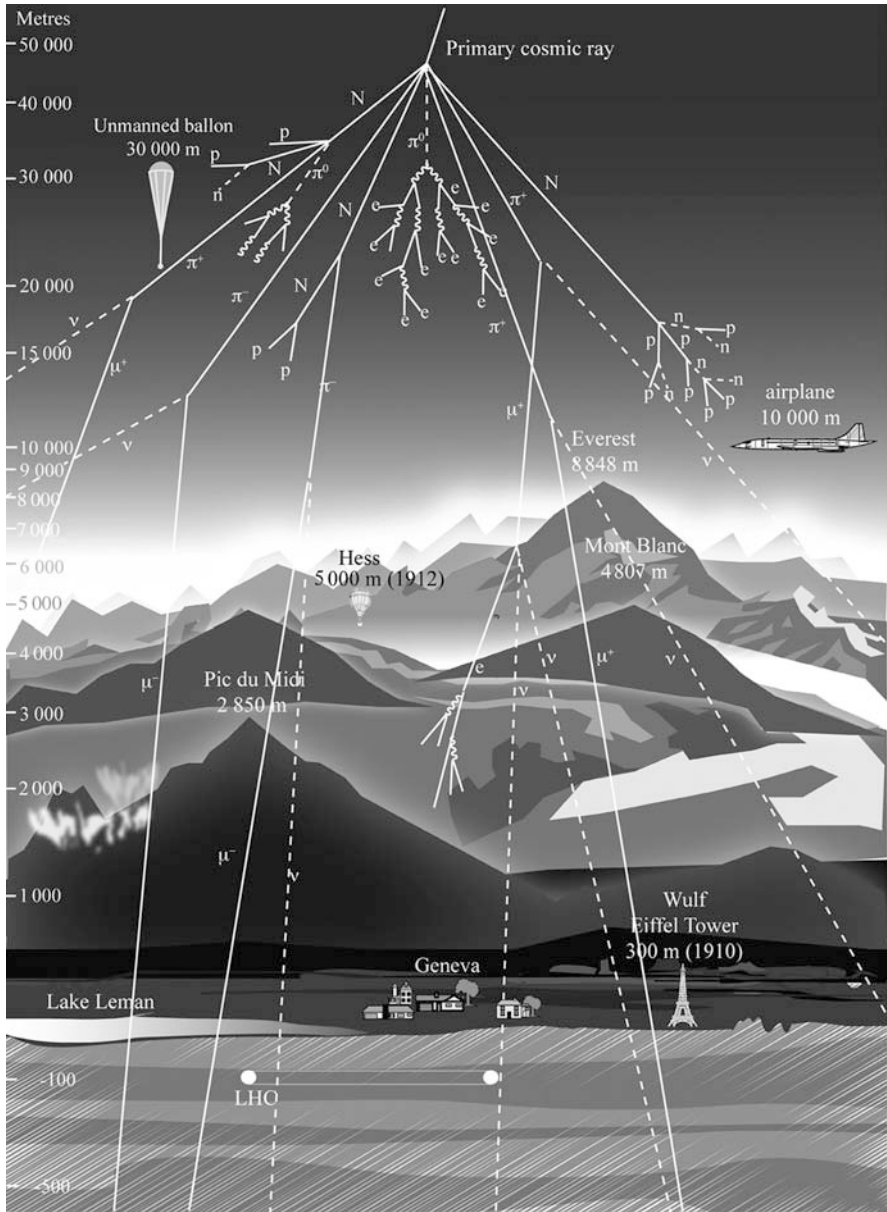
Fig. 3.1 Artist's view of the interaction of a very high high-energy cosmic ray in the upper atmosphere and the subsequent production of secondary particles. Most of the time only muons, neutrinos and some low-energy gamma rays will reach the surface of the Earth. For the sake of clarity, the distance travelled by neutral pions is shown much larger than reality

$$\frac{E_\pi}{m_\pi} c \, 2.6 \, 10^{-8} \, \text{s} = \frac{E_\pi}{m_\pi} \, 7.8 \, \text{m} \, .$$

The symbols E_π, m_π and c stand for the energy and mass of the pion and the velocity of light. The pions are sensitive to the strong colour force, but high in the atmosphere the mean free path before a nuclear interaction for a pion is several kilometers; therefore, most of the pions will have decayed before they can interact. A pion decays into a muon and a neutrino as indicated below:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu$$
$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$$
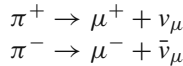
The neutrino is almost unobservable. A muon has an electric charge and is therefore easily observable. However, a muon is insensitive to the strong colour force; a muon will almost never undergo a nuclear interaction. The muon will lose energy according to the Bethe–Bloch equation discussed in Sect. 2.2. In travelling from the upper atmosphere down to the surface of the Earth, it will lose about 2 GeV. Many of the muons have more than 2 GeV of energy and can, therefore, reach the surface of the Earth. The muon has a lifetime of 2.2×10^{-6} s and decays into an electron and two neutrinos. However, the average energy of muons at sea level is about 4 GeV and the mean free path before decay of a muon with this energy is about 25 km. Many of the muons produced high in the atmosphere will therefore reach the surface of the Earth.

In the interaction of primary cosmic rays, neutral pions are also produced. Such neutral pions will decay in 8.4×10^{-17} s into two gamma rays. This lifetime is so short that a neutral pion will only travel a microscopic distance before decaying. At high energy, a gamma ray will initiate an avalanche consisting of a large number of electrons, positrons and secondary gamma rays. On average in about one radiation length the original gamma ray gives rise to an electron–positron pair. This electron and positron will create a large number of secondary gamma rays by bremsstrahlung. In one radiation length, an electron or a positron will radiate about half of its energy in this way. Many of these secondary gamma rays will again create electron–positron pairs and these will again undergo bremsstrahlung and so on. If the energy of the initial gamma ray is large enough, the number of particles in the shower will grow exponentially. However, at each step the average energy of the particles in the shower decreases, and fewer of the secondary gamma rays have sufficient energy to produce electron–positron pairs. After a few radiation lengths, the number of the particles in the shower reaches a maximum and thereafter starts to decrease. Eventually, all electrons, positrons and gamma rays are absorbed or stopped.

Because of these cosmic rays, everywhere on the Earth's surface there is a constant flux of muons. The intensity of this flux is of the order of $1/(\text{cm}^2 \cdot \text{min})$. The energy spectrum of cosmic ray muons is shown in Fig. 3.2. These muons typically have a few GeV of energy, but the spectrum extends beyond 100,000 GeV. In addition to the muons, we will also see the end of the electromagnetic shower caused by gamma rays from neutral pion decay or by primary electrons. This will give electrons and gamma rays with energy rarely exceeding a few 10 MeV.
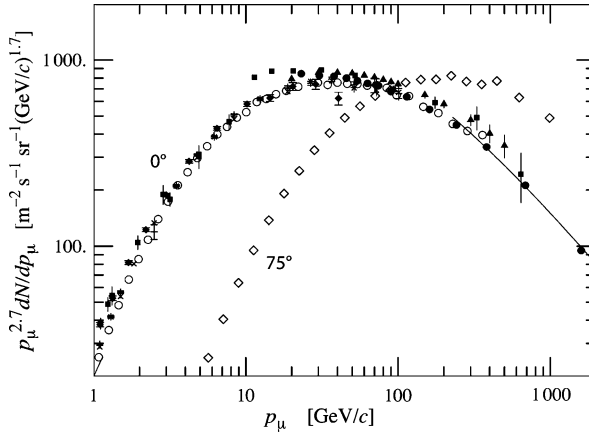
Fig. 3.2 Momentum spectrum at sea level of muons produced in cosmic rays. For energies above 1 GeV, the energy and the momentum of a muon are almost the same when expressed in GeV and GeV/c, respectively. Notice that this plot shows the muon flux multiplied by the muon momentum to the power of 2.7; the curve therefore peaks at a much larger value than the actual muon momentum spectrum. The angle indicated in the figure is the angle of the muon relative to the vertical direction. Figure by courtesy of the particle data group [6] in Chap. 1

The origin of cosmic rays is not known with any certainty. It is widely believed that most cosmic rays have been accelerated in the plasma shock waves caused by supernova explosions in our galaxy. There are indications that the cosmic rays with energy above 10^{14} eV are of extragalactic origin, possibly accelerated in the extremely intense electromagnetic fields that are known to exist near massive black holes.

The radiation dose from cosmic rays is small for people living on the surface of the Earth. At an altitude of 10 km, cosmic radiation is much more intense and reaches an average of 5 μSv/h. This is negligible for an occasional traveller, but is of some concern for airline crews. Astronauts in low orbits are at moderate risk because the magnetic field of the Earth shields out most cosmic rays. Outside low Earth orbit, this radiation is much more intense, it is an important concern for astronauts and it represents a major obstacle to future long-term human exploration of the Moon or Mars.

3.2 Units of Radiation and Radiation Protection

Much of this section is reproduced from the 'Review of Particle Physics', Ref. [6] in Chap. 1. The International Commission on Radiation Units and Measurements (ICRU) recommends the use of SI units. We also mention CGS units, and some other non-SI units, because these units are still widely used.

- *Unit of activity:* The amount of radioactivity present in a sample can be characterised by the number of radioactive decays per second. The corresponding unit is the becquerel (Bq); the corresponding non-SI unit is the curie (Ci):

$$1 \text{ Bq} = 1 \text{ disintegration s}^{-1} = 1/(3.7 \times 10^{10})\text{Ci}$$

- *Unit of absorbed dose:* The amount of radiation absorbed in a sample can be characterised by the amount of energy deposited by the radiation in the sample. The corresponding unit is the gray (Gy); the corresponding non-SI unit is the 'rad':

$$1 \text{ Gy} = 1 \text{ joule kg}^{-1} = 6.24 \times 10^{12} \text{ MeV kg}^{-1} = 100 \text{ rad}$$

- *Unit of exposure:* This unit is somewhat obsolete, but it continues to appear on many measuring instruments. It is a measure of the photon fluence at a certain point in space integrated over time, in terms of ion charge pairs produced by secondary electrons in a small volume of air around the point. The name of the unit simply is 'unit of exposure'; the corresponding non-SI unit is roentgen (R). One 'unit of exposure' creates 1 Coulomb of ionisation charges in one kilogram of air.

$$1 \text{ R} = 1 \text{ esu cm}^{-3}\text{in air} = 2.58 \times 10^{-4}\text{'unit of exposure'}$$

Implicit in the definition is the assumption that the small test volume is embedded in a sufficiently large and uniformly irradiated volume, and that the number of secondary electrons entering the volume equals the number of secondary electrons leaving the volume, i.e. there is charged particle equilibrium.

- *Unit of equivalent dose:* The amount of biological damage caused by ionising radiation in a sample depends on the type of radiation and on the amount of energy deposited by the radiation in the sample. The corresponding unit is the sievert (Sv), the corresponding non-SI unit is 'rem' (roentgen equivalent for man). The conversion factor is 1 Sv = 100 rem. The equivalent dose H_T in an organ T is equal to the absorbed dose in the organ in gray, times the radiation weighting factor ω_R, formerly called the quality factor Q. The equivalent dose expresses the long-term risks, primarily cancer and leukaemia, from low-level chronic exposure. It depends on the type and energy of the radiation as indicated in Table 3.2 [1]:

Table 3.2 Radiation weighting factors

Radiation	ω_R
X- and γ-rays, all energies	1
Electrons and muons, all energies	1
Neutrons, energy < 10 keV	5
Neutrons, energy 10 keV–100 keV	10
Neutrons, energy > 100 keV–2 MeV	20
Neutrons, energy 2 keV–20 MeV	10
Neutrons, energy > 20 MeV	5
Protons (other than recoils) > 2 MeV	5
Alphas, fission fragments, and heavy nuclei	20

Table 3.3 Tissue weighting factors ω_T

Tissue or organ	ω_T
Gonads	0.20
Colon	0.12
Bone marrow	0.12
Lung	0.12
Stomach	0.12
Chest	0.05
Liver	0.05
Thyroid gland	0.05
Oesophagus	0.05
Skin	0.01
Bone surface	0.01
Adrenals, brain, small intestine, kidney	0.05
Muscle, pancreas, spleen, thymus, uterus	

- *Effective dose:* The amount of biological damage an irradiated person suffers is called the 'effective dose' *E*. This is the sum of the equivalent doses in each tissue H_T, weighted by the tissue weighting factors ω_T of the organs and tissues in the body that are considered to be the most sensitive [1]:

$$E = \sum_T \omega_T \times H_T$$

The tissue weighting factors are listed in Table 3.3.

- *Radiation levels* [2]: The natural annual background radiation dose, summed over all sources, in most world areas, amounts to a whole-body equivalent dose rate in the range 0.4–4 mSv/year. The world average is 2.5 mSv/year. It can reach up to 50 mSv/year in certain areas. The most important component, ≈ 2 mSv, comes from the inhaled natural radioactivity, mostly radon and radon daughters. The average quoted is for a typical house but varies considerably. It can be more than two orders of magnitude higher in poorly ventilated mines. It is only 0.1–0.2 mSv/year in open areas. The US average is ≈ 3.6 mSv/year. In Europe it varies from 2 mSv/year in the UK to 7.5 mSv/year in Finland.

  Table 3.4 gives some typical average values for the different contributions to the annual background doses received by an average person. The average contribution from medical interventions has increased in recent years and is probably underestimated in this table.

- *Cosmic ray background in counters:* At sea level a detector for charged tracks will count: < 1 min^{-1} cm^{-2} counts due to cosmic rays penetrating in the detector. Most of the counts are due to muons, the rest is due to low-energy electron or positron tracks caused by gamma interactions.

- *Dose from external gamma emitting sources:* The dose rate in air from a gamma point source of '*C*' Curies emitting one photon of energy *E* MeV, with energy $0.07 < E < 4$ MeV, per disintegration, at a distance of 30 cm is about $6 \times C \times E$ rem/h, or $60 \times C \times E$ mSv/h. The uncertainty on this number is $\approx 20\%$. The dose rate from

Table 3.4 Typical average values for the contribution of different sources to the radiation dose for an average person

Natural external gamma radiation	19%
Natural internal radiation	17%
222Rn (radon, alpha emitter 3.8 d)	32%
220Rn (radon, alpha emitter 55 s, also called thoron)	5%
Medical	10.1%
Fallout nuclear explosions	0.5%
Professional	0.4%
Nuclear waste	0.1%
Other	0.5%

a point source decreases approximately proportional to $1/r^2$ as a function of the distance r. The dose rate in air from a semi-infinite uniform photon source of specific activity C (in µCi/g) and gamma energy E (in MeV) is about $1.07 \times C \times E$ rem/h or $10.7 \times C \times E$ mSv/h.

- *Recommended limits to exposure of radiation workers (whole-body dose):*
    EU&Switzerland: 20 mSv year^{-1}
    US: 50 mSv year^{-1}
- *Lethal dose:* The whole-body dose from penetrating ionising radiation resulting in 50% mortality in 30 days (assuming no medical treatment) is 2.5–4.5 Gy. For this number it is assumed that the dose is measured internally on body longitudinal centre-line. The surface dose varies due to variable body attenuation and may be a strong function of energy.
- *Cancer induction by low-LET radiation:* The probability to induce cancer, on average, is about 5% per Sv. [1]

3.3 Electrostatic Accelerators

Most of the radiation we use is not natural but made by artificial means. Nuclear reactors produce a huge amount of radioactive material, but this subject is not discussed in these lecture notes. We will only discuss the production of high-energy particles with accelerators. The book 'An Introduction to Particle Accelerators' by Edmund J. N. Wilson [3] contains an excellent introduction to the physics of particle accelerators. A more advanced discussion of accelerator technology can be found in [4–9].

The most straightforward way to accelerate charged particles is by using an electrostatic potential difference. A particle with a charge $Z = 1$ travelling a potential difference of X volts receives a kinetic energy of X eV. All that is needed is a high-voltage power supply and a source of charged particles.

The only charged particles that are easily available are electrons and nuclei including protons. Any material heated to a high temperature under vacuum will copiously emit electrons. The hot filament is covered with a suitable substance with

low work function (usually alkali oxides) to increase the electron yield. There are many different designs of ion sources, but essentially they are all based on causing a glow discharge in a low-pressure gas. The pressure is typically 0.01 mbar. The glow discharge is a plasma, i.e. a state of matter where most atoms are ionised. There is usually a hot filament producing electrons and a magnetic field. The electrons spiral in the magnetic field and ionise the gas, helping the plasma formation. The beam of positive ions radiates from the plasma chamber through one or more small holes.

If a particle only needs to be accelerated to a few 100 keV, standard commercial high-voltage power supplies can be used. To obtain a much higher energy, special devices are necessary. There are basically two methods to generate the very high potential difference that is needed to accelerate a particle to high energy. This gives rise to two different types of accelerator; the Cockcroft–Walton and the Van de Graaff. These are discussed below. The first successful particle accelerator was built by Cockcroft and Walton and was used in 1932 to create the first example of transmutation of elements using an accelerator.

The Cockcroft–Walton accelerator. The layout of a Cockcroft–Walton accelerator is illustrated in Fig. 3.3. The high voltage is generated with the circuit shown on the right-hand side of the figure. This type of circuit is commonly used in many applications that need a high DC voltage source. It uses a voltage multiplier ladder network of capacitors and diodes to generate a high voltage. The principle is as follows: A moderately high-voltage transformer creates an alternate voltage at one end of the secondary winding, the other end being connected to ground. The capacitor transmits this voltage to point A by capacitive coupling. At first the voltage at point A will also oscillate between +V and −V, but every time this point is at a voltage below zero, a current will flow through the diode, charging the point A to a positive voltage oscillating between 0 V and +2 V. The diode between point A and point B will then cause the charging of point B to the potential +2 V. The same scheme can then be repeated many times to reach higher and higher voltages, but eventually the
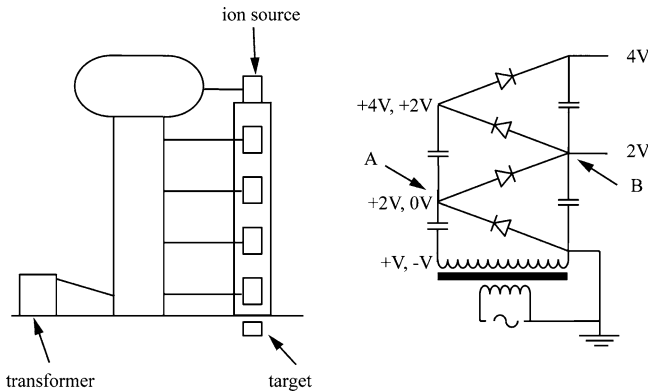
Fig. 3.3 Working principle of a Cockcroft–Walton accelerator

problems associated with very large electrostatic potentials will also limit this type
of accelerator to ≈10 MV.

Until recently a Cockcroft–Walton accelerating structure was often used as the
first acceleration step in the high-energy accelerators that will be discussed in the
next sections. Today this method is abandoned in favour of RF quadrupole acceler-
ation structures. Only the first ≈100,000 V of acceleration is usually still obtained
with an electrostatic voltage difference.

The Van de Graaff accelerator. A completely different approach to reach a high
voltage is used in the Van de Graaff accelerator. This device looks like a 19th century
electrostatic instrument, but it is still used today for accelerating ions. The working
principle of a Van de Graaff accelerator is schematically shown in Fig. 3.4. The
high-voltage electrode is a hollow sphere. A circular rubber band runs continuously
between the high-voltage electrode and the low-voltage side of the accelerator. A
high-voltage power supply of a few 10 kV at the low-voltage side of the accelerator
provides the positive charge. The electric charges are generated by field emission at
the tip of fine needles and sprayed on the rubber belt. The belt transports the charges
mechanically to the high-voltage electrode. Inside this high-voltage electrode,
the charges are collected by reverse field emission. The charge collection is done
inside the hollow electrode, where the potential is constant, regardless of how large
the potential of this electrode is. The ion source is inside the high-voltage elec-
trode. The ions are accelerated in high vacuum inside a straight tube connecting the

Fig. 3.4 (**a**) Schematic representation of a Van de Graaff accelerator. (**b**) Practical realisation of
a Van de Graaff accelerator. The accelerator is usually placed horizontally and is contained in a
pressure vessel to reduce the dimensions of the system

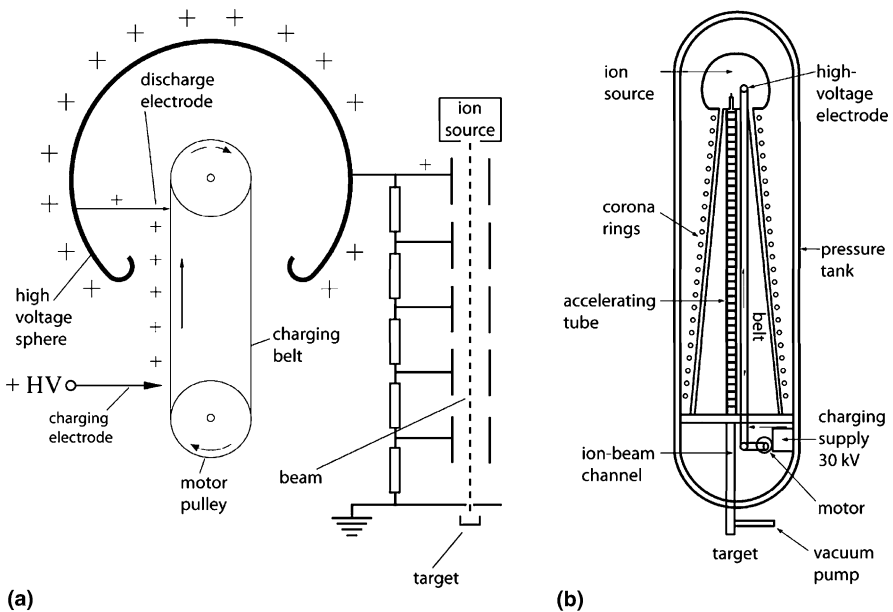high-voltage electrode to the target area outside the accelerator. In the tube there are field-shaping electrodes ensuring that the electric field lines guide the ions towards the target. In modern Van de Graaff accelerators, to avoid sparking between the high-voltage electrode and any other metal object at ground potential, the complete accelerator is in a pressure vessel filled with a suitable gas such as freon or SF_6. However, also with such precautions, high static potentials in excess of 1 million volts are an enormous technical challenge. A voltage of 10 million volts will cause sparking over a distance of the order of 10 m in air at atmospheric pressure; the exact value depends on the shape of the electrodes. These technical problems with very high voltages limit the maximum energy that can be reached with electrostatic accelerators to about 25 MeV.

An interesting variant on this instrument is the 'Tandem Van de Graaff' accelerator. This instrument takes advantage of the tendency of protons to form negative ions by capturing two electrons. The high-voltage electrode is brought to a large positive potential. A negative hydrogen ion source is outside the detector at zero potential, and the negative ion is accelerated towards the centre of the Van de Graaff. There the negative ion is stripped of its two electrons by letting it pass through a thin metal foil, and the resulting positive ion is accelerated a second time by the same potential difference. This machine allows doubling of the energy of the protons, and it also has the advantage that the delicate proton source is easily accessible outside the accelerator structure.

3.4 Cyclotrons

The difficulties with very high voltages led Rolf Wideröe in 1928 to propose accelerating particles by using a lower voltage difference several times. The principle is illustrated in Fig. 3.5. A beam of particles passes through a succession of metallic tubes. The voltage difference between the tubes is changed while the particles are inside the tube, in such a way that the particles always see an accelerating field on passing from one tube to the next. Several accelerators using this principle were actually built, but the highest radio frequency (RF) that could be made in the 1930s was ≈10 Mz. With this frequency, the linear accelerator becomes impractically long. Today, frequencies in the GHz range are possible and thus make this a practical proposal. Such linear accelerators will be discussed later.
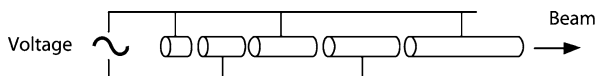

Fig. 3.5 The linear accelerator proposed by Rolf Wideröe consists of a series of tubes alternatively connected to the two poles of an alternating high-voltage power supply. The particles travel inside the tube. The length of the tubes is adjusted such that the particles always see an accelerating field when passing from one tube to the next

Fig. 3.6 Schematic representation of a cyclotron. The distance between the pole pieces of the magnet is shown larger than reality to allow seeing what is inside
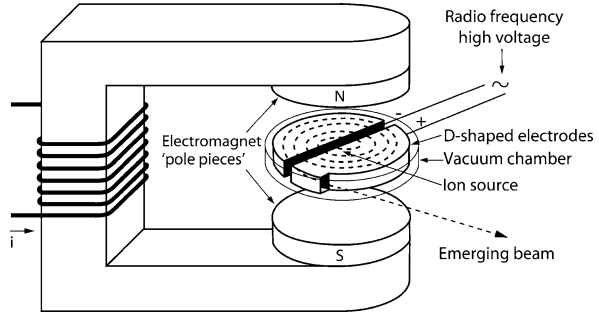
To obtain a more compact accelerator, in 1930 Lawrence proposed bending the particles into a circular path with a magnetic field. In this way the same electrodes can accelerate the particles several times. The idea is illustrated in Fig. 3.6. The essential components of a cyclotron are a homogeneous and parallel magnetic field that forces the particles to travel in circles and an accelerating cavity in the shape of a pillbox cut into two halves. The two electrodes are called 'Dee's' because of their shape.

A large and alternating electric potential difference is applied on the two D-shaped electrodes. The particles emanating from the ion source in the centre are accelerated by the field in the gap between the two D-shaped electrodes. The magnetic field bends the charged particles back towards the gap. If the frequency is right, the field will have reversed when the particles pass the gap a second time and the particles receive a second acceleration. This is repeated again and again. The energy of the particles increases and so does the radius of the particle trajectory. Eventually, the radius becomes too large and the particles leave the cyclotron with a high energy.

When a charged particle is travelling in a plane perpendicular to a magnetic field, the particle will travel in a circular orbit. The radius of the orbit is found by requiring the centrifugal force and the Lorentz force to be equal. Using MKSA units we have

$$M\frac{v^2}{r} = Ze\,\vec{v} \times \vec{B} = Ze\,v\,B$$

The notations used are

M: mass of the particle
r: radius of the trajectory of the particle
v: velocity of the particle
e: charge of the proton
Ze: charge of the particle
B: magnetic induction

The equation above is only correct in the non-relativistic limit. To make it relativistic we need to make the substitution

$$M \Rightarrow M\gamma = \frac{M}{\sqrt{1 - (v/c)^2}} = M\frac{E_{\text{kinetic}} + Mc^2}{Mc^2}$$

We therefore obtain a very simple relation between the momentum P of the particle, the magnetic induction B and the radius of curvature of the trajectory r.

$$P = Mv\gamma = Ze\,B\,r$$

In convenient units this relation is written as

$$\{Pc\}\ [\text{GeV}] = 0.29979 \cdot Z \cdot B[\text{tesla}] \cdot r[\text{m}] \tag{3.1}$$

The rotation frequency of the charged particles is therefore given by

$$f = \frac{v}{2\pi r} = \frac{ZeB}{2\pi M\gamma}$$

In convenient units this relation is written as

$$f[\text{MHz}] = 14.3\frac{Z\,B[\text{tesla}]}{\gamma\ \{Mc^2\}[\text{GeV}]} \tag{3.2}$$

It follows from Eq. (3.2) that the frequency is independent of the energy of the particle as long as the particle remains non-relativistic and therefore $\gamma \approx 1$. With a field of about 1 tesla, the frequency needed is of the order of 10 MHz for protons. Note that in the 1930s a large RF field with this frequency was feasible.

In the simplest version of the cyclotron, there is a dipole magnet with a soft iron core. The pole faces are flat and parallel and create a constant and parallel magnetic field. In between the pole faces there are two accelerating electrodes, and the RF field has a constant frequency given by Eq. (3.2).

The cyclotron just described will certainly accelerate particles, but we would expect that only very few particles will be accelerated. Only particles that start off travelling exactly in the mid-plane between the two magnet poles will eventually reach the exit port. If the particle has a small momentum component parallel to the magnetic field, it will soon hit one of the D-shaped electrodes. The more turns the particles make, the more severe this limitation becomes. What is required is a mechanism to force the particles back towards the mid-plane of the accelerator.

We need to study more closely the geometry of the magnetic field to understand how this is achieved. The field in the dipole will not be exactly constant and parallel. In the centre, the field will have the maximum value and it will decrease slowly from the centre towards the edge of the poles. As a result, the magnetic field lines cannot be exactly parallel, but must have a shape as shown in Fig. 3.7. To prove this let us consider the Maxwell equations in the integral form. The loop integral over the loop shown in Fig. 3.7 must be zero, since there is no current inside this loop. We therefore can write

Fig. 3.7 The shape of the
magnetic field lines in the gap
between the pole pieces
results in a focusing effect on
the particles being accelerated
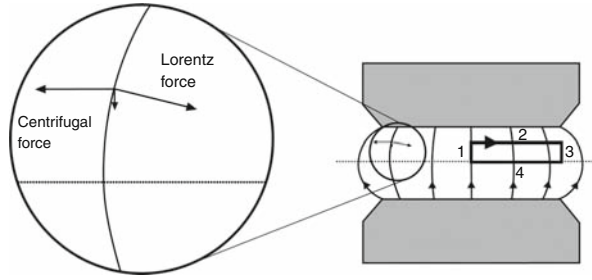
$$\oint \vec{H} \, d\vec{l} = \int_1 \vec{H} \, d\vec{l} + \int_2 \vec{H} \, d\vec{l} + \int_3 \vec{H} \, d\vec{l} + \int_4 \vec{H} \, d\vec{l} = 0$$

The four parts of the integral represent the line integrals over the four sides of
the loop. Because of the overall symmetry, the line integral over part four is equal
to zero and we have

$$\int |H_1| \, l_1 + \int \vec{H}_2 \vec{l}_2 - \int |H_3| \, l_3 = 0$$

If the field in the centre is larger than at the edge we have $|H_1| > |H_3|$. If follows
that

$$\int \vec{H}_2 \vec{l}_2 < 0$$

This shows that the field lines must be bending outward as shown in Fig. 3.7.

This shape of the field will have a focusing effect. This is made clear on the left-
hand side of Fig. 3.7. The Lorentz force is always perpendicular to the magnetic field
lines, while the centrifugal force always points radially outwards. In the mid-plane
of the magnet, both forces exactly compensate, but away from the mid-plane, a small
component towards mid-plane remains. To a good approximation, this restoring
force is proportional to the distance of the particle from the mid-plane. The particles
will have a harmonic oscillation in the vertical direction around the mid-plane.

Also the radial trajectory of the particles in the plane should be stable around the
nominal trajectory. It can be shown that the condition for radial stability is

$$\frac{dB_z}{dr} \leq 0$$

For the proof of this statement, I refer the reader to [10]. In the simple geometry
we are discussing, this condition is also satisfied. It is remarkable that a simple
dipole with flat pole pieces has exactly the magnetic field that is needed to have
stable particle acceleration conditions.

However, there remains one problem. If the field is radially decreasing, the rota-
tion frequency is no longer constant. At a small radius, we should have a slightly

Fig. 3.8 Phase of the
particles relative to the RF
phase in a cyclotron with
planar pole faces
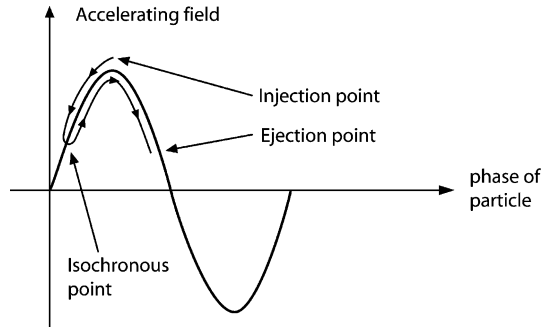
larger RF frequency than at a large radius. If the particles are only making a small
number of turns, we can get away with taking an average value for the RF frequency.
Assume a particle starts its journey exactly in phase with the RF field as shown in
Fig. 3.8. In the beginning, the particle will travel too fast and it will gradually be
more and more early relative to the maximum of RF phase. However, at the same
time, its trajectory will have a larger radius and the mismatch between the rota-
tion frequency and the RF frequency will decrease. If the number of turns is not
very large, the particle will reach the point where the rotation frequency and RF
frequency are equal before it is completely out of phase. From this point on, the
particle will be too slow and will start lagging behind relative to the maximum of
the RF phase. If the number of turns is not very large, the particle will reach the exit
before it is too much out of phase. If the cyclotron should only accelerate particles
to a modest energy, this method is possible, and early cyclotrons worked in exactly
this way.

As the energy increases, more turns are necessary and the method described
above can no longer be used. Moreover, as the energy increases, the relativistic
correction to the frequency in Eq. (3.2) can no longer be neglected. It is no longer
possible to have a constant RF frequency.

The most straightforward solution is to follow one particular bunch of particles
from the source to the ejection and adjust the frequency throughout as the particles
are accelerated. A cyclotron using this principle is called a synchro-cyclotron.

Many years ago high-energy cyclotrons worked in the way just described. The
drawback is that only a small fraction of the ions produced at the source is
accelerated. The higher the energy, the more severe this effect becomes.

For this reason, most cyclotrons today have a very different focusing mecha-
nism using a much more complicated magnetic field shape, namely, focusing with
azimuthally varying fields.

In this design, the magnet is subdivided into azimuthal sectors with alternatively
larger and smaller values for the magnetic field, as shown in Fig. 3.9. In this figure,
darker sectors represent large values of the field and lighter sectors smaller values of
the field. In this field geometry, the particles no longer travel in circles, but according
to a trajectory as shown in the figure. The particle, therefore, acquires a periodically
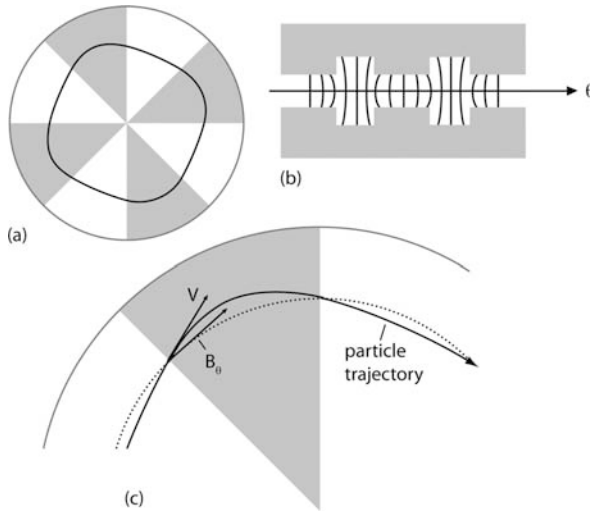
Fig. 3.9 (**a**) In a cyclotron using focusing with azimuthally varying fields, there are alternating sectors with smaller and larger value of the field. The trajectory of a particle is no longer a circle but has a shape as shown by the solid black line in the figure. (**b**) The magnetic poles are not flat but have hills and valleys. The field develops an azimuthal component. (**c**) A positive particle travels to the right, the magnetic field points upwards. At the edge of the high-field region, the field acquires an azimuthal component. Above the mid-plane of the magnet the direction of this azimuthal component is as shown in the figure. The radial component of the velocity and the azimuthal component of the field provide a force pushing the particle back to the mid-plane of the magnet

varying radial component of the velocity. The magnetic field lines acquire a variable azimuthal component, as shown in Fig. 3.9(b). At the edge of the high-field section, the combination of the radial component of the velocity and the azimuthal component of the field together cause a force that pushes the particles back to the mid-plane of the magnet. This force is focusing both when the particles enter and leave a high-field sector.

In addition, the sectors are also given a spiralling shape as shown in Fig. 3.10. It is possible to show that this will further enhance the focusing effect on the beam. The focusing effect obtained with this azimuthal variation of the field is strong enough to compensate the defocusing effect due to a magnetic field that slightly increases with the radius. In this way, it is possible to make an isochronous cyclotron, i.e. a cyclotron where the rotation frequency of the particles remains constant during the acceleration cycle. With an isochronous cyclotron, a much larger beam current can be achieved. Figures 3.11 and 3.12 show examples of cyclotrons.

For the extraction of the beam, one can use electrostatic fields, but often one prefers to accelerate H^- ions and remove the electrons with a thin metal foil to convert these ions into protons. The magnetic field deflects the positive proton in the opposite direction from the negative H^- ion and therefore immediately ejects it from the magnet.

Fig. 3.10 Pole pieces of the GANIL injector cyclotron. The hills and valleys creating the azimuthally varying fields are given a spiralling shape to further enhance the focusing effect. Photograph by courtesy of Eric Baron [11]

Fig. 3.11 (*Top*) The first successful cyclotron was built by Lawrence and Livingston in 1932. It measured 13 cm in diameter and accelerated protons to 80 keV. Figure by courtesy of the Lawrence Berkeley National laboratory. (*Bottom*) Modern isochronous cyclotron for proton therapy producing protons beams of up to 230 MeV. The external diameter of the magnet is 434 cm. Figure by courtesy of IBA
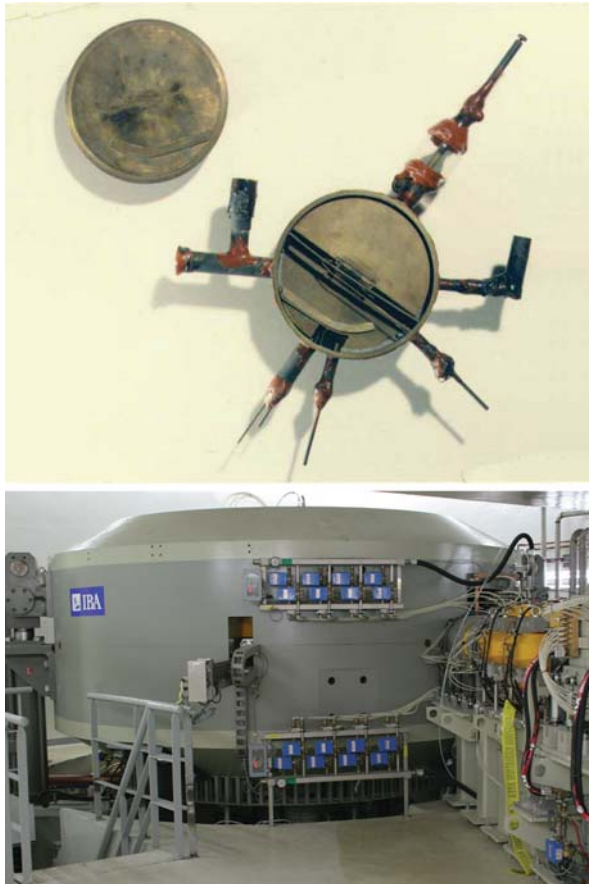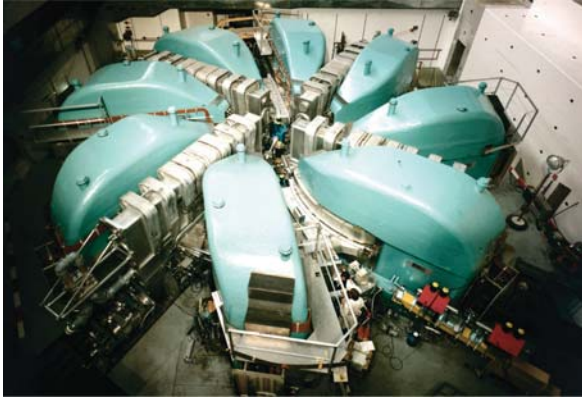
Fig. 3.12 One of the most powerful cyclotrons in the world today is at the Paul Scherrer Institute in Switzerland near Zurich. The PSI 590 MeV ring cyclotron for high-intensity proton beams started its operation in 1974. It is based on a design proposed by Hans Alexander Willax [12] in 1963. Originally designed for 100 microampere, it now produces beam intensities of 2 mA, with a further increase in preparation. The 20-fold increase became possible by replacing the original injector (a one solid pole cyclotron) with a new separate sector cyclotron and by increasing the RF power for acceleration. The main parameters of the machine are given in Table 3.5. Figure by courtesy of the Paul Scherrer Institute, Switzerland

Table 3.5 Main parameters of the cyclotron at PSI Switzerland

Injection energy	72 MeV
Extraction energy	590 MeV
Extraction momentum	1.2 Gev/c
Energy spread (FWHM)	ca. 0.2%
Beam current	2 mA DC
Accelerator frequency	50.63 MHz
Time between pulses	19.75 ns
Bunch width	ca. 0.3 ns
Extraction losses	ca. 0.03%

3.5 The Quest for the Highest Energy, Synchrotrons and Colliders

The maximum energy that can be reached with a cyclotron is limited by Eqs. (3.1) and (1.1). For a given magnet, there is a maximum radius the trajectory of the particles can have and therefore a maximum momentum and a maximum energy. To reach higher energies, one must either increase the magnetic field or increase the diameter of the magnet. Therefore, we will first discuss what magnetic fields can be achieved.

Most large magnets used in accelerators are electromagnets. The magnetic induction B in a solenoid, in MKSA units, is given by

$$B[\text{tesla}] = \mu_0 \, \mu_r \, I[A] \frac{N}{L[m]}$$

The notations are

N: total number of windings of the solenoid
μ_0: magnetic permeability of vacuum ($\mu_0 = 4\pi \times 10^{-7}$)
μ_r: relative magnetic permeability of the core
I: current in the windings
L: length of the solenoid

The shape of a magnet for a cyclotron is not a simple solenoid, but the maximum field that can be reached in the gap between two pole pieces is given by a similar equation. Note that throughout this book, the term 'magnetic field', will usually mean the magnetic induction B. If no ferromagnetic core is used, the maximum field that can be reached is limited by the amount of heat produced in the coils. Even with forced water-cooling, it is difficult to reach a magnetic field larger than ≈ 0.1 tesla. All magnets in accelerators therefore use a ferromagnetic yoke such that, for the same current, the field is increased by the relative magnetic permeability of the yoke material. The values of μ_r for some commonly used ferromagnetic materials are shown in Fig. 3.13. Soft iron has $\mu_r > 1000$ and allows much larger fields to be reached. However, in this case, the maximum field that can be reached is limited by the saturation of the ferromagnetic material. All ferromagnetic materials saturate at

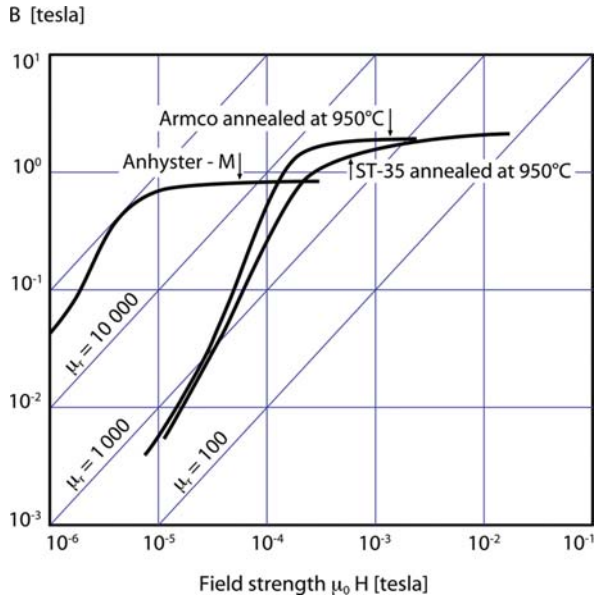Fig. 3.13 Magnetic induction B versus $\mu_0 H$ in a solenoid with a ferromagnetic core. The quantity $\mu_0 H$ is the magnetic induction that would exist in the solenoid in the absence of a ferromagnetic core. Armco and ST-35 are types of soft steel similar to the types of steel commonly used in magnets. Anhyster, often called mu-metal, is a nickel–iron alloy with extremely high values for μ

≈2 tesla, setting an upper limit of about 2 tesla to the field that can be reached with conventional (i.e. non-superconducting) magnets.

Superconducting magnets can reach higher fields, because almost no heat is generated in the coils; therefore, there is also no need for a ferromagnetic yoke. In this case the maximum field is limited by the properties of the superconducting material. Indeed, the magnetic field destroys the superconductive property of the superconducting wires. With presently used niobium-based superconductors ≈20 tesla is the upper limit for the magnetic induction that can be reached. Practical considerations limit the field to values well below this number.

When accelerating protons in a cyclotron with a conventional magnet with a diameter of 2 m, the maximum energy that can be achieved is ≈100 MeV (see Exercise 1). To achieve higher energies, the diameter of the magnet needs to be increased and the cost of the magnet will increase faster than the energy! A different approach is necessary. The solution is the synchrotron and we now discuss this type of accelerator.

Figure 3.14 shows the layout of a synchrotron. In a synchrotron, the particles have a fixed trajectory. The beam pipe has the shape of a torus and all around this ring there are bending magnets to keep the particles on the circular track. The particles must already have a minimum energy before they can be accelerated in a synchrotron. If a bunch of particles is injected into the synchrotron, these particles will circulate inside the torus. The magnetic field, the curvature of the track and the particle energy must obey equations (3.1) and (1.1). If the bending magnets are designed such as to have a magnetic field that slightly decreases with increasing radius, the trajectory of the particles will be stable, as we have shown when discussing the cyclotron. The bunch of particles can remain stored in this stable orbit for a very long time. To be used as an accelerator, the synchrotron also needs an RF cavity. An RF cavity is a large enclosure, usually made of copper, with a precise shape. In the box a standing electromagnetic wave is generated. The geometry of the RF cavity is such that this standing electromagnetic wave will have an oscillating electric field pointing either parallel or anti-parallel to the direction of the beam.
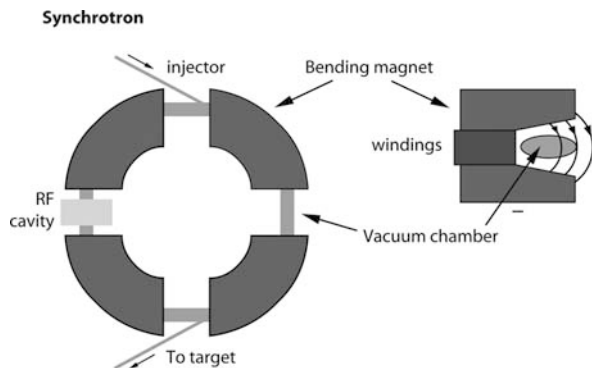

Fig. 3.14 Schematic representation of a simple synchrotron with weak focusing. The inclination of the pole pieces of the dipole magnet is exaggerated for better visibility on the drawing

When the beam passes through the RF cavity, the particle will either be accelerated or be decelerated, depending on its phase relative to the RF field. RF cavities will be discussed further in Sect. 3.6 on linear accelerators.

For the sake of argument, I will now assume that the stored particles are electrons of several 100 MeV. In this case, the energy and the momentum of the electrons are, to a very good approximation, related by $E = Pc$, and the velocity is very close to the velocity of light. The revolution frequency is determined by the length of the trajectory divided by the speed of the particle. The frequency of the RF cavity should be equal to, or a multiple of, the revolution frequency of the particles.

Figure 3.15 shows the field experienced by an electron passing through the RF cavity. Consider an electron passing exactly through the centre of the beam pipe.

This electron has the nominal trajectory and it will have the same phase relative to the RF field every time it passes the RF cavity. Consider now an electron with the nominal trajectory passing through the cavity with a phase as indicated by the arrow 'stable point'. This electron is neither accelerated not decelerated and it can continue to turn for a long time. Consider now an electron with the nominal trajectory passing through the cavity with the phase indicated by the arrow 'particle is accelerated'. This electron will acquire energy each time it passes through the RF cavity. The radius of curvature will become larger and the trajectory will become longer. It will need longer to make one full turn and each time it passes the RF cavity somewhat later. It will move in the direction of the 'stable point' and continue in that direction beyond the 'stable point'. There the electron will experience a decelerating electric field. It will lose energy and as a result the trajectory will become shorter and it will need less time to make a complete revolution. The electron will again move in the direction of the 'stable point'.

The result is that the energy of the electron will be making oscillations around the 'stable point'. All electrons will converge to this point and will therefore become grouped in bunches. The number of bunches is equal to the ratio of the RF frequency over the particle revolution frequency.

To turn the machine just described into an accelerator, all that is necessary is to increase the magnetic field in the bending magnets very slowly. If the magnetic field is slightly increased, the trajectory of the particles is shorter and the particles come early relative to the 'stable point' and will again experience an accelerating
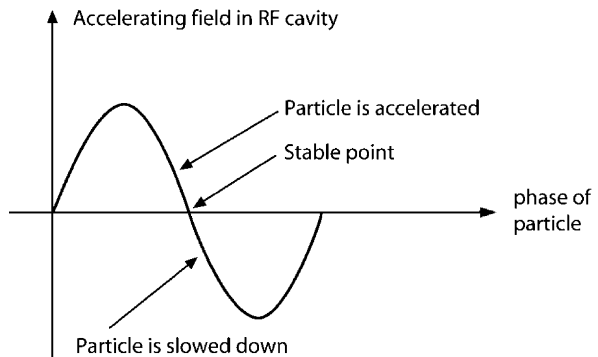
Fig. 3.15 Phase of the particle relative to the RF phase in a synchrotron

field. The electrons acquire more energy and a new equilibrium is reached, but this time with a slightly higher energy for the electrons. We can again slightly increase the field, the electrons will again acquire a higher energy, and so on. This can be continued until the maximum magnetic field in the bending magnets is reached.

To use a synchrotron as a particle accelerator, one must first turn the field in the bending magnets to some low value and inject a bunch of particles at the energy corresponding to the value of the field. It is not possible to make this field much lower that ≈ 0.1 tesla, because then it becomes impossible to accurately control the field parameters. After injection, the field is increased slowly to the maximum value possible. The beam is then ejected to the target area. The acceleration cycle can be quite long, e.g. it is 12 s for the CERN SPS synchrotron with a maximum energy of 450 GeV.

So far, we have been considering the acceleration of high-energy electrons. These electrons move at a speed very close to the speed of light, such that the revolution time of a particle with the nominal trajectory will always be the same, regardless of the energy. The frequency of the RF cavity is constant. However, protons of the same energy will have a speed that is lower and the speed will change with energy. As a result, for protons it is necessary to slightly change the frequency of the RF cavity during the acceleration cycle.

The synchrotron we have just described is a 'weak focusing' synchrotron, and such machines have been used in the past. Modern synchrotrons, however, use a different focusing system called 'strong focusing'. Strong focusing is based on the use of quadrupole magnets such as shown in Fig. 3.16. A quadrupole magnet has four poles that are alternatively of north and south magnetic type. The beam passes through the magnet perpendicularly to the plane of the drawing and the centre of the beam passes through the centre of the magnet. From the geometry it is clear

Fig. 3.16 (**a**) Schematic representation of a quadrupole magnet. The dark area represents the soft iron yoke of the magnet. The direction of the current in the windings is such that the poles are alternatively of the south and north type. The lines with arrows are magnetic field lines. (**b**) Quadrupole magnets in a beam transport line. Photograph by courtesy of IBA

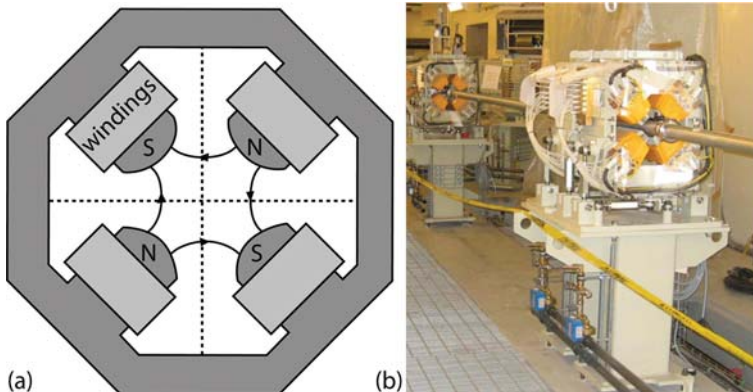that the magnetic field in the centre is zero. A few magnetic field lines are drawn in Fig. 3.16. Let us assume that a beam of protons is passing through the quadrupole magnet shown in Fig. 3.16 from the front to the back. From simple inspection of the direction of the field lines, we see that the Lorentz force on the beam particles is focusing in the horizontal plane and defocusing in the vertical plane. With a correct shape of the pole pieces of the magnet, the magnetic field will increase linearly with distance from the centre. A quadrupole magnet system will therefore behave like an optical lens for the particle beam, except that it will be focusing in one plane and defocusing in the other plane. We now have to recall a well-known property of optical lenses. Two optical lenses, one focusing and one defocusing, will behave like a focusing lens, regardless of the order of the two lenses. More generally, if we have two lenses with focal length f_1 and f_2, separated by a distance d, this system of lenses behaves like a single lens with focal length F given by

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2}$$

A positive value for f means a focusing lens and a negative value a defocusing lens. A doublet of two lenses with equal focal length, one focusing and one defocusing, will behave like a focusing lens with focal length given by

$$F = \frac{f^2}{d}$$

To obtain a system that will focus the beam both in the horizontal and vertical direction, we only need to use two identical quadrupoles with reversed magnetic fields. The focusing effect of a doublet in both planes is not symmetric, because the two effective lenses will seem to be displaced by a distance $2f$ relative to one another. Therefore, it is often preferable to use triplets of quadrupoles. In such a triplet, we have lenses with focal lengths $(2f, -f, 2f)$ in one plane and $(-2f, f, -2f)$ in the other plane. It is straightforward to show that such a triplet will also behave like a focusing lens in both planes. For triplets the optical properties of the two planes are much more similar.

All very high-energy accelerators in use today are synchrotrons. The layout of these machines is similar to what is shown in Fig. 3.14, except that there are many more bending magnets and sets of quadrupole lenses are added between the bending magnets. The advantage of using quadrupoles is that a much stronger focusing effect on the beam is obtained. This will result in a larger particle flux and/or a smaller beam pipe diameter and therefore smaller magnets. This type of synchrotron is therefore called a 'strong focusing synchrotron'. As an example of a synchrotron, the properties of the Super Proton Synchrotron (SPS) of the international research centre CERN are listed in Table 3.6. This accelerator is entirely installed in an underground tunnel near the city of Geneva, Switzerland. This accelerator was completed in 1976 and will certainly still be used for many years. All the magnets are conventional magnets. Figure 3.17 shows a view of the inside of the tunnel housing the accelerator.

Table 3.6 Main properties of the CERN Super Proton Synchrotron

Ring diameter	2.2 km
Revolution time	23 µs
Maximum proton energy	450 GeV
Number of bending magnets	744
Number of quadrupoles	316
RF cavities	4 of 20 m, $f = 199.4$ to 200.2 MHz
Intensity	10^{13} protons per cycle
Acceleration	3 MeV/turn
Injection energy	10 GeV
Duration of one cycle	12 s

Fig. 3.17 Inside the tunnel housing the Super Proton Synchrotron (SPS) in CERN. The blue blocks are quadrupoles and the red blocks are bending magnets. All magnets are conventional ferromagnets. Photograph copyright CERN

Proton synchrotrons are the highest energy accelerators available in the world. The energy is limited by the maximum magnetic field obtainable in the magnets and the diameter.

Such high-energy accelerators are mainly used for fundamental research. The aim is to study interactions between elementary particles such as protons on protons or electrons on positrons at the highest possible energy. The energy that really matters is the energy in the centre of mass system.

Consider a beam of particles with mass m_b and energy E_b colliding with particles of mass m_t at rest. For any particle, the quantity $E^2 - \vec{P}^2 c^2$ is an invariant. If we consider a system of two particles, the quantity

$$(E_1 + E_2)^2 - (\vec{P}_1 + \vec{P}_2)^2 c^2$$

is also an invariant. Evaluating the quantity in the centre of mass frame we see that this invariant is simply the square of the centre of mass energy. We therefore have

$$(E_b + m_t c^2)^2 - P_b^2 c^2 = E_{cm}^2$$

$$E_{cm} = \sqrt{(E_b - m_t c^2)^2 - P_b^2 c^2}$$

For a very high energy accelerator the beam energy is much larger than the mass of the beam and target particles, and the centre of mass energy can be written as

$$E_{cm} \approx E_b \sqrt{\frac{2 m_t c^2}{E_b}}$$

We see that the centre of mass energy is smaller than the beam energy by a large factor. The importance of the effect increases with the energy. It is much better to use two beams of particles colliding head on. In that case, the laboratory system is also the centre of mass system and all the energy is useful.

A particle accelerator arranged in such a way as to allow studying head-on collisions between two beams of particles is called a 'collider'. One could build two accelerators and let them send two particle beams against one another. A better way is as follows: We have already seen that a synchrotron accelerator can store particle beams. In addition, particles and their anti-particles have the same mass but opposite charge and these can circulate in the opposite direction in the same machine. If we have two stored beams, the particle bunches will meet each other at a number of points along the circular orbit. How many events will we observe when two beams meet each other? Let us assume that the bunch of particles has the shape of a cylinder with length l and section s. Let us further assume that the particle density all over this volume is constant and that the two bunches meet head-on. The total number of particles in the bunches are n_1 and n_2. Using the definition of the cross section (Eq. 2.1) and assuming there is only one particle in bunch 2, we have

$$dW = \frac{n_1}{s\,l} \sigma\, l = \frac{n_1}{s} \sigma$$

If there are n_2 particles in bunch 2, the number of interactions that will be observed is

$$\frac{n_1\, n_2}{s} \sigma$$

To have a large number of events there must be as many particles as possible in each beam and the transverse dimensions of the beam should be as small as possible.

If the beams meet with a frequency f, the number of interactions per second will be given by

$$f \frac{n_1 n_2}{s} \sigma = L\sigma$$

The quantity $L = f \frac{n_1 n_2}{s}$ is called the luminosity.

Of course, the density of particles in the beam is not a constant over the volume of the beam, but rather has a Gaussian shape with standard deviation σ_h and σ_v

Table 3.7 Main parameters of the LHC accelerator and collider

Start date	2009
Circumference	26659 m
Maximum beam energy	7 TeV
Luminosity	10^{34} cm^{-2} s^{-1}
Number of bending magnets/ring	1232
Maximum field in bending magnets	8.3 tesla
Length bending magnets	14.3 m
Number of quadrupoles	482 (2 in 1) + 24 (1 in 1)
Time between collisions	24.95 ns
Bunch radius	16.6 μm
Bunch length	7.55 cm
Injection energy	450 GeV
RF frequency	400.8 MHz
Particles per bunch	1.15×10^{11}
Number of bunches/ring	2808

in the horizontal and the vertical directions. It is straightforward to show that the expression for the luminosity in this case becomes

$$L = f \frac{n_1 n_2}{4\pi \, \sigma_h \sigma_v}$$

As an illustration we give the main parameters of the new Large Hadron Collider (LHC) now being finalised in CERN. It will be the largest energy accelerator in the world and is due to start operation at the end of 2009. This machine is designed to study proton–proton collisions. Collisions of Pb on Pb ions will also be possible. LHC has two interleaved accelerators that cross the beams under a small angle in eight points along the rings. The accelerators are synchrotrons that will first accelerate the beam to 7 TeV and then store the beams to let the protons interact in the crossing points. The parameters of the LHC accelerator and collider are summarised in Table 3.7.

3.6 Linear Accelerators

Circular machines such as cyclotrons and synchrotrons have been very successful, but linear accelerators have not been abandoned. There are many reasons for this. In 1928 when Wideröe first proposed linear accelerators, the technology did not allow the production of very high frequency, high power electromagnetic fields. This technology was developed in the 1940s and 1950s, mainly for the use in radar systems. Today linear accelerators can be much shorter than what was possible in 1930. Moreover, cyclotrons are not well suited for the acceleration of electrons. The reason is the γ factor in Eq. (3.2). For electrons, this effect is very important and the only possible solution is using the cyclotron as a synchrocyclotron, but this results in a strong reduction of the beam intensity. As a result, linear accelerators

are the preferred solution for accelerating electrons up to a few 10 MeV. The synchrotron is well suited for accelerating electrons up to a few GeV. However, at much higher energy the 'synchrotron radiation' makes it more and more difficult to use circular accelerators for the acceleration of electrons or positrons. The phenomenon 'synchrotron radiation' is briefly discussed below.

Any charged particle undergoing acceleration will emit electromagnetic radiation. If a charged particle is travelling in a circular orbit in a magnetic field, it is being accelerated in the direction perpendicular to its direction of motion and the particle will emit electromagnetic radiation. The energy radiated per turn by a particle of mass m, unit charge, energy E, and velocity v is given by

$$\frac{\Delta E}{turn} = \frac{4\pi}{3} \frac{\alpha \, \hbar c}{r} \, (v/c)^3 \left(\frac{E}{mc^2}\right)^4$$

Notice the effect is proportional to $(E/mc^2)^4$, therefore synchrotron radiation is usually negligible for all particles except electrons and positrons.

For electrons this equation can be written in convenient units as

$$\frac{\Delta E}{turn}[\text{MeV}] = 0.0885 \frac{\text{E}^4[\text{GeV}]}{\text{r[m]}} \tag{3.3}$$

The electromagnetic radiation is emitted with a broad spectrum with a maximum around

$$h\omega[\text{keV}] \approx \frac{\text{E}^3[\text{GeV}]}{\text{r[m]}} \tag{3.4}$$

We can see that the amount of energy radiated becomes quite large if the energy of the electron exceeds 1 GeV. This equation makes it clear that it is impossible to accelerate an electron to very high energy in a circular accelerator!

The accelerating structures in synchrotrons or linear accelerators consist of one or more resonant radio frequency cavities. In a synchrotron, the magnetic field bends the beam such that the particles pass many times through the same accelerating element. In a linear accelerator, the beam passes the same structure only once and one usually needs a large number of RF cavities.

We will now briefly discuss the main properties of such resonant RF cavities. Electromagnetic waves can be induced inside any conducting box. We are all familiar with this phenomenon, because a microwave oven is nothing else that a conducting box in which electromagnetic waves are induced. For the purpose of particle acceleration, we are mainly interested in resonant waves in cylindrical cavities, and we will discuss this case in more detail. Consider a box with walls made of conductive material. Any electromagnetic wave in this cavity should satisfy the Maxwell equations in vacuum with the boundary condition that, at the surface of the cavity, the electric field component parallel to the surface must vanish. The general solution of the Maxwell equations in a cavity is rather lengthy and we will only consider the most important solution for the purpose of particle acceleration. For this

application, we naturally consider a resonating cavity with the shape of a cylinder with length 'L' and radius 'R_0'. Because of the cylindrical geometry we will use cylindrical coordinates, 'z' denoting the direction along the axis of the cylinder, 'r' the distance to the axis of the cylinder and 'θ' the angle in the plane perpendicular to the axis of the cylinder. The Maxwell equations are

$$\vec{\nabla} \bullet \vec{D} = \rho$$
$$\vec{\nabla} \times \vec{H} - \frac{\partial \vec{D}}{\partial t} = \vec{J}$$
$$\vec{\nabla} \bullet \vec{B} = 0$$
$$\vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0$$

Using the well-known vector relation

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla}(\vec{\nabla}\vec{A}) - \vec{\nabla}^2\vec{A}$$

it is straightforward to show that the electric field should satisfy the following differential equation

$$\vec{\nabla}^2\vec{E} - \frac{1}{c^2}\frac{\partial^2\vec{E}}{\partial t^2} = 0 \text{ with } c = \frac{1}{\sqrt{\varepsilon_0\mu_0}} \tag{3.5}$$

We are looking for solutions of Eq. (3.5) satisfying the following conditions:

(1) Stationary solutions (also called standing wave solutions), i.e. solutions where the electric field can be written as a product of a spatial and a temporal function

$$\vec{E}(z, r, \theta, t) = \vec{f}(z, r, \theta) \times g(t)$$

(2) Solutions where the electric field has no longitudinal or azimuthal variation
(3) Solutions satisfying the boundary condition that the component of the electric field parallel to the surface of the cavity vanishes

From these conditions it immediately follows that only the component of the field in the z-direction is different from zero and that the function $f(z, r, \theta)$ is a function of 'r' only. Using the first condition, Eq. (3.5) can be written as the sum of two terms, one term depending only on 'r' and one term depending only on 't'.

$$\frac{\nabla^2 f(r)}{f(r)} = \frac{1}{c^2}\frac{\frac{\partial^2 g}{\partial t^2}}{g(t)}$$

This equation can only be satisfied if both terms are equal to the same constant. As it will turn out, this constant must be negative in order to obtain a stationary solution and therefore the constant is written as '$-k^2$'. Writing the Laplace operator in cylindrical coordinates we obtain the following two equations

$$\nabla^2 f(r, \theta, z) = \frac{1}{r}\frac{\partial f}{\partial r} + \frac{\partial^2 f}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2}$$

$$\begin{cases} \dfrac{\partial^2 g}{\partial t^2} + c^2 \, k^2 \, g(t) = 0 \\[2mm] \dfrac{\partial^2 f}{\partial r^2} + \dfrac{1}{r}\dfrac{\partial f}{\partial r} + k^2 \, f(r) = 0 \end{cases}$$

The solution for the first equation is given by: $g(t) = A \exp(i\omega t)$ with $\omega = kc$. The second equation is an equation of the Bessel type with $\alpha = 0$ and the solution is a linear combination of the zero-order Bessel functions $J_0(kr)$ and $Y_0(kr)$. The Y_0 term is eliminated by the requirement that the electric field has to have a finite value on the axis. A solution for the electric field satisfying Eq. (3.5) and satisfying the three conditions therefore exists and is given by

$$\begin{cases} E_{zn}(r, t) = E_{0n} \, J_0\,(k_n r) \, \exp(i\omega t) \\ E_r = 0 \\ E_\theta = 0 \end{cases}$$

The constant k is not a free parameter, but is constrained by the requirement that the electric field must vanish on the surface of the cylinder mantel, i.e. $J_0(k \, R_0) = 0$. There is one possible value of k for each zero of the Bessel function, and this is indicated by the index n. If the first zero of the Bessel function coincides with the wall of the cylinder, we obtain the condition $k_1 \times R_0 = 2.405$. The other values of k_n are given by

$$\begin{aligned} k_2 \times R_0 &= 5.520. \\ k_3 \times R_0 &= 8.654. \\ k_4 \times R_0 &= 11.792. \\ &\text{etc.} \end{aligned}$$

The magnetic induction B obeys a similar equation as the electric field, and the solution to this equation can be derived from the fourth Maxwell equation. This gives

$$\begin{cases} B_z = 0 \\ B_r = 0 \\ B_{\theta n} = (-i/c)E_{0n} \, J_1(k_n r) \, \exp(i\omega t) \end{cases}$$

Together with the conditions $k_1 \, R_0 = 2.405$ and $\omega = kc$ this completes the solution of the problem. The shape of the corresponding electric and magnetic fields is illustrated in Fig. 3.18. This mode of oscillation is referred to in the literature as the TM_{010} mode. Here TM stands for 'transverse magnetic'. Clearly it should be understood that the true fields are the real parts of the expressions above. For the TM_{010} mode, the frequency of the oscillation is given by
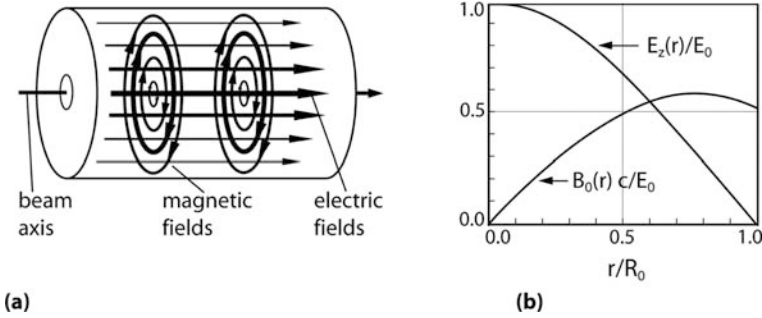
Fig. 3.18 (**a**) Geometry of the electric and magnetic field for the TM_{010} mode standing wave in a cylindrical cavity. The electric field is oriented in the axial direction and the magnetic field circles around the axis. (**b**) Dependence of the electric and magnetic fields on the radius r for the TM_{010} mode

$$f = 2.405 \frac{c}{2\pi R_0}$$

The symbol R_0 stands for the radius of the cylinder and c is the velocity of light. The length of the cylinder does not enter in the equations. However, if this RF cavity is to be used for accelerating particles, the length is constrained by the condition that the particle should only see an accelerating field while passing through the cavity. The time spent inside the cavity must be less than, or equal to, $\frac{1}{2}$ period of the oscillation. For a particle of velocity 'v', we therefore have

$$L \le \frac{v}{2f} = \frac{\pi R_0}{2.405} \frac{v}{c}$$

A numerical example is instructive. If the frequency of the RF cavity is 200 GHz, the diameter of the cylinder is 115 cm. For accelerating particles that travel at almost the speed of light, the length of the cavity should be less than or equal to 75 cm. These dimensions scale inversely proportional to the frequency.

In its simplest geometry a linear accelerator hence consists of a series of aligned RF cavities as shown in Fig. 3.19. If each cavity has the maximum allowable length, the phase difference between two successive cavities must be equal to π. In the accelerating structure we must create a stationary wave with a phase difference of π between any two successive cavities. Sometimes one prefers to use shorter cavities

Fig. 3.19 Very schematic representation of a standing-wave linear accelerator structure
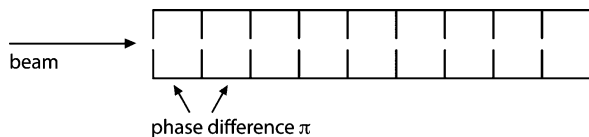
and in that case the phase difference between two successive cavities is less than π. This corresponds to an RF wave travelling along the structure.

A very important consideration in accelerating structures is the heat dissipation. Let us again consider a cylindrical cavity. The energy content and the heat dissipation of a stationary wave inside this cavity can be calculated as follows: The energy present in the electromagnetic wave is switching back and forth between the electric and the magnetic fields. When the electric field reaches its maximum, the magnetic field is zero and vice versa. The calculation of the energy content of the electromagnetic field 'U' is therefore reduced to a volume integral over the electric field at its maximum value

$$U = \int \int_{\text{volume}} \int \frac{\varepsilon_0 \, E_{\text{max}}^2}{2} \, dv$$

$$U = L \int_0^{R_0} (\varepsilon_0 [E_0 \, J_0(kr)]^2 / 2) \, 2\pi r \, dr$$

$$U = (\pi R_0^2 L)(\varepsilon E_0^2 / 2) \, J_1^2(2.405)$$

The heat dissipation is due to Ohmic heating caused by the currents in the walls of the cavity. These currents can be found by considering the second Maxwell equation in the integral form over the loop shown as a dotted line in Fig. 3.20.

$$\oint \vec{H} d\vec{l} = \int \int \left(\vec{j} + \frac{\partial \vec{D}}{\partial t} \right) ds$$

The line integral of the magnetic field for the part of the loop inside the metal is zero, because the magnetic field has no time to penetrate in the metal. Also the surface integral of the derivative of the dielectric displacement D over the whole surface of the loop is zero, because the tangential component of the electric field is zero. Therefore, the magnitude of the current in the metal is equal to the magnetic field component parallel to the surface and the orientation of the current is perpendicular to this magnetic field. Consider an infinitesimal surface element $dxdy$ as shown

Fig. 3.20 The tangential component of the magnetic field on the surface of the cavity is related to the current inside the cavity wall. (**A**) The loop integral over the dotted line gives a relation between the magnetic field and the surface current in the wall, (**B**) Surface element of the cavity wall
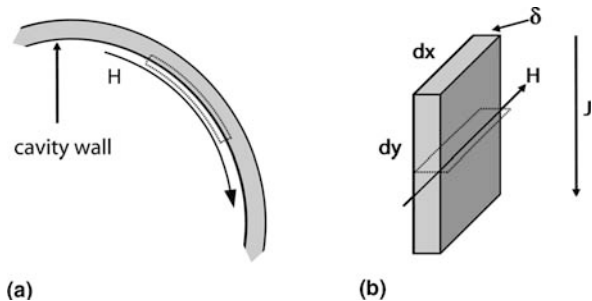
on Fig. 3.20(B). The current flowing through the surface element $dxdy$ is given by $J = Hdx$ and is flowing in the y direction. This current will be restricted to a thin layer of thickness δ by the skin effect. The skin depth δ depends on the frequency f and on the resistivity of material. If ρ is the resistivity of the wall of the cavity, the resistance 'R' of the surface element $dxdy$ is given by

$$R = \rho \frac{dy}{\delta \, dx}$$

The power dissipation in this surface element is given by

$$dW = RJ^2 = \rho \frac{dy}{\delta \, dx} (H \, dx)^2 = \frac{\rho}{\delta} H^2 \, dxdy$$

The total power dissipation in the walls of the cylinder is therefore given by

$$W = \frac{\rho}{\delta} \iint H^2 \, ds$$

The magnetic field is proportional to $\sin(\omega t)$ and only the time averaged power dissipation matters. Since we have

$$\frac{1}{2\pi} \int_0^{2\pi} \sin^2 (\omega t) \, dt = \frac{1}{2}$$

This time averaged power dissipation is given by

$$W = \frac{\rho}{2\delta \mu_0^2} \iint B_{max}^2 \, ds$$

The integral is to be taken over the total surface of the cylinder.

For a numerical example, consider again the cylindrical cavity oscillating at 200 MHz and with a length of 75 cm; the total power dissipation for a maximum electric field of 1 MV/m is $\approx$ 20 kW! To find this result we used that the resistivity of copper is $\rho = 1.7 \, 10^{-8}$ Ωm and that the skin depth of copper is given by

$$\delta = \frac{66 \, \text{mm}}{\sqrt{f}}$$

Notice that the power dissipation is increasing proportionally to the square of the electric field. Notice also that the power dissipation per unit length of the accelerator is decreasing inversely proportional to the square root of the frequency.

The Q value of an oscillator is the energy content of the system divided by the energy dissipated in one half oscillation. This quantity is a measure of how accurately the oscillation frequency is determined. In our example the energy content of the field is $\approx$ 1 J; therefore $Q = 1.6 \times 10^{-4}$.

The above calculation underlines the importance of the power dissipation in accelerating cavities. The maximum field that can be maintained in a cavity is ultimately limited by the extraction of electrons from the metal walls. This phenomenon depends on the surface smoothness and the practical limit is about 100 MV/m. At such fields the power dissipation is enormous, of the order of 300 MW per meter! It is clear that it is impossible to operate cavities continuously at this value of the accelerating field.

As an example we will describe the structure of SLAC 2-mile linear accelerator. This accelerator is located near San Francisco, California, and it is the largest linear electron accelerator in the world today. The accelerating structures in this accelerator are of the travelling wave type and consist of a series of coupled oscillating cavities that support travelling waves. This structure is illustrated in Fig. 3.21. Inside each of the cavities there is an oscillating electromagnetic field with a geometry similar to what we discussed before. The oscillations in each cavity are coupled and the distance between the discs and the diameter of the iris adjusts the degree of coupling. This coupling determines the phase velocity along the structure.

The accelerating structure is made of pure copper and each element is about 3 m long. A high frequency wave from a high power klystron is injected at one end. The phase velocity equals the speed of the electrons, so that all the electrons stay in phase with the electric field in the cavities as they travel along the structure. The wave is attenuated because of resistive losses in the copper. After 3 m, the amplitude is reduced so much that it is not useful to make the structure any longer. The remaining energy is absorbed in a dump. The average accelerating electric field in the structure is $\approx$15 MeV/m.

The main properties of SLAC accelerator are listed in Table 3.8. The heat dissipation in the structure is of the order of 1 MW/m. It is clear that the amount of heat produced is so large that it is impossible to use the accelerator continuously. The accelerator therefore produces beam pulses lasting 2.5 μs with a repetition rate of 360 Hz.

Iris-loaded waveguides cannot be used for accelerating particles with a velocity much below the speed of light. If the particles travel at a few percent of the speed of light, each cavity becomes much shorter than its diameter, and it becomes difficult to avoid exciting other modes than the T_{010} mode in the structure. A possible way to
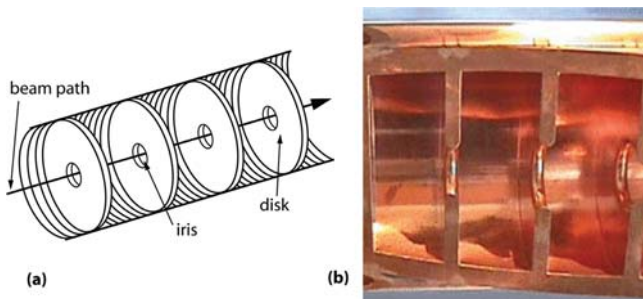

Fig. 3.21 (a) Geometry of the accelerating structure in an electron accelerator of the travelling wave type. (b) Cut away view of a section of the 'SLAC National Accelerator Laboratory' electron accelerator showing the internal structure

Table 3.8 Main properties of the SLAC linear accelerator

Length	3100 m
Structure	Iris-loaded wave guide
Outer diameter wave guide	10.5 cm
Diameter iris	1.9 cm
Number of accelerating sections	960
Number of klystrons	245
Peak power per klystron	24 MW
Operating frequency	2.856 GHz
Maximum pulse rate	360 pulses/s
Pulse duration	2.5 μs
Average electron beam current	30 μA
Maximum energy	50 GeV

accelerate slow ions is the structure shown in Fig. 3.22(a). In this geometry, there are a number of cavities, and around the beam there are drift tubes shielding the beam from the electric field when it has the wrong orientation. In such an accelerator the length of the drift tubes must vary to stay in step with the changing velocity of the particle. The shape of each cavity must therefore change in such a way that all the cavities oscillate at the same frequency. If the oscillations in the cavities are all in phase, the wall separating two successive cavities carries zero net current. This wall can therefore be omitted, leaving only bars to support the drift tubes. This is called an 'Alvarez structure', and it is shown in Fig. 3.22(b). In a linear accelerator with an Alvarez structure, a standing electromagnetic wave in a large conductive tube is created, with drift tubes in the centre containing the beam. Figure 3.23 shows the inside of an accelerating structure of the Alvarez type.

This type of linear accelerator is often used as an injector for large proton synchrotron accelerators. The linear accelerator will typically receive a beam of ≈1 MeV from an RF quadrupole accelerator, and it will accelerate the beam to a few 100 MeV. After this the energy is sufficient for the beam to be injected into
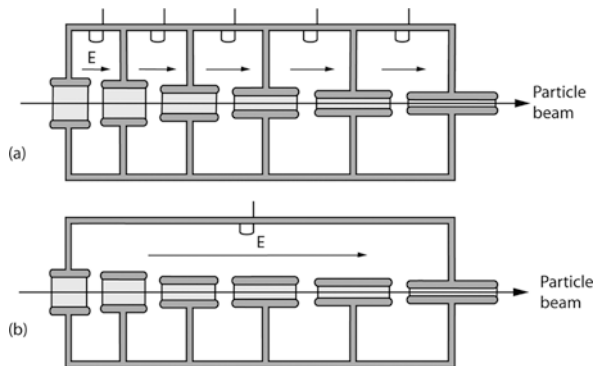

Fig. 3.22 (**a**) Hypothetical structure for a linear accelerator for non-relativistic ions. (**b**) Alvarez structure for a linear accelerator for ions
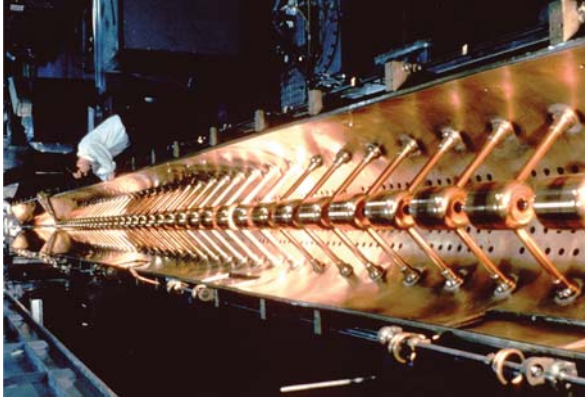
Fig. 3.23 This linear accelerator of the Alvarez type was used as an injector for the CERN acceler-
ator complex. It accelerated protons from 520 keV to 50 MeV. It is no longer in operation. Similar
devices are part of the standard layout of a modern synchrotron. It forms the next acceleration step
after the RF quadrupole and is used to bring the protons from $\approx$1 MeV to $\approx$50 MeV. Photograph
copyright CERN

a synchrotron. The RF cavities in Super Proton Synchrotrons also have a similar
geometry.

The most important recent advance in linear accelerator technology is the
development of superconducting accelerating cavities. Figure 3.24 shows such a
superconductive cavity working at 1.3 GHz. It consists of nine sub-cavities and in
each sub-cavity a standing wave with a field geometry similar to what is shown in
Fig. 3.18 is generated. There is a phase difference of 180° between any two succes-
sive cavities. If the centre-to-centre distance between two successive sub-cavities
equals the distance travelled by a particle in one half period of the oscillation, a
particle experiencing a maximum accelerating field in sub-cavity one, will again be
in phase with the field in the next sub-cavity two, and so on. Accelerating fields of
35 MeV/m have been achieved in such structures.

Fig. 3.24 Superconductive
accelerating cavity in
niobium for the acceleration
of electrons. Niobium is
superconductive at the
temperature of liquid helium.
This resonator was designed
for the TESLA test facility.
Photograph from [13], with
permission

3.7 Secondary Beams

So far we have only considered accelerators producing beams of electrons, protons or nuclei. These are the only charged particles that are easily available. However, often one is interested in other particles.

Positrons or gamma rays are produced in large numbers whenever an electron beam penetrates any piece of target material. If the electrons have more than 10 MeV of energy, they will produce a large number of gamma rays by bremsstrahlung. These gamma rays have a $1/E$ spectrum, but the photons with a very low energy are absorbed in the target. Most of the gamma rays are going in the direction of the initial electron beam. The higher the energy, the better the collimation in the forward direction. The gamma rays, in turn, will produce electron–positron pairs by the pair creation process. If the target has a thickness of a few tenths of a radiation length, the beam spot in the target will be a copious source of both gamma rays and positrons.

A magnet placed behind the target will remove the electrons and positrons, leaving only the gamma beam (see Fig. 3.25). This is the standard method for the production of gamma beams. This gamma beam will have a broad energy spectrum. By using suitable absorbers one can somewhat reduce the bandwidth of this spectrum.

This setup can also serve as a source of positrons. Positrons can be accelerated by any of the methods that can be used for accelerating electrons. However, there is a problem. Any accelerator has a very limited acceptance in energy and direction of the particles it will accelerate. The positrons produced in the target have a broad energy spectrum. In electron accelerators, the electron source produces large amounts of electrons coming from the same point and all with about the same energy, i.e. almost zero energy. The result is that the number of positrons available for acceleration is many orders of magnitude smaller than the number of electrons and can only produce fairly low-intensity positron beams. To increase the intensity
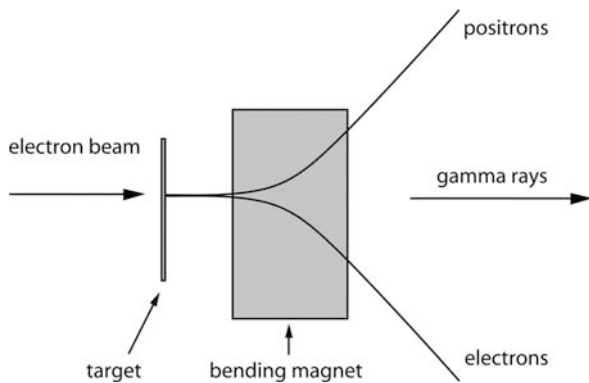

Fig. 3.25 Principle of the production of secondary gamma ray beams or positrons beams

of the positrons, the particles are often kept in a storage ring. This is basically a synchrotron, but optimised for storing particles rather than for accelerating particles. Immediately after injection, the positrons fill the complete phase space of position and momentum that the storage ring will accept. The positron will oscillate around their equilibrium trajectory; the amplitude of the oscillations being determined by the aperture of the magnets and the diameter of the beam pipe. However, positrons emit significant amounts of synchrotron radiation and this emission is equivalent to some sort of friction; therefore, with time the amplitude of these oscillations is damped, and the positrons all converge to the equilibrium trajectory. Because of the stochastic nature of the synchrotron radiation, the particles will not all end up in the equilibrium trajectory; there will remain some spread in position and momentum. If a second bunch of positrons is injected in the storage ring, they will at first have a different trajectory from those already stored, but with time they will converge to the same equilibrium trajectory. In this way one can accumulate a large number of positrons in one bunch and eventually have a positron bunch with the same intensity as that possible for electron bunches. Such a storage ring for positrons is often called a damping ring.

It is also possible to produce anti-protons by a method very similar to the one just described. If a proton beam of well above 10 GeV interacts with a target, a very large number of secondary particles will be produced. About 90% of these particles will be pions, the rest will be a collection of other hadrons; heavier hadrons being less abundant than lighter ones. Among these there will also be some anti-protons. With the help of magnets we can now select those anti-protons with an energy and direction interval that the next accelerator will accept. However, the number of anti-protons produced is very small and their energy is spread over a large range. Very few of these anti-protons will be accepted by the storage ring or the accelerator. Because of the absence of synchrotron radiation, there is no natural damping mechanism for anti-protons. There are methods to reduce the oscillations of such stored anti-protons, but discussion of these is beyond the scope of the present lecture notes.

It is also possible to produce beams with unstable particles such as pions, but because of their short lifetime it is impossible to store these particles in a storage ring. The intensity of such beams will therefore be many orders of magnitude lower than the intensity of proton beams. The principle of the production of secondary pion beams is illustrated in Fig. 3.26. To produce a beam of pions, a primary beam of protons is allowed to hit a solid target. A large number of secondary particles, mostly pions, are produced. With the help of bending magnets, collimators and quadrupoles, pions in a certain energy range are selected. To obtain a beam of muons, one starts from a beam of pions and keeps the pions with the help of quadrupoles in a sufficiently long decay tunnel. A pion decays in a muon and a muon–neutrino with an average decay path equal to 7.8 m multiplied by the relativistic gamma factor of the pion, $\gamma = E/m$. Depending on the energy of the primary proton, it can require several 100 m before most of the pions have decayed. At the end of the decay tunnel, the beam is sent into a dump to remove the remaining pions.
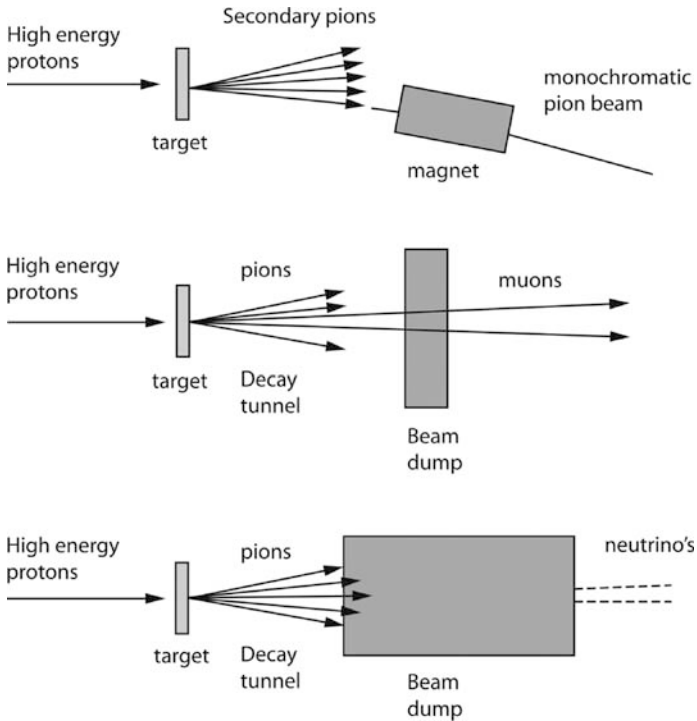
Fig. 3.26 Secondary beams of pions, muons and neutrinos can be made with the help of a primary proton beam

After about 10 hadronic interaction lengths, all of the pions have undergone interactions and the secondary hadrons have been absorbed, thus a beam containing only muons emerges from the dump.

To make a neutrino beam, the same set-up is used but with a much more massive dump. A sufficiently massive dump will not only stop all the pions but also all the muons. After such a massive dump only neutrinos are left.

All these secondary beams are available in the CERN accelerator complex shown in Fig. 3.27. In this accelerator complex, an RF-quadrupole accelerating section delivers protons of 750 keV. These protons are injected into a linear accelerator of the Alvarez type that brings the energy to 50 MeV. The protons are subsequently injected into a first synchrotron, the PS-booster that accelerates them to 1.4 GeV. A second synchrotron, called the PS, increases the energy further to 13.1 GeV. The Super Proton Synchrotron (SPS) accelerates the protons to 450 GeV. This last accelerator injects the protons into the (LHC) Large Hadron Collider, where the protons are finally accelerated to 7 TeV.

The PS and SPS each provide a wide variety of secondary beams for particle physics experiments. One neutrino beam is sent in the direction of the Gran Sasso laboratory at 730 km distance, where these neutrinos are detected.
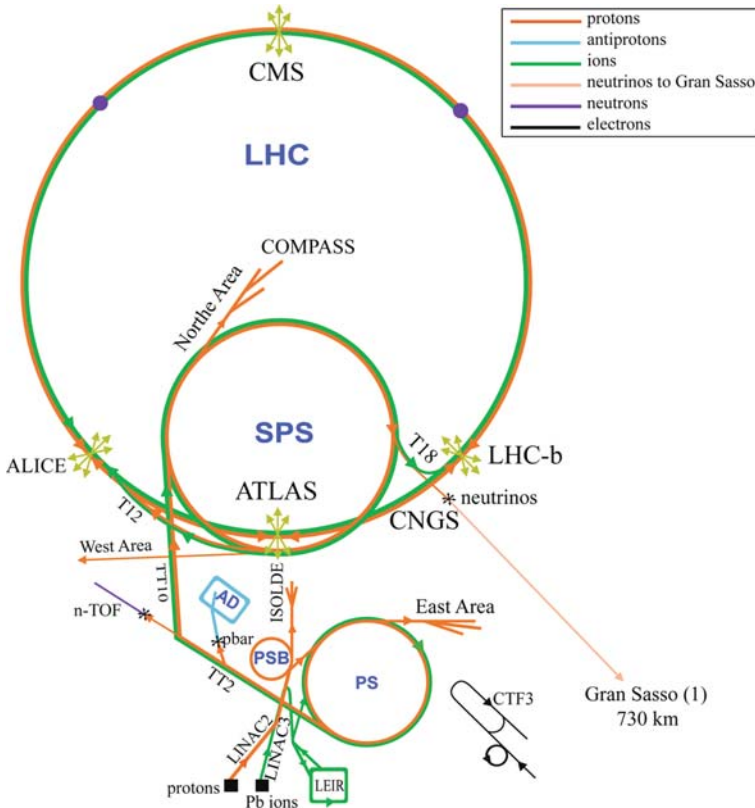
Fig. 3.27 The CERN accelerator complex provides protons up to 7 TeV and also provides a wide variety of other particle beams. Figure copyright CERN

3.8 Applications of Accelerators

The initial motivation for developing particle accelerators was to probe the structure of matter at the subatomic scale. However, these machines have found numerous applications in many other fields.

Accelerators are widely used for the implantation of ions in the semiconductor industry or for hardening steel objects such as ball bearings or cutting tools. Ions of tungsten, chromium, tantalum, nitrogen and boron, among others, are used for this purpose. This application usually only requires accelerating particles to a few 100 keV and electrostatic acceleration is the most economic approach.

When particles with a larger energy are required, it is more appropriate to use one of the methods for particle acceleration described in the previous sections.

Accelerators for medical applications. One important use of accelerators is as sources of X-rays and gamma rays. This is achieved by accelerating electrons and letting them interact with a target. The X-rays and gamma rays are produced in

the target by bremsstrahlung and by photoelectric effect followed by the emission of X-rays. The energy of the electrons, the target thickness and the target material together determine the spectrum of the gamma rays. Worldwide, many thousands of electron accelerators are in use in hospitals for gamma radiation therapy. The energy of these accelerators varies typically from 4 to 25 MeV. In radiation therapy, one needs an arrangement where the beam can come from all directions on a circle around the patient. By rotating the beam one ensures that a maximum of radiation dose is delivered at the position of the tumour and a lower dose is delivered to the surrounding tissue.

Figure 3.28(a) shows the general layout of an electron accelerator for this application. If the beam energy is not very large, the accelerator has a modest size and it is possible to rotate it around the patient with a rotating gantry as shown on Fig. 3.28(b). If the energy of the electrons is larger, the accelerator is very bulky. It is better to have the accelerator in a fixed position on the rotation axis of the gantry and steer the beam with magnets fixed on the gantry as shown in Fig. 3.28(c). Figure 3.29 shows an example of an accelerating cavity used in this type of electron accelerators. Very similar accelerators are used for the purpose of sterilisation of food or other materials.

Another approach to radiation therapy is the use of protons or light ions such as carbon. The motivation of this method comes from the property of protons or heavy ions to have an increase of the energy loss by ionisation towards the end to the trajectory. In this case there is also no radiation damage behind the irradiated area. One therefore obtains a distribution of the radiation dose as shown on Fig. 3.30. This is particularly important if the area to be irradiated is close to a vital organ. The range of the hadrons in tissue should be up to about 30 cm, and this requires 220 MeV and 5280 MeV for protons and carbon ions, respectively. This energy can be reached either with a cyclotron or with a synchrotron, and systems of either type are in use in several places. Figure 3.11 shows a cyclotron for used proton therapy. The limited integrated beam flux obtainable with a synchrotron is not a problem in this application because only a small beam flux is needed. Figure 3.31 shows the layout of a typical synchrotron for radiation therapy. Typically a synchrotron for hadron therapy consists of an ion source and an electrostatic potential that bring the protons to 80 keV. Next there is an RF-quadrupole accelerating structure and a drift tube linac that brings the protons to a few 10 MeV. After this point, the protons can be injected in the synchrotron to be accelerated to their final energy. One acceleration cycle takes 2 s and is followed by a slow extraction of the beam in a time that can be up to 10 s. The number of protons accelerated in one cycle is $\approx 10^{10}$, the number of carbon ions is $\approx 10^{9}$.

One of the most expensive components in such a treatment system is the massive gantry needed to bend the beam such that the patient can be irradiated from all directions. This requires heavy magnets that are rotating around the patient.

Another important medical application of accelerators is the production of isotopes for Positron Emission Tomography (PET). This medical imaging modality will be discussed in Chap. 6. To apply this imaging method, one needs to produce short-lived positron-emitting isotopes such as ^{18}F or ^{11}C with decay times of 109.7 min and 20 min, respectively. Because of the short lifetime of these isotopes
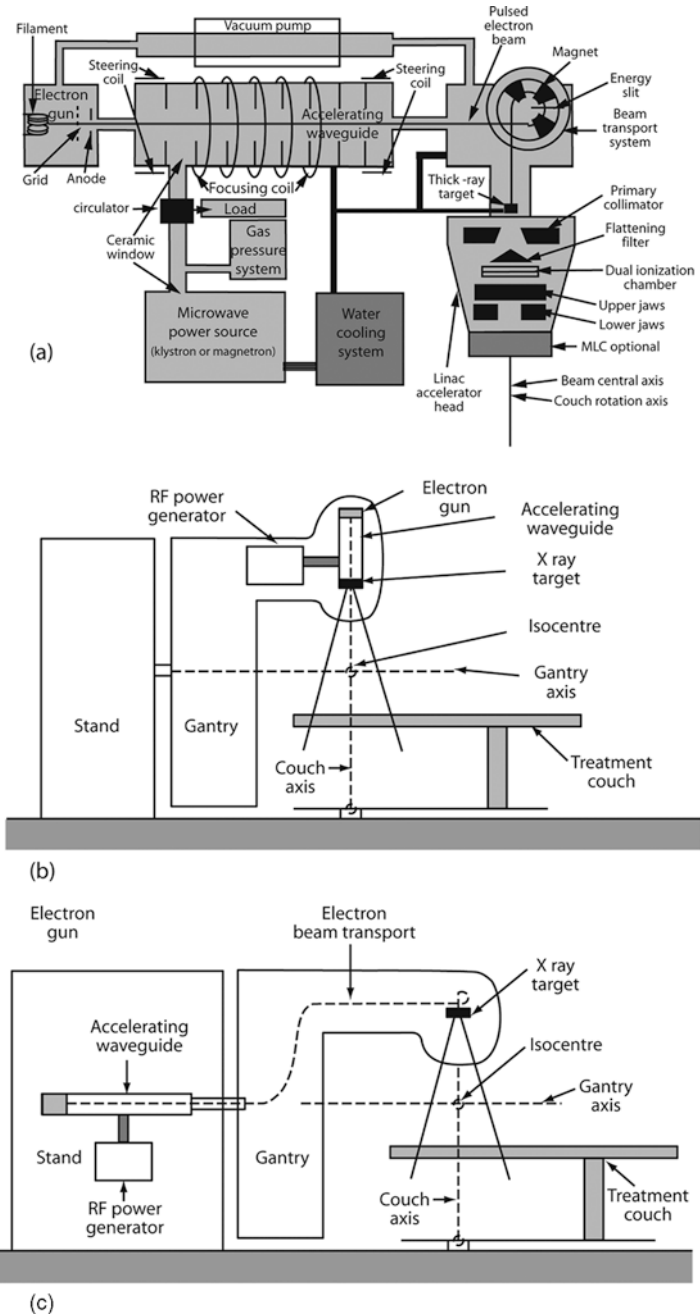
Fig. 3.28 Equipment for gamma radiation therapy. (**a**) Layout of the electron accelerator and target. (**b**) The electron accelerator is mounted on the rotating gantry. (**c**) If a higher energy is need the electron accelerator is mounted on the rotating gantry. (**c**) If a higher energy is need the electron accelerator is too large and is mounted on the axis of the gantry. In this case a beam transport system with magnets brings the electrons to the target. Figure by courtesy of IAEA

Fig. 3.29 Cutaway view of a standing wave accelerating waveguide for a 6 MV electron linac. The accelerating cavities are on the central axis; the coupling cavities are offside. The electron gun is on the left, the target on the right, both permanently embedded. Photograph by courtesy of IAEA
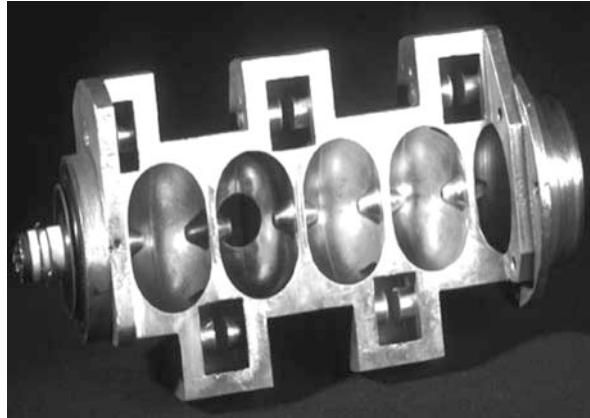
the cyclotron has to be near to the hospital. The energy needed for the production of isotopes is typically in the range 8–40 MeV and a cyclotron is perfectly suited for this application. Worldwide several 100 cyclotrons are being used for the production of isotopes for nuclear medicine.
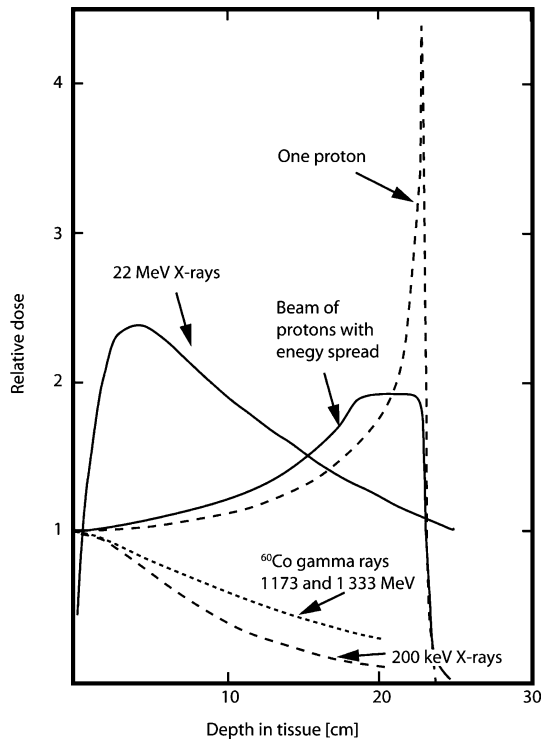
Fig. 3.30 Comparison of the dose delivered by protons and gamma rays in tissue
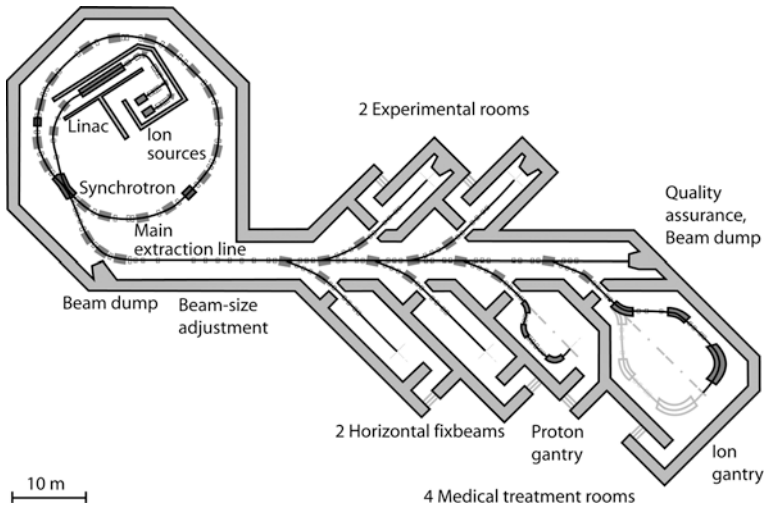
Fig. 3.31 Layout of the synchrotron for hadron therapy for the MedAustron project in Austria. Image courtesy of FOTEC Company, Viktor Kaplanstrasse 2, A-2700 Wiener Neustadt

Accelerators for nuclear power stations. A potentially very important application of proton accelerators is as 'drivers' for nuclear fission reactors in ADS (Accelerator Driven Systems). In a nuclear power station using this approach, the reactor core is sub-critical, i.e. not enough neutrons are produced by the fission process to keep a steady chain reaction going. A spallation source provides the additional neutrons to sustain the nuclear chain reaction. Figure 3.32 shows that the optimal proton energy for this application is about 1 GeV. At this energy about 25 neutrons are produced for each proton incident on a heavy target such as lead.
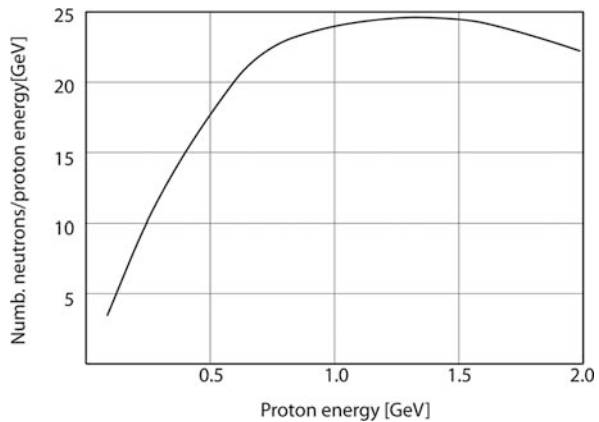
Fig. 3.32 Number of neutrons produced by a proton in a lead target divided by the proton energy in GeV. The data for this figure are obtained from [15]

The total power in the beam is therefore 10 MW or more. The advantages of this approach are the following

– The reactor is subcritical; as soon as the accelerator is switched of it will stop immediately. This will give an additional safety margin to the reactor.
– Different fuel compositions can be used compared to a conventional reactor. In particular spent fuel from reactors could be used, or thorium. Thorium is about three times more abundant in the world than uranium and consists of 100% of ^{232}Thorium. This isotope itself cannot sustain a nuclear chain reaction, but it can absorb neutrons and form the fissile isotope ^{233}U.
– The most attractive aspect of ADS is that the energy spectrum of the neutrons in such a reactor is different from the neutron spectrum in a conventional reactor, and as a result a large fraction of the very long-lived isotopes in the reactor will be converted to short lived ones. ADS systems built for the sole purpose of 'nuclear waste transmutation' are being considered.

For this kind of application, the average beam intensity needed is 10 mA or more, which excludes the use of synchrotrons and probably also the use of cyclotrons. The only proven accelerator technology allowing to reach the energy and beam intensity needed for ADS is the linear accelerator. The advantage of linear accelerators over cyclotrons is due to the fact that in a linear accelerator it is possible to obtain a strong focussing effect with the help of quadrupole magnets and this in turn allows obtaining much higher beam intensities. The much weaker focussing effect in a cyclotron limits the beam intensity that can be reached. However, the accelerating gradient that can be achieved in room temperature linear acceleration cavities is limited to about 1 MeV/m if we want to keep the power dissipation in the cavity walls at an acceptable level. Such a linear accelerator therefore will be more than 1 km long and will be too expensive. A possible solution is the use of superconducting linear accelerator structures as shown in Fig. 3.24. Such accelerating structures are used successfully in high-energy electron synchrotrons, such as synchrotron radiation sources. For particles travelling at a velocity close to the speed of light, average accelerating gradients of 25 MeV/m and conversion efficiencies of electrical power to beam energy as high as 50% have been routinely achieved. The theoretical limits for the accelerating gradient that can be reached with niobium cavities are 55 MeV/m, 44 MeV/m and 37 MeV/m for particles with $\beta = v/c$, of 1, 0.65, and 0.5, respectively. A value of $\beta = 0.5$ corresponds to a proton energy of 145 MeV. At lower energy it is necessary to have a different geometry for the accelerating cells, and it is more difficult to obtain large accelerating gradients. This fact, and the need to have a very reliable accelerator, probably limits the accelerating gradients that can be used to values well below 25 MeV/m. A considerable research effort is under way to design superconductive accelerators for ADS.

Circular accelerators such as fixed field alternating gradient (FFAG) accelerators are also being considered for ADS. This type of accelerator would be more compact and therefore possibly less expensive than linear accelerators. Demonstrating the feasibility of this approach to ADS is also an active field in accelerator research today. This type of accelerator will probably need a proton source followed by a RF-quadrupole accelerating structure and linear accelerator of the Alvarez type

accelerating the protons to about 50 MeV. The protons will then be injected in two (or more) circular fixed field machines in succession.

Synchrotron radiation sources: Synchrotron radiation is a major problem when trying to accelerate electrons to the highest possible energy, but the phenomenon is very interesting in its own right. It allows making very intense beams of electromagnetic radiation. The radiation is tuneable in frequency from the visible region to the X-ray region, by choosing the energy of the beam and the strength of the bending magnet. The emission is concentrated in a narrow cone with an opening angle given by $\approx m_e/E$ in the direction of motion of the electron. The radiation is also polarised.

The first synchrotron radiation sources were initially built as electron synchrotrons for research in high-energy physics, later converted to synchrotron radiation sources and these used the synchrotron radiation emitted from the bending magnets. The synchrotron is used as an accelerator and as a storage device. The beam is accelerated to the desired energy and kept in orbit for a long time. The lifetime of the electron beam in such a storage ring is several 10 h. The energy radiated by synchrotron radiation is compensated by the RF cavities.

Current third-generation synchrotron radiation sources typically have several insertion devices called wigglers or undulators. In an undulator the straight sections in the storage ring are used for inserting periodic magnetic structures composed of many magnets that form a repeating row of magnetic fields with alternating direction. Instead of a single bend, many tens or hundreds of 'wiggles' with a fixed period add up and multiply the total intensity that is seen at the end of the straight section. In such a structure stimulated emission can occur. It can therefore work as a free electron laser.

There are more than 50 electron accelerators in the world dedicated for the use as synchrotron radiation sources.

3.9 Outlook

There are two parameters determining what future accelerators will be possible: the maximum magnetic field that can be achieved and the maximum accelerating gradient that can be achieved.

Regarding magnetic fields, only modest progress is to be expected in the coming decades. With the currently used type of superconductors, 20 tesla is the upper limit for the field that can be achieved. New high temperature superconductors could possibly one day reach a field of 30 tesla, but experience shows that progress in superconductive magnets is very slow.

Regarding accelerating gradients, the situation is very different. It is well known that the electric field in lasers is many orders of magnitude larger than the maximum field that can now be reached in cavities. However, nobody has been able to come up with a realistic proposal on how to use these fields for the purpose of accelerating particles. The electric fields in plasma waves are also very large and this seems a realistic approach to particle acceleration. In a recent experiment, electrons were accelerated to 1 GeV in a plasma wave over a length of only 4 cm. The idea is

Fig. 3.33 Particle
acceleration based on plasma
waves. An igniter laser pulse
forms a 'wire' of plasma in a
plume of hydrogen gas; a
heater pulse expands the wire
to a plasma channel; the drive
pulse accelerates bunches of
electrons inside the channel
to nearly uniform high
energy. Figure by courtesy of
Berkeley Lawrence Lab [16]
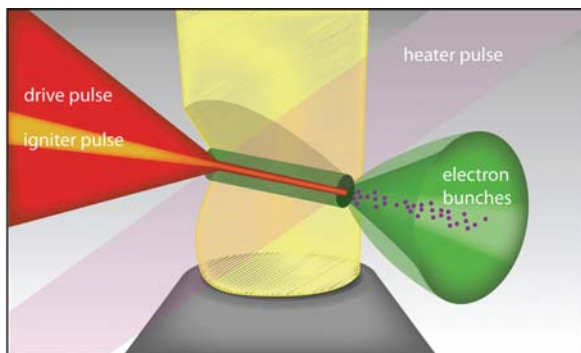
illustrated in Fig. 3.33. The plasma, and the plasma wave, are caused by a powerful
laser pulse in a low-pressure gas.

It is possible that electron accelerators based on this principle will replace con-
ventional linear accelerators as the source of electrons reaching a few 10 MeV of
energy. Such machines could be operational by the year 2020.

A somewhat similar idea is the plasma wake field accelerator. In this approach,
a high intensity low-energy electron beam excites the plasma. This beam causes a
strong wake field that can be used for accelerating particles. A high-intensity low-
energy electron beam could, in this way, give rise to a low-intensity high-energy
beam of electrons or other charged particles.

For the future of very high-energy accelerators for basic research, there are two
approaches: proton colliders, such as the present LHC machine and electron linear
accelerators used for studying electron–positron collisions. It is very unlikely that
a proton–proton collider of higher energy than the present LHC accelerator will be
built in the near future, if ever.

The international scientific community agrees that the next high-energy accel-
erator to be built should be the International Linear Collider (ILC). This machine
will be based on the use of superconductive accelerator cavities similar to the one
shown in Fig. 3.24. The accelerating field will be 31.5 MeV/m. The main parame-
ters of the ILC are summarised in Table 3.9. Such a machine could optimistically be
operational by 2019. ILC will allow the study of electron–positron collisions with
a centre of mass energy of 500 GeV. This seems modest compared to the LHC, but
a proton is composed of quarks, it is not an elementary object. What matters is the
energy in the quark–quark collision and this energy is only about 1.2 TeV at LHC.
The interpretation of events observed in proton–proton collisions is also much more
difficult than the interpretation of electron–positron collisions.

A much more ambitious proposal is the CLIC project [18], see Table 3.10. It
aims at building an electron–positron collider that could reach a total centre of
mass energy of 3 TeV. The two accelerators and associated equipment could fit in
a site 38 km long. Figure 3.34 shows the layout of the CLIC collider and Table 3.9
summarises its main properties.

Table 3.9 Main parameters of the International Linear Collider

Type	Electron–positron collider
Centre of mass energy	205 + 250 GeV
Luminosity	2×10^{34} cm^{-2} s^{-1}
Accelerating gradient	31.5 MeV/m
Cavities	Superconducting, standing wave
Frequency	1.3 GHz
Repetition rate	5 Hz
Pulse duration	1 ms
Beam current in pulse	9 mA
Energy of damping rings	5 GeV
Circumference of damping rings	6.7 km
Length of each linac line	11 km
Total site length	31 km
Total power consumption	230 MW

Table 3.10 Main parameters of CLIC (Compact Linear Collider)

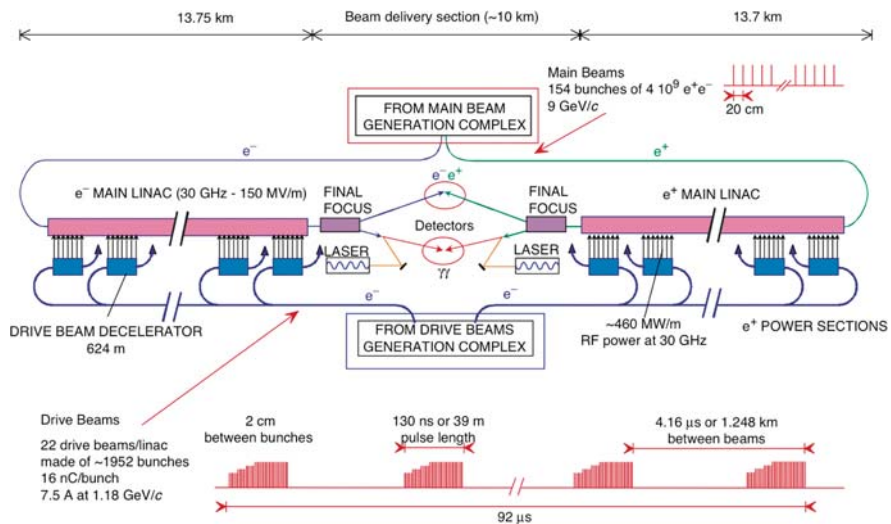| Type	Electron–positron collider
Centre of mass energy	1500 + 1500 GeV
Design Luminosity	80×10^{33} cm^{-2}s^{-1}
Linac repetition rate	100 Hz
Total 2 linac length	38 km
RF frequency	30 GHz
No. of particles/bunch	0.4×10^{10}
No. of bunches/pulse	154
Bunch train length	102 ns
Accelerating gradient	150 MeV/m

Fig. 3.34 Overall layout of the (Compact Linear Collider) CLIC. Figure from [17], copyright CERN

The size of the budgets involved makes it clear that this will, in the end, be a political decision. Many different options are possible. It is difficult to be certain what the outcome will be, but it remains very likely that, if a new very high-energy accelerator is built one day in next 30 years, it will be a linear electron–positron collider.

3.10 Exercises

1. Assume a linear accelerator as shown in Fig. 3.5 and an alternating voltage source of 10 MHz. Assume that we want to use it to accelerate electrons. After a few steps, the electrons will have a velocity close to the velocity of light. How long should each of the tubes be to accelerate each electron further?
2. Assume that you have a cyclotron with a magnet of 1.5 tesla field. The useful diameter of the magnet is 2 m. What is the maximum energy you can reach for protons with this machine?
3. Show that in a cyclotron the distance between the successive orbits becomes smaller as the energy of the particles becomes larger.
4. Show that the equation for the radius of curvature of the track of a charged particle in a magnetic field: $P = Ze\,B\,r$, can be rewritten as Eq. (3.1).
5. Assume that to drive a nuclear reactor one needs a beam of protons with an energy of 1 GeV with a beam current of 20 mA. Assume that the accelerator has an efficiency for converting electrical energy to beam energy of 33%. How much electrical power will this accelerator use?
6. What is the speed of the train that has the same kinetic energy as the energy stored in one of the proton beams of LHC. A typical train weighs 400 metric tons.
7. Assume that we accelerate protons and make them collide with protons at rest. What should be the energy of the proton beam to produce the same centre of mass energy as the LHC collider.
8. In the SPS proton synchrotron, the frequency of the RF cavities at the maximum energy of 450 GeV is 200.2 MHz. How much should the frequency be at the injection energy of 10 GeV?
9. Assume a synchrotron for electrons with a beam energy of 1 GeV. What is the power dissipated by synchrotron radiation? Assume that the bending magnets have a field of 2 tesla that the number of particles stored is 10^{12} and that 33% of the circumference is occupied by the bending magnets. The rest of the circumference has quadrupoles and straight sections. Neglect the power dissipated in the quadrupoles.

References

1. IRCP publication 103, Recommendations of the International Commission on Radiological Protection, Elsevier (2008).
2. E. Pochin, Nuclear radiation: risks and benefits, Clarendon Press, Oxford, (1983).

3. E.J.N. Wilson, An introduction to particle accelerators, Oxford University press (2001).
4. P.J. Bryant and K. Johnsen, The principles of circular accelerators and storage rings, Cambridge UniversityPress (1993).
5. D.A. Edwards and M.J. Syphers, An introduction to the physics of high energy accelerators, John Wiley& sons, Inc. (1993).
6. H. Wiedemann, Particle accelerator physics, Springer-Verlag, Berlin, (1993).
7. M. Reiser, Theory and design of charged particles beams, John Wiley& Sons (1994).
8. A. Chao and M. Tigner, Handbook of accelerator physics and engineering, World Scientific (1998).
9. K. Wille, The physics of particle accelerators: an introduction, Oxford University Press (2000).
10. S. Humphries, Jr, Principles of particle accelerators, John Wiley & Sons; (1986), Also downloadable from http://www.fieldp.com/cpa/cpafull.pdf
11. Panorama des accélérateurs et de leurs utilisations, Eric Baron, Ecole IN2P3 « De la Physique au Détecteur » , Bénodet 15–22 (novembre 2006).
12. H.A. Willax, Proposal for a 500 MeV Isochronous Cyclotron with Ring Magnet, Proc. Int. Conf on Cyclotrons, Geneva 1963, CERN Rep 63–19, pp. 386–397.
13. D. Proch, Superconducting cavities for accelerators, Rep. Prog. Phys. 61, 431–482 (1998).
14. E.B. Podgorsak, Technical Editor. Radiation oncology physics: handbook for teachers and students, International Atomic Energy Agency, Vienna (2005).
15. A.N. Didenko, A.A. Glazkov, A.D. Koljaskin, G.L. Horasanov, V.E. Kalantarov, Conception of secure atomic energy plant with subcritical reactor and 100 MeV proton accelerator, Paper presented at the accelerator conference, CERN (1996).
16. Berkely Lab Research News, Laser Wakefield Acceleration: Channeling the Best Beams Ever (September 29, 2004).
17. A 3 TeV e^+ e^- Linear Collider Based on Clic Technology, Proton Synchrotron Division, CERN 2000–008, (28 July 2000).
18. Extensive information of high-energy electron accelerators and the CLIC project can be found on http://clic-study.web.cern.ch/CLIC-Study/.

Chapter 4
Detectors Based on Ionisation in Gases

4.1 Introduction to Detectors for Subatomic Particles

Charged subatomic particles can be detected by their electromagnetic interactions with matter. Neutral subatomic particles can only be detected if they first undergo an interaction. The charged particles produced in the interaction reveal that a neutral particle was present.

Today there are three main methods for detecting charged subatomic particles that are important in nuclear science or in high-energy physics: detection based on gas ionisation, detection based on semiconductors and detection based on scintillation. All these methods are based on the detection of electron ionisation or electron excitation produced by the coulomb interaction of the charged particle with the medium.

Many years ago, there were several other useful detection methods, such as bubble chambers and cloud chambers, but these have all become obsolete because these methods did not allow advantage to be taken of the dramatic improvements in electronics.

One significant exception to this is the nuclear emulsion. A nuclear emulsion is essentially the same as a common black-and-white photographic emulsion. The very first observations of subatomic particles and ionising radiation were, in fact, made with photographic emulsions. In 1895, Wilhelm Conrad Roentgen discovered X-rays by observing that images could be obtained from the bones in a human body, by using what is now called a cathode ray tube and a photographic emulsion. One year later, Henri Becquerel observed that certain uranium salts emit a kind of penetrating radiation that can be detected with photographic emulsions. Ordinary photographic film consists of silver halide grains (mainly silver bromide) suspended in a gelatine matrix and supported with a backing of glass or cellulose acetate film. The action of ionising radiation in the emulsion is similar to that of visible light: some of the grains become sensitive through the excitation of electrons in the silver halide crystal. In the subsequent development process, the sensitive grains become metallic silver and are visible as black grains. The silver grains in emulsions can be made quite small, less than 1 μm, so that images of excellent resolution can be recorded. This excellent resolution is one of the main reasons why nuclear emulsions have not yet been abandoned as a detection technique.

Nuclear emulsions can be used for observing a radiation flux impinging on the emulsion, and the higher the radiation flux the darker the emulsion. This is the basis of the use of emulsions as film badge dosimeters. A long time ago, nuclear emulsions were also used to make X-ray images for medical applications, i.e. for clinical radiography. This method has long been abandoned because the probability of the X-ray interacting in the emulsion is small, typically only a few percent, resulting in the need of a large dose to obtain good images. For this application, emulsions have been replaced by methods that will briefly be described in Chap. 6. However, emulsions are still used for making X-ray images of objects, for example in non-destructive material testing, where dose is usually not important and advantage can be taken of the excellent spatial resolution of emulsions.

A limitation of the use of emulsions in radiography or in dosimetry is the limited dynamic range, i.e. the range in dose the emulsion can record. At higher or lower doses, the only information the emulsion provides is that the dose was above or below the range of sensitivity. Typically, emulsions have a dynamic range of ≈ 100.

Another application of emulsions is as a method to make trajectories of charged particles visible. If a charged particle travels in an emulsion, the silver halide crystals along its trajectory will become sensitive and after development the track of the particle becomes visible under a microscope. For a long time this was a very important method for the study of nuclear interactions. It is still used today, but only in cases where one wants to observe very short tracks that are difficult to observe by other methods. Figure 4.1 illustrates how emulsions can be used to observe tracks of subatomic particles.

If a charged particle penetrates in a medium, many atoms along the trajectory of the particle are ionised. If the medium is a gas, an electric field of the order of a few 100 V/cm will be sufficient to collect the charges and, in this way, detect the presence of charged particles. In a solid, it is usually not possible to collect
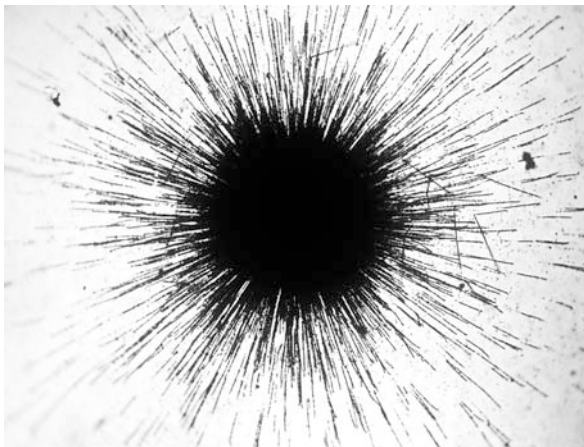

Fig. 4.1 Alpha particles shoot out from a speck of uranium salt on the surface of a nuclear emulsion. The area shown corresponds approximately to 0.3×0.2 mm^2. Figure from [1]./Science Photo Library
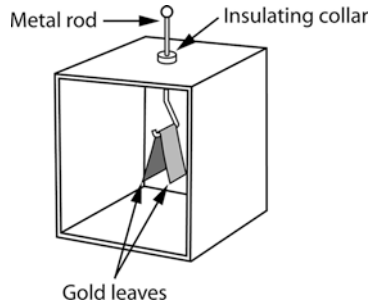
Fig. 4.2 The gold leaf electroscope was one of the earliest instruments used in studying subatomic particles. It consists basically of a box with a window, or of a glass bottle. A metal rod, which passes through an insulating collar in the top of the box, has two thin sheets of gold foil attached. When the rod is electrically charged the two gold leaves acquire the same charge and repel each other. If the air in the box is ionised by radiation, the charge on the gold leaves leaks away to the walls of the box and the leaves collapse together

the charges. This is because in a solid there are always a very large number of traps that will capture the charges and prevent the collection. Only for a small number of carefully engineered materials, such as silicon and germanium, is it possible for charges to move freely over distances of several millimetres with negligible loss. Semiconductor detectors are discussed in Chap. 5. In liquefied noble gases, it is also possible to collect the ionisation charges. This is because the noble gases have a very low affinity for electrons and the electrons move freely under the influence of an electric field if the material is extremely pure. This has a number of important applications in high-energy physics and this is briefly discussed in Sect. 6.6.

Observation of ionisation in gases is also one of the oldest ways of observing ionising radiation. In the very beginning of nuclear science, this was done with the help of an instrument called an electroscope (see Fig. 4.2).

Today the use of ionisation in gases is one of the common methods used to detect the presence of high-energy subatomic particles. This is the subject of the rest of this chapter.

4.2 Ionisation and Charge Transport in Gases

Before we proceed with a discussion of detectors based on gas ionisation, it is useful to briefly review the main physical phenomena associated with the creation and the transport of charges in a gas. See Ref. [7] for a more in depth review of the subject.

When a charged particle travels in a gaseous medium, the coulomb interaction between this charged particle and the gas atoms will cause excitation or ionisation of the gas molecules. If the gas molecule is ionised, a free electron and a positive ion are produced. The ionisation energy for most gases is between 10 and 20 eV (see

Table 4.1 Energy loss characteristics in some commonly used gases. Energy loss, the number of electron–ion pairs and the number of primary electrons is for charged particles at minimum ionisation

Gas	Ionisation potential [eV]	Mean energy /electron–ion pair 'W' [eV]	Energy loss [keV/cm]	Number of electron–ion pairs [cm^{-1}]	Number of primary electrons [cm^{-1}]
Ar	15.7	25.0	2.53	106	25
Xe	12.1	22	6.87	312	41
He	24.5	41.6	0.345	8.3	5
H_2	15.6	36.4	0.32	8.8	5.2
N_2	15.5	34.8	1.96	56.3	10
Air		33.8	2.02	59.8	
O_2	12.5	30.2	2.26	74.8	22
CH_4	12.6	30	1.61	54	37
C_2H_6	11.5	26	2.91	112	48
Isobutane/i-C_4H_{10}	10.6	26	5.67	220	90
CO_2	13.8	34	3.35	100	35

Table 4.1), but the average energy needed to produce an electron–ion pair is typically about two times larger. This is because the energy transferred to the electron is usually larger than what is needed to ionise the atom, and part of this kinetic energy will dissipate as heat. If this electron has acquired sufficient energy, it will itself cause further ionisation of gas molecules. Each primary electron will typically give rise to 3 electron–ion pairs.

If a charged track travels a fixed distance in a gas, the amount of energy deposited fluctuates from one event to the next. This is partly due to the fluctuation on the number of primary electrons ejected from the gas molecules and partly due to the fluctuation on the amount of energy received by each electron. This will give rise to a distribution of the number of charges produced that is wider than what can be expected for a Poisson distribution.

In the presence of an electric field, the electrons and ions created by the radiation are accelerated towards the anode and cathode, respectively. This acceleration is interrupted by collisions with the gas molecules. In these collisions, the electrons and ions will completely change direction. The collision resets the average drift velocity of the electron or ion in the direction of the field back to zero. After the collision, the charges will again be accelerated in the direction of the electric field. At the microscopic scale the movement is quite chaotic, but at the macroscopic level it appears as if the charges drift with constant velocity in the direction of the electric field and as if a diffusion component is superimposed on the otherwise uniform drift velocity. The cross section for the collision of ions with gas molecules is determined by the dimensions of the molecules involved and these cross sections are therefore of the order of a few 10^{-15} cm^2. The corresponding mean free path of the atoms or ions is about 100 nm. Moreover, these cross sections are independent of the kinetic energy of the ion. For the electric fields of

interest the drift velocity is always much smaller than the thermal velocity of the particle. We can therefore derive an approximate expression for this drift velocity as follows.

The average thermal kinetic energy of the molecules, or ions, in a gas is $3kT/2$. The average thermal velocity v_t of an ion is therefore given by

$$E_{\text{kinetic}} = \tfrac{1}{2}M\,v_t^2 = (3/2)\,kT$$

$$v_t = \sqrt{\frac{3kT}{M}}$$

In this equation k is the Boltzmann constant, T the absolute temperature and M the mass of the molecule or ion. Applying this equation to nitrogen gas molecules, we find that the thermal velocity is $\approx$500 m/s. If λ represents the mean free path of the ion, the average time between two collisions is given by

$$\Delta t = \frac{\lambda}{v_t}$$

If there is an electric field, the ion will be accelerated in the direction of the electric field. The acceleration is given by

$$a = \frac{eE}{M}$$

The average velocity of the ion in the direction of the field is therefore given by

$$v_d = \frac{a}{2}\Delta t = \frac{eE}{2\,M}\frac{\lambda}{v_t} = \frac{e\lambda}{\sqrt{12kT\,M}}E \tag{4.1}$$

We see that the drift velocity is proportional to the electric field and it is therefore useful to introduce the mobility 'μ' defined by

$$v(E) = \mu E$$

Applying Eq. (4.1), we find that the drift velocity of charged nitrogen molecules in an electric field of 1000 V/cm is 34 m/s if we assume a mean free path of 100 nm. This is indeed much smaller than the thermal velocity of the molecules. The ion mobilities of a few ions are listed in Table 4.2.

For electrons, the drift velocity is not proportional to the electric field and the mobility therefore becomes a function of the electric field. This difference in behaviour is due to the fact that the cross section for collisions of the electrons on gas molecules is strongly dependent on the kinetic energy of the electrons, as is illustrated in Fig. 4.3. Moreover, electron–molecule cross sections are, on average, much smaller than ion–molecule cross sections. Figure 4.4 shows the drift velocity of electrons in an electric field, for a few gases of interest. Electron drift velocities

Table 4.2 Ion mobility in a few gases. Data taken from Refs. [6] in Chap. 1 and [2]

Gas	Ion	Mobility [cm^2 V^{-1} s^{-1}]
Ar	Ar^+	1.54
He	He^+	10.4
CO_2	CO_2^+	1.09
Ar	CH_4^+	1.87
Ar	$C_2H_6^+$	2.06
Ar	$iC_4H_{10}^+$	2.15
Ar	CO_2^+	1.72

are typically of the order of a few cm/μs for an electric field of 1 keV/cm. This is about 10,000 times faster than the typical ion drift velocities.

While drifting in an electric field, the electrons and ions can be involved in a number of interactions with the gas molecules of the medium:

– Several gases, e.g. oxygen, are strongly electronegative. This means that electrons will become attached to oxygen atoms and form a negative ion. After capture, the negative oxygen ion displays similar drift behaviour as positive ions.
– During a collision between different molecules, one charged and one neutral, the molecule with the largest ionisation energy will tend to capture the electron from another molecule, where the electron is less tightly bound. As a result the charge can be transferred from one molecule to another.
– Electrons and positive ions, or positive and negative ions can also recombine their charges to produce neutral atoms.
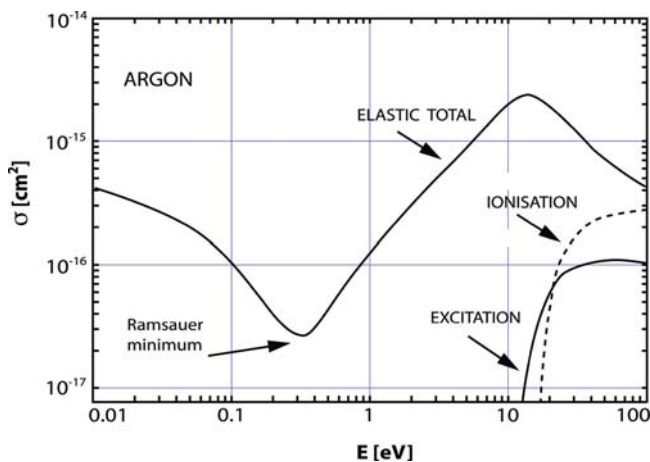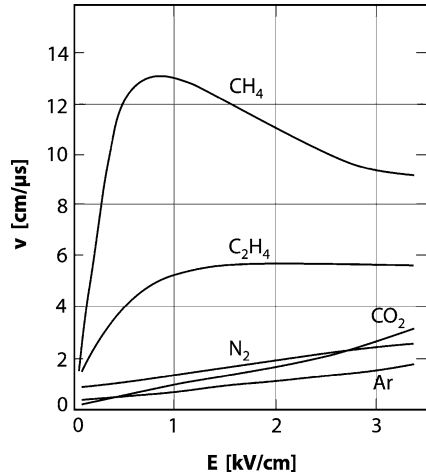

Fig. 4.3 Electron–argon cross section in argon gas as a function of the kinetic energy of the electron. The data for this figure were obtained from [2]

Fig. 4.4 Drift velocity of electrons as a function of the electric field in several gases. Figure from [3], copyright CERN

Diffusion causes the drifting charges to deviate from the direction given by the electric field. If the diffusion is perpendicular to the motion of the particle, this is called lateral diffusion. However, the diffusion will also cause some particles to move slightly faster or slower than others. This is called longitudinal diffusion. If a group of particles is created at the same moment in one point in the gas, the electric field will cause these particles to drift in the direction of the field. After some time these particles will no longer be at one point but will be spread over a certain volume. Since this dispersion is due to a large number of uncorrelated random collisions along the track, the positions of the particle will have a Gaussian distribution, and the dispersion is proportional to the square root of the drift distance. It can be shown [7] that this dispersion is related to the diffusion constant D by the relation

$$\sigma_x = \sqrt{(2D/v_d)l}$$

In this equation σ_x represents the dispersion of the projected distance on some direction x, v_d represents the drift velocity, and l the path length of the particle.

If the electric field is increased, at some point the moving charges in the gas can acquire sufficient energy to ionise other atoms. In this way, the number of charges will increase. The mean free path of the electrons is much larger than the mean free path of ions; therefore, electrons can acquire a much larger energy than the ions, and electrons will start multiplying at a lower value of the electric field than the ions. As is discussed in Sect. 4.4, this charge multiplication effect is exploited to amplify the very small signals that are produced by ionisation in gases.

4.3 Ionisation Chambers

The amount of ionisation in a gas volume is a measure of the amount of radiation present. This provides a commonly used method to measure gamma ray exposure. The SI unit of X-ray or gamma ray exposure is defined as the amount of radiation
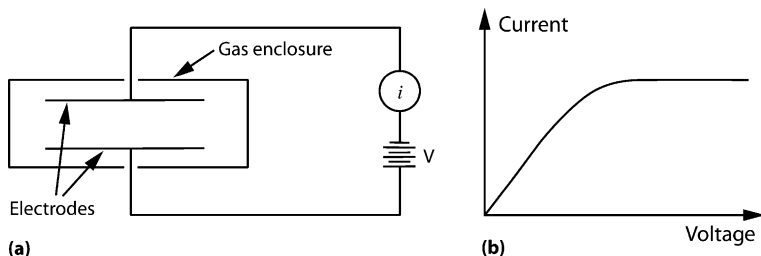
Fig. 4.5 Schematic representation of an ionisation chamber. (**b**) Current–voltage diagram in an ionisation chamber. A voltage of a few 100 V/cm is usually sufficient to collect all the charges

producing one coulomb of charge per kilogram of air at standard temperature and pressure.

An ionisation chamber to measure gamma ray exposure is a very simple device and is illustrated schematically in Fig. 4.5. A certain volume of air is enclosed between two electrodes and a voltage difference of several 100 V is applied on these electrodes. If the voltage is too low, the electrostatic force on the charges is not sufficient to overcome the random thermal movement of the gas molecules, and we have incomplete collection of the ionisation in the gas. In addition, the charges have to be collected fast enough so that recombination of the charges remains negligible. A moderate field of the order of a few 100 V/cm is usually sufficient to obtain full charge collection. Increasing the voltage over the electrodes further will not lead to an increase in the collected current. If the voltage over the electrodes is increased to a value much larger than 1000 V/cm, at some point amplification of the charges will occur and the current will again increase. In that case, the amount of charge collected is no longer equal to the ionisation produced by the gamma rays. The voltage at which charge multiplication will occur depends strongly on the geometry of the ionisation chamber. In practice, it is fairly easy to find a working voltage where the charges are collected and no charge amplification occurs.

Implied in the definition of gamma ray exposure is the assumption that the sample air volume is taken in a sufficiently large volume of air. Sufficiently large here means large compared to the typical trajectories of the electrons produced by the gamma rays. These trajectories can be up to several metres long! Ideally, one should therefore make the gas enclosure and the electrodes sufficiently thin such as to produce a negligible perturbation of the radiation present. In a practical measurement set-up, the ionisation chamber and the electrodes are unavoidably made of some suitable solid material. This material will perturb the measurement because, on the one hand, it shields the gas in the ionisation chamber from electrons produced outside the sample air volume, and on the other hand, it adds electrons produced by interactions of the gamma rays in the walls of the ionisation chamber. This perturbation will be minimised if the walls are made of 'air equivalent material', i.e. of material with approximately the same atomic charge as nitrogen and oxygen. The electrodes should therefore preferably be made of aluminium and the detector walls of plastic.
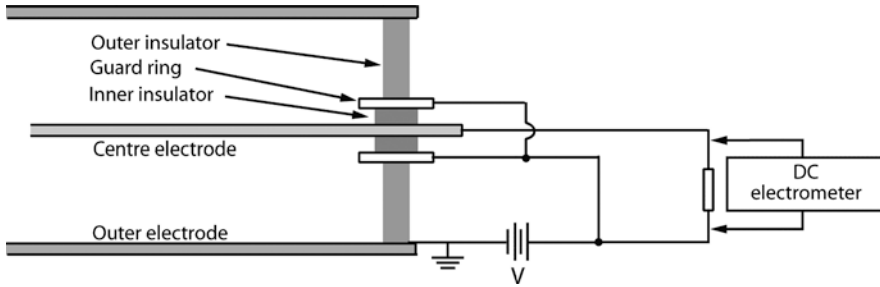
Fig. 4.6 Cross-sectional view of an ionisation chamber with guard rings. The leakage current between the guard ring and the ground is not included in the measured ionisation current. The very small ionisation currents can be measured with a large series resistor and an electrometer to record the voltage over the resistor

The currents that need to be measured with ionisation chambers are usually extremely small. If we assume that there are 1000 electron tracks per second in the gas volume and that each track is 2 cm long. The corresponding current in the ionisation chamber will be 18×10^{-15} A. This is an extremely small current, and without adequate precautions, the leakage current in the system will be much larger that the ionisation currents one wants to measure. The solution that is universally adopted is the use of guard rings. The principle of guard rings is schematically illustrated in Fig. 4.6. The measurement electrode is completely surrounded by another electrode called the guard ring. This guard ring is at the same voltage as the measurement electrode. Because there is no voltage difference between the high-voltage electrode and the guard ring there is no leakage current.

The method used for the measurement of the small ionisation currents is illustrated in Fig. 4.6. The ionisation current causes a voltage difference over a resistor, and this voltage difference is measured with a sensitive voltmeter (DC electrometer). The values of the resistors used are typically between 10^{+9} Ω and 10^{+12} Ω.

Battery-operated portable ionisation chambers are commonly used as radiation-monitoring instruments to measure gamma ray exposure. Figure 4.14 shows some commercially available systems. Because the walls of these detectors can never be made exactly equivalent to air, each instrument has to be calibrated for its sensitivity as a function of the gamma ray energy.

When an ionisation chamber is used to measure gamma ray exposure, the total charge pulse corresponding to one electron trajectory of a few centimetres long corresponds to only a few 100 electron charges. However, an ionisation chamber can also be used to measure the concentration of air-borne alpha emitters. An alpha particle of 3.38 MeV will produce about 100,000 electron–ion pairs in air. This is still a very small electrical pulse, but large enough to be observed with suitable low noise electronics.

It is always better to count individual pulses rather than measuring just a total current. The example at hand illustrates this. If one were to use the current to measure the number of alpha emitters present in a gas, it would be impossible to distinguish the current caused by the alpha particles from the current caused by electrons

Fig. 4.7 Under the influence
of the electric field an
electron and a positive ion
drift towards the electrodes
delimiting the volume of an
ionisation chamber. These
two electrodes form a parallel
plate capacitor
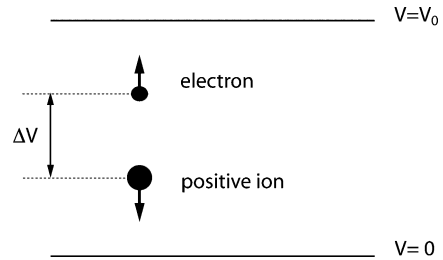
produced by gamma rays. If the individual pulses are observed, it is possible to distinguish the two types of events because of the very large difference in amplitude between the corresponding pulses. Another advantage of counting pulses is that this method is much less sensitive to changes in the collection efficiency of the charges, either variations over the volume of the detector or variations with time. On the other hand, counting pulses needs a more elaborated electronics and becomes impossible if the rate of events exceeds a few 10 MHz.

To understand the time development of the electrical pulse caused by ionisation in a gas, let us consider a detector consisting of two parallel plates enclosing a gas volume. These two plates form a capacitor with a capacitance denoted by C. A potential difference V_0 between the plates causes a constant and parallel electric field between them (see Fig. 4.7). Let us assume that one electron–ion pair forms in the gas. Under the influence of the electric field the electron and the positive ion will drift towards the anode and cathode, respectively. After a certain time the electron and the positive ion will have reached their electrodes, and if the plates are not permanently connected to an external voltage source, the voltage difference over plates changes by

$$dV_0 = \frac{-e}{C} \tag{4.2}$$

In this equation '$-e$' represents the charge of the electron.

The drift of the electron and the ion in the gas to their respective electrodes is equivalent to a current source connected between the two plates and injecting a charge '$-e$' and 'e' in the anode and cathode, respectively. However, it is not immediately obvious at what moment exactly this charge appears on the electrodes. Energy conservation will help us in answering this question.

In the interest of clarity let us only consider the motion of the electron. If the electron has only travelled a fraction of the distance between the plates, it will have crossed a potential difference ΔV and the energy it received in doing so, $e\Delta V$, has been dissipated as heat in the gas. Because of energy conservation, this energy must have been supplied by the energy stored in the capacitor $1/2CV_0^2$. The voltage over the capacitor has therefore changed by dV_0 and we have the relation:

$$d(\tfrac{1}{2}CV_0^2) = V_0 C dV_0 = -e\Delta V$$
$$dV_0 = -\frac{e\Delta V}{CV_0}$$

The motion of the charge $-e$ has injected an apparent charge in the electrode given by

$$\Delta Q = -\frac{e\Delta V}{V_0}$$

The motion of a particle with charge '$-e$' in the detector is therefore equivalent to a current source injecting a current in the anode given by the Eq. (4.3):

$$i = \frac{dQ}{dt} = \frac{-e}{V_0}\frac{dV}{dt} = \frac{-e}{V_0}\frac{dV}{ds}\frac{ds}{dt} \tag{4.3}$$

In this equation, s represents the trajectory of the electron. Notice that, if a current 'i' is injected on the anode, necessarily a current '$-i$' is injected on the cathode. One is therefore free to choose either of the electrodes to read the signal. In fact, we can use both signals and this is often used in detectors.

We have seen in the above example that the signal in a detector is caused by the motion of the charges between the electrodes. This is completely general and applies to all detector types and all detector geometries. In the simple case, where the detector only has two electrodes, the signal can be found using energy conservation as we have done here. Later in this book we will see many examples of detectors where there are more than two electrodes. In that case, we need to use the Shockley–Ramo theorem to find the signal induced in one particular electrode. Assume that there are many electrodes in the detector and we want to know the current injected in one of these electrodes by the motion of a charge somewhere in the space between the electrodes. The Shockley–Ramo theorem [4, 5] states that the current injected in this electrode is also given by Eq. (4.3), but using the weighting potential field for this electrode in calculating dV/ds, not the true potential field in the detector. The weighting potential field is the potential field obtained when all other electrodes are set at zero potential and only the electrode under consideration is at the potential V_0. If there are only two electrodes, the weighting potential field is the same as the true potential field in the detector. If there are more than two electrodes, the weighting potential field is, in general, different from the true potential field in the detector. In the calculation of the current, the trajectory 's' used in Eq. (4.3) must of course be the true trajectory caused by the real potentials used when operating the detector.

The time development of the electrical pulse caused by one electron–ion pair produced in the electric field of an ionisation chamber with parallel plate electrodes is shown in Fig. 4.8. The electron travels much faster than the positive ion and therefore reaches the anode in a short time. The movement of the positive ion is much slower and the signal induced by the positive ion lasts much longer. The current pulse therefore also consists of two parts, one fast part corresponding to the electron movement and one slow part corresponding to the movement of the positive ions.
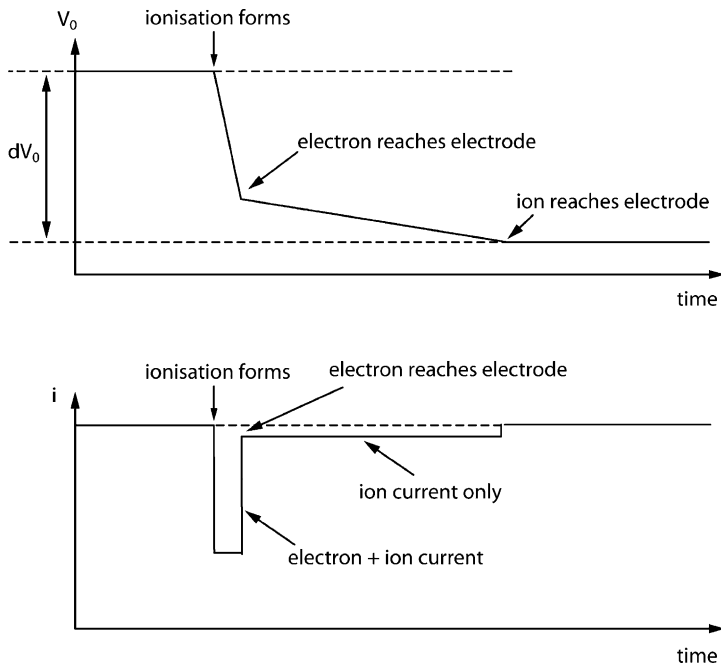
Fig. 4.8 Time development of the signal caused by the motion of the electron and the ion in a parallel plate ionisation chamber. In the interest of clarity, the figure is drawn as if the electron drift velocity were only ten times faster than the ion drift velocity. In reality, the difference in velocity is much larger. The top figure shows the change of voltage over the plates caused by the motion of the charges if the plates are disconnected from the power supply. The bottom figure shows the apparent current caused by this motion of the charges

The voltage difference caused by the creation of one electron–ion pair in a detector with a capacitance of 10 pF (a typical value) is 1.6×10^{-8} V. This is a very small signal and is not observable. An alpha particle of 3 MeV will give a signal of 15×10^{-4} V. This is still very small but is observable with suitable low noise electronics.

If a high-energy electron causes the ionisation in the detector, the signal will typically only be of the order of 100 electron charges and such a signal is unobservable using purely electronic methods. One therefore needs to use gas amplification as described in the next section.

4.4 Counters with Gas Amplification

The electrical signals produced by charged particles in a gas are often too small to be observable. In 1928, Geiger and Müller invented a detector that takes advantage of the phenomenon of charge amplification in gas to obtain much larger electrical signals. Counters based on the same principle are still in common use today. These

Fig. 4.9 A Geiger tube or a proportional tube is a conductive cylinder with a thin wire in its centre. The dotted line shows an imaginary cylinder used in calculating the electric field in the tube
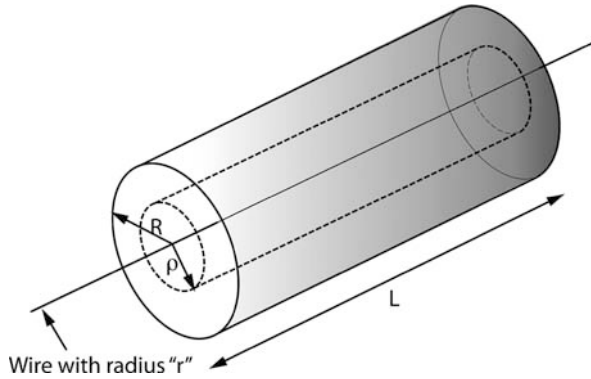
counters are now usually of a type called 'proportional counters'; the difference between a proportional counter and a Geiger counter will be explained later.

A proportional counter or a Geiger counter consists of a metal tube of radius R with a thin metallic wire of radius 'r' in its centre (see Fig. 4.9). The central wire is brought to a large positive voltage, i.e. it is used as an anode, while the tube itself is at ground potential and is used as a cathode.

The electric field in the counter can be found by applying the Maxwell equations in the integral form over an imaginary cylindrical volume of radius ρ shown as a dotted line in Fig. 4.9.

$$\int_{\text{surface}} \vec{D}.\vec{ds} = \int_{\text{volume}} q.dv$$

We assume the cylinder to be a finite section taken out of a very long cylinder. The electric field over the flanges at both ends therefore points radially outwards from the central wire and does not contribute to the surface integral. We therefore have

$$\varepsilon E(\rho)2\pi\rho L = Q$$
$$E(\rho) = \frac{Q}{2\pi\varepsilon L}\frac{1}{\rho} \tag{4.4}$$

Equation (4.4) gives the expression for the electric field as a function of the electric charge present on the central anode wire. In practice, this wire is charged by connecting it to a voltage power supply. We therefore need to calculate this charge on the wire as a function of the externally applied voltage V_0. This is the same as calculating the capacitance of the system. The electric field in the cylinder is given by

$$E = -\frac{dV}{d\rho}$$

Integrating this equation along a line pointing radially outwards from 'r' to R, we obtain

$$V(R) - V(r) = -\int_r^R E(\rho)d\rho$$

The outer cylinder is at ground potential, therefore $V(R) = 0$ and the central wire is connected to the external powers supply, hence $V(r) = V_0$. We hence have

$$V_0 = \int_r^R E(\rho)d\rho = \frac{Q}{2\pi\varepsilon L}\ln\left(\frac{R}{r}\right)$$

By definition, the capacitance of a system is given by Q/V. Therefore, the capacitance C of the proportional tube is given by

$$C = \frac{Q}{V_0} = \frac{2\pi\varepsilon L}{\ln(R/r)} \tag{4.5}$$

Eliminating Q from Eqs. (4.4) and (4.5), we readily find the expression for the electric field in the cylinder:

$$E(\rho) = \frac{V_0}{\rho\ln(R/r)}$$

We see that the electric field in the tube varies inversely proportional to the distance to the centre of the tube. For a central wire with a radius of 10 μm and a power supply of only 1000 V, the electric field close to the central wire is 1.5×10^7 V/m. This is sufficient to cause electron multiplication.

Let us assume that a charged particle produces a number of electron–ion pairs in the gas inside the tube. Under the influence of the electric field, the electrons and the ions move towards the anode and cathode, respectively. As an electron comes closer to the anode, the electric field increases, and in the last few 10 μm, very close to the anode wire, the field is sufficiently strong to give rise to charge multiplication. This means that the electrons acquire enough energy that they can ionise other atoms and create additional free electrons. In this avalanche, the number of electrons increases exponentially. Because of charge diffusion, the electron avalanche spreads more or less evenly around the anode wire.

A large number of electrons and an equal number of positive ions are formed. These electrons move much faster than the positive ions; therefore, the electrons reach the anode in about 1 ns and leave the positive ions behind. The ions move slowly towards the cathode. In doing so, the ions do not cause any charge multiplication, because the onset of charge multiplication for ions is at much larger electric fields than for electrons. The ions need several 100 μs to reach the cathode; the exact value depends on the dimensions of the system and on the gas filling of the tube. This is illustrated in Fig. 4.10.
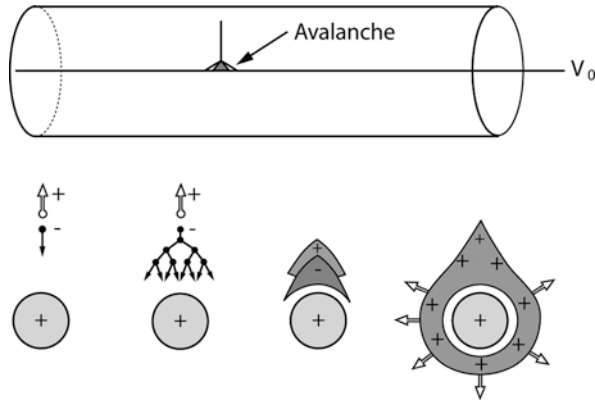
Fig. 4.10 Development of an electron avalanche in a proportional tube. An electron formed some-where in the gas volume drifts towards the positive anode wire. In the intense electric field close to the anode, the electrons multiply. Because of diffusion, the avalanche more or less evenly sur-rounds the anode wire. The electrons created in the avalanche reach the anode in less than 1 ns. The positive ions drift slowly outwards towards the cathode

Pulse shape in counters with gas amplification. Let us first imagine that the charged particle has created only one electron–ion pair in the gas. The signal caused by this single electron–ion pair is negligible compared to the signal produced by the many electrons and ions produced in the avalanche. The electrons in the avalanche reach the anode wire in less than 1 ns, but these electrons only travel a few $10 \, \mu m$, and therefore the fraction of the total potential difference travelled by the electrons is very small. As a result, the contribution of the electron motion to the total signal is, at most, a few percent. Most of the potential difference is travelled by the ions, and therefore the motion of the ions causes the largest part of the electrical pulse.

Let us consider a proportional tube with a signal readout as shown in Fig. 4.11(a). We assume that the current is measured with an ideal current meter with zero impedance. We, furthermore, suppose leakage currents are negligible. As long as no ionisation is produced in the gas, there is no current. If one electron–ion pair forms, the electron will quickly drift to the anode wire. Close to the anode wire this electron will give rise to the formation of an electrons' avalanche and the number of charges is multiplied by a large factor. The motion of the positive ions determines the evolution of the signal. In the beginning, the ions are in a very large electric field, and both the velocity of these ions and the potential gradient are large. Therefore, in the beginning, the induced signal current is large. As the ions move away from the anode, the change in potential slows down and the current decreases. Eventually all the ions reach the cathode and the current stops.

A more realistic electrical readout scheme is shown in Fig. 4.11(b). The capaci-tance C_a, shown in dotted lines, represents the capacitance of the anode wire itself. The external resistor R_e and the external capacitance C_e represent external compo-nents. The triangle to the right of Fig. 4.11(b) represents an amplifier. This amplifier
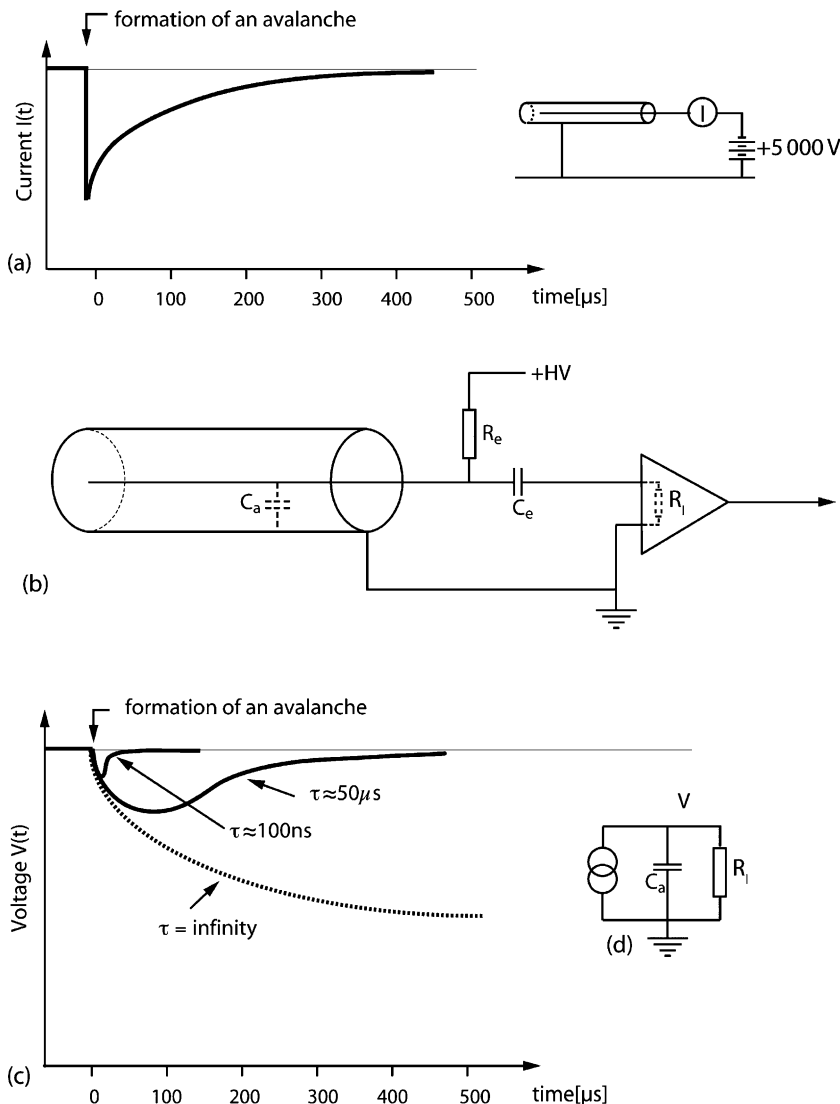
Fig. 4.11 (**a**) Idealised electrical circuit for the readout of a proportional tube. The anode current I is read with an ideal current meter with zero internal resistance. (**b**) Realistic electrical circuit for the readout of a proportional tube. The capacitance of the anode wire C_a and the input impedance R_I of the amplifier are drawn as dotted lines to show that these are not actual components in the circuit. These are only drawn as a reminder that the anode wire has a certain capacitance and that the amplifier has a certain input impedance. (**c**) Pulse shape of a proportional tube for different values of the shaping time $\tau = C_a R_I$. (**d**) Simplified equivalent circuit allowing to derive the shape of the output pulse

has an input impedance represented by the resistor R_I, also shown in dotted lines in the figure.

The value of the external resistor, R_e, should satisfy the condition $R_e \gg R_I$. This condition guarantees that most of the signal goes to the amplifier instead of going to the voltage power supply. Indeed, a voltage power supply has negligible internal impedance. The external resistor R_e decouples the anode wire from the power supply for short pulses. This means that a short pulse will develop as if the anode were not connected to the external power supply. After the arrival of a pulse, the anode potential will slowly return to its nominal value V_0 with a time constant given by $R_e\, C_a$. Obviously, this time constant should be large compared to the duration of the pulse. On the other hand, the resistor cannot be too large because at high rates, the voltage of the anode wire will not be restored to its normal value before the next pulse arrives, resulting in a drop of the average level of the anode voltage. This will introduce a rate dependence of the gain of the proportional tube. In practice, the value of the external resistor R_e will be a compromise between these conflicting requirements. The external capacitor C_e is necessary to isolate the amplifier from the high voltage of the anode wire. The value of this capacitor is not critical; the only condition is that $C_e \gg C_a$. This will ensure that the charge is effectively transferred to the amplifier instead of staying on the capacitor formed by the anode wire.

Finding the exact output signal that will be produced by the circuit shown in Fig. 4.11(b) is rather involved and the tools allowing to do so will only be introduced in Chap. 8. However, if the capacitance C_e and the resistor R_e are sufficiently large, to a good approximation these components can be ignored and the circuit becomes equivalent to the much simpler circuit shown in Fig. 4.11(d). The current source in Fig. 4.11(d) delivers a current equal to the current signal shown in Fig. 4.11(a). The shape of the output pulse is completely determined by the shape of the current pulse and the value of the time constant $\tau = C_a R_I$. Figure 4.11(c) shows the influence of the value of this time constant on the output pulse of the proportional tube. If the time constant τ is much longer than the duration of the current pulse, the total current pulse is integrated and the signal amplitude is large. After the pulse, the voltage returns to zero with a time constant τ. The result is a pulse that is extremely long, leading to a severe limitation of the maximum pulse rate the detector can handle. If this time constant τ is made shorter, this will lead to pulses that are much shorter, but are also much smaller. If the time constant τ is made short compared to the physical formation time of the pulse, the output is directly proportional to the current. Depending on the application, different values for this time constant will be used. One often chooses a time constant τ of the order of 100 ns.

When using a time constant much shorter than this, one will see the individual avalanches caused by the individual electrons arriving at the anode wire. Indeed, these individual electrons will not arrive at the same time on the anode, giving rise to a pulse as shown in Fig. 4.12.

We have seen that the motion of the relatively slow ions causes the pulse in a proportional tube. As a result the rise time of the pulse is not very fast and the time resolution that can be achieved is not very good. A time resolution of a few 10 ns (r.m.s.) is typical.
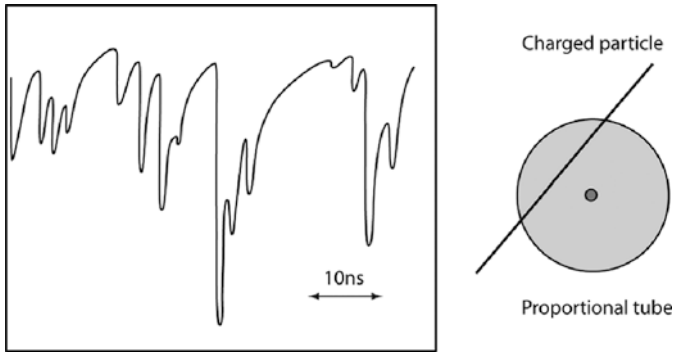
Fig. 4.12 Signal of a proportional tube as it would be seen if a very short shaping time were used. In this case, the individual pulses caused by each individual electron are visible

Pulse amplitude and amplitude fluctuations. Charge multiplication occurs if the electron acquires sufficient energy between two collisions with gas molecules to ionise one of these molecules. The probability that an electron creates an additional electron in an infinitesimal path length dx is αdx and the quantity α is called the 'first Townsend coefficient'. If $n(x)$ represents the number of charges as a function of the distance travelled by an electron in a constant electric field, we have

$$\frac{dn(x)}{dx} = n(x)\, \alpha$$

The number of charges increases exponentially

$$n(x) = n_0\, e^{\alpha x}$$

More generally, if the particle moves in a non-uniform electric field, the first Townsend coefficient becomes a function of x and the expression for the gas gain generalises to

$$n(x) = n_0\, e^{\int \alpha(x)dx}$$

This equation is only valid for moderate values of the gas gain. If the gas gain becomes too large, the space charge represented by the cloud of charges modifies the electric field and therefore the value of the first Townsend coefficient. If the gas gain exceeds the value of about 10^8, the simple mechanism described breaks down and the charge multiplication ends in a discharge. This is called the Raether limit.

Until now we have only considered the average gain in an avalanche. The charge multiplication is a stochastic process, and obviously all primary electrons will not be multiplied by exactly the same gas gain factor. It can be shown that, if all electron multiplication is only dependent on the local electric field strength in the absence of an avalanche, and for large values of the multiplication, the number of electrons in

an avalanche produced by one primary electron has an exponential distribution.

$$P(n) = \frac{1}{\langle n \rangle} e^{-n/\langle n \rangle}$$

This expression is a fair approximation of the amplitude distribution for moderate values of the gas gain. However, the primary signal in a gas amplification detector consists of many electrons, and the fluctuations on the pulse height in the total pulse is often dominated by the fluctuation on this primary number of electrons.

Gas mixtures: The gas mixture used in a proportional tube is very important. The gas should not contain any electronegative component. An electronegative molecule tends to form negative ions by capturing the electrons. The result is that the positive and the negative charge carriers are both ions and gas multiplication will only begin at much larger voltages. Moreover, as soon as there is charge amplification, both the positive and the negative ions give rise to multiplication resulting in an avalanche that never stops growing, ending in a discharge. Since oxygen is electronegative, air is not a good working gas for proportional tubes.

An obvious choice for the gas filling of a proportional tube is a noble gas. A noble gas certainly is not electronegative; moreover noble gases can easily be purified avoiding impurities that give rise to complications. Finally, a noble gas has the advantage that the collisions of the electrons with the gas atoms are elastic below the ionisation threshold. Since noble gas molecules are single atoms, there are no rotation or vibration states that can absorb the electron energy during the collisions. As a result, avalanche multiplication occurs at a lower voltage in a noble gas than in other gases. Of all the noble gases, argon is the least expensive, therefore nearly all proportional tubes use a gas filling based on argon. However, with pure argon the gain of a proportional tube is limited to a few times 100. This is due to the following mechanisms. The ionisation potential of argon is 11.6 eV, while the ionisation potential of all metals is less than 11.6 eV. The ionisation potential of copper, for example, is 7.7 eV. Therefore, electron–ion recombination, or exited argon atoms, will give rise to VUV photons that will be able to extract electrons from the cathode. If the average number of such secondary electrons is larger than one, each avalanche will, on average, give rise to more than one new avalanche. The number of avalanches will grow exponentially until the tube is filled throughout with avalanches, and the voltage drops to a very low value due to the large current drawn by all these avalanches.

Another reason why pure argon is not suitable as a fill gas is the fact that the argon ions arriving at the cathode will form neutral argon atoms by extracting an electron from this cathode and can dissipate the energy liberated in doing so by extracting an additional electron from the cathode. Again, if for one avalanche the average number of electrons extracted from the cathode is larger than one, the number of avalanches keeps increasing until breakdown occurs.

To prevent these phenomena from occurring, a small amount (typically 10%) of a polyatomic gas is added. This is called a quenching gas and several polyatomic gases can be used for this purpose. Isobutane (C_4H_{10}) or methane (CH_4)

is often used as quenching gas. Since these polyatomic gases have many rotation and vibration degrees of freedom, they will readily absorb these UV photons without being ionised. Moreover, in the collisions between the argon ions and quenching gas molecules, the charge will be transferred from the argon atom to the quenching gas molecule because the binding energy of the electron is lower in the quenching gas. When the charged quenching gas molecule arrives at the cathode, the likelihood of extracting an electron is much smaller, because this molecule has many other ways of dissipating the extra energy. In a gas mixture consisting of argon and ≈10% of a suitable quenching gas, stable operation with electron multiplication of 10^6 is possible.

The choice of gas mixture is very important if the detector has to work at a high count rate for a long time. The avalanche produces a large number of excited molecules and in this way initiates complex chemical reactions. In particular, the quenching gas, usually a gas containing carbon, participates in the formation of more complex molecules. These complex molecules are deposited on the electrodes in the proportional tube and, in particular, on the anode. This forms a thin insulating layer on the surface of anode wires, and this prevents the normal operation of the proportional tube. Finding a gas mixture that minimises these problems is largely done empirically. It is found that not only the gas mixture, but also all other materials used in the construction of the detector have a very strong influence on the ageing effect. Trace amounts of impurities in the gas or in other construction materials can have a dramatic effect on the ageing properties of a wire chamber. With correct choice of materials and proper care, it is found that a wire chamber can cope with a total accumulated charge of more than 1 coulomb/cm of anode wire. This means that with a gas gain of 10^4, and a rate of 10^4/s minimum ionising tracks per mm of anode wire, the chamber will work for 10 years without problems.

Figure 4.13 shows the signal amplitude in a proportional tube as a function of the externally applied voltage. In this plot, it was assumed that the primary charged particles produce 100 electron–ion pairs in the gas. The following operational regimes of the chamber can be distinguished.

Ionisation chamber: at a voltage of a few 100 V, the device works as an ionisation chamber. All the charges are collected but there is no charge amplification. The amplitude of the pulse is therefore only 100 electron charges.

Proportional regime: As the voltage is increased, charge multiplication sets in. The charge amplification increases more or less exponentially with the voltage. As long as the gain is less than about 10^5, each electron arriving at the anode receives the same amplification, regardless of how many electrons there are in total. In this case, the amplitude of the output pulse is proportional to the amount of primary ionisation in the gas. The names 'proportional regime' and 'proportional counter' refer to this property of the counter.

Non-proportional regime: As the voltage is increased further, the gain becomes very large and the proportionality property no longer holds. This is because the first electrons arriving at the anode undergo a very large amplification. After the electrons from this first avalanche have reached the anode, they leave behind a large
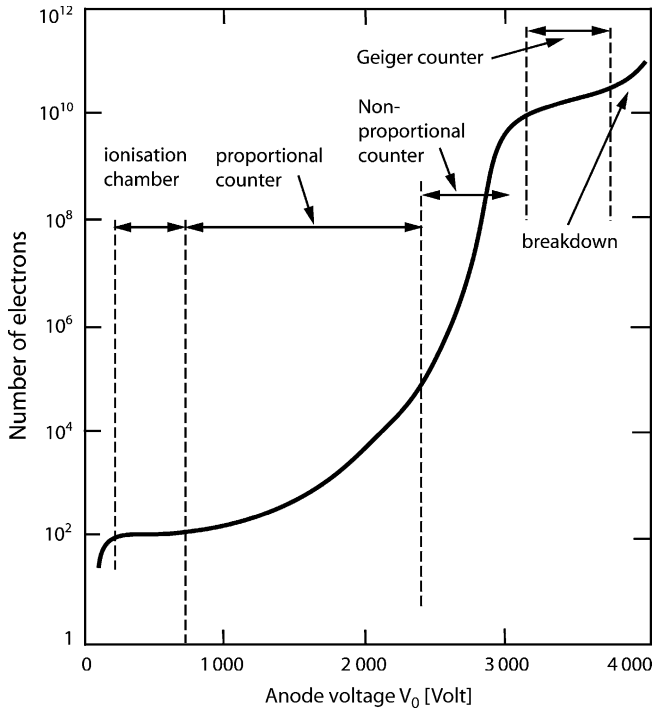
Fig. 4.13 Number of electrons in the signal caused by one minimum ionising particle in a counter with gas amplification. The different operating regimes of the counter as a function of the anode voltage V_0 are shown

number of positive ions that drift slowly towards the cathode. These positive ions will weaken the field close to the anode wire. Therefore, the electrons arriving later see a smaller field and are amplified less than the first electron.

The Geiger regime: If the voltage is increased further, at some point the Geiger regime is reached. A stable Geiger regime is only reached for suitable gas mixtures.

In this regime, the avalanche extends laterally along the anode wire and eventually fills the whole tube. The extension of the avalanche is due to UV photons formed in the process of avalanche multiplication. These UV photons will ionise molecules some distance away from the original avalanche, inducing further avalanches. Eventually, the anode wire is surrounded along its full length with avalanches.

The avalanche formation stops because the space charge of all these positive ions left behind reduces the electric field close to the wire to a point where there is no more avalanche formation. The voltage drop over the external resistor in series with the voltage power supply also contributes to stopping the avalanche multiplication process.

In the Geiger regime, the pulses are very large and can be several volts. The drawback is that the counter has a very large dead time because one must wait

Fig. 4.14 Some commercial radiation monitoring devices based on ionisation in gas. Photograph by courtesy of IAEA, from Ref. [14] in Chap. 3
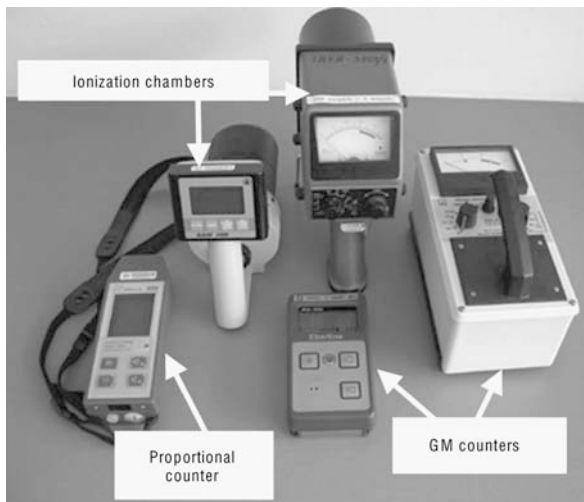
until all the positive ions have been evacuated, and this typically needs several 100 µs. In the proportional regime, the number of charges in each avalanche is much smaller and a new event at the same point is possible before all the charges have been evacuated. Moreover, the avalanche is localised at one point along the wire. The rest of the length of the wire is not affected at all and is ready to accept new events. The count rate achievable in proportional tubes is therefore about 10^7 pulses per meter of wire, several orders of magnitude larger than what can be achieved with a Geiger counter.

Figure 4.14 shows a few, commercially available, gas based counters.

4.5 Applications of Counters with Gas Amplification

Detectors for subatomic particles based on gas amplification have found many applications. The main application is as a device for localising trajectories of high-energy charged particles, i.e. as 'tracking' detectors. However, proportional tubes also make excellent detectors for thermal neutrons, for low-energy X-rays and for beta electrons. In the present section, I will discuss the use of proportional tubes and detectors derived from them, such as multi-wire proportional chambers or drift chambers, as tracker detectors and as X-ray detectors. The discussion of their use as neutron detectors is deferred to Chap. 7.

Gas amplification-based detectors are usually relatively simple devices and therefore tend to be inexpensive. In particular, if a large detector area is needed, the lower cost of gas amplification systems makes this often the preferred solution. In a proportional tube, the signal is generated by the slow motion of the ions, the signal

therefore has a slow rise-time and a long duration. This results in poor time resolution and large dead times. A number of different gas amplification-based devices have been developed to overcome these limitations, and these are briefly discussed in Sect. 4.6.

4.5.1 Proportional Counters for X-Ray Detection

Proportional counters are not very suitable for the detection of high-energy gamma rays, since the probability for the gamma ray to interact in the gas of the counter is small. The main sensitivity to high-energy gamma rays comes from interactions of the gamma rays in the walls of the detector. The interaction probability of X-rays in argon drops to a very low value above 20 keV. With krypton or xenon, sizeable detection efficiencies up to100 keV can be obtained. This is illustrated in Fig. 4.15.

For X-rays of the order of 10 keV, the photoelectron can be fully contained in the gas and the counter signal will be proportional to the energy of the X-ray. The energy resolution that can be obtained depends on the number of primary electron–ion pairs formed. However, the fluctuations on this number are not well described by Poisson statistics. The fact that the total energy used to create electron–ion pairs must equal the energy of the gamma reduces the fluctuations. It is customary to express this with the help of an empirical factor called the 'Fano factor'. For a Poisson distribution we have $F = 1$. The r.m.s. dispersion σ on the number of charges can be written as
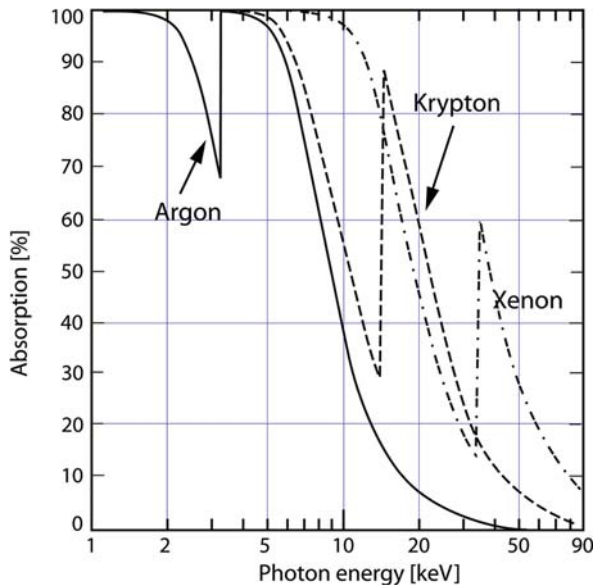

Fig. 4.15 Probability to absorb X-rays in 5 cm of argon gas, krypton gas and xenon gas at standard temperature and pressure. The data for this figure were obtained from [9] in Chapter 1

$$\sigma = \sqrt{Fn} = \sqrt{\frac{FE}{W}}$$

where n is the number or electrons produced in the gas and W is the energy needed to produce one electron–ion pair given in Table 4.1. The relative energy resolution, expressed as 'full width at half maximum' (FWHM) is therefore given by

$$FWHM = 2 \times 2.35 \frac{\sigma}{E} = 2 \times 2.35 \sqrt{\frac{FW}{E}}$$

In deriving this expression we have used that the 'full width at half maximum' for a Gauss distribution is 2.35σ. The Fano factor F in gases is typically about 0.1. On the other hand, each electron is not amplified by the same amount. These fluctuations in gain degrade the energy resolution by about a factor of 2. This is the excess noise factor and this effect will be discussed more extensively in Sect. 6.4.

The energy resolution of X-rays of 5.89 keV in argon is about 11% FWHM.

4.5.2 Gas Counters for the Tracking of High-Energy Charged Particles

In high-energy physics experiments, one often wants to measure the direction and the energy of all the particles produced in a collision. In nearly all experiments, there is a magnetic field and the momentum of the charged particles can be obtained from the curvature of the trajectory. Determining the trajectory of charged particles is called tracking. This requires the measurement of the space coordinates of a sufficient number of points along the track. Measuring three points is sufficient to determine the trajectory, but there are usually many tracks in the same event, and with only three points per track it is impossible to know which points belong to the same track. For this reason one usually needs to measure many more points along each track.

The momentum resolution of the particle is related to the position resolution on the points along the track. In the simplest case, where only three equidistant points are measured along the trajectory, this momentum resolution is obtained as follows. Let P denote the projection of the momentum on the plane perpendicular to the magnetic field. If the multiple scattering and energy loss are negligible, the trajectory of the particle projected on this plane is a circle. The radius R and the momentum P are related by Eq. (3.1), $Pc = 0.3\,Z\,B\,R$, where Pc is expressed in GeV, R in meter, B in tesla and Z in proton charges. The sagitta 's' is related to the length of the track 'L' and the curvature 'R' by (see Fig. 4.16(a)):

$$s = (1 - \cos\theta)R \approx \frac{\theta^2}{2}R \approx \frac{L^2}{8R}$$

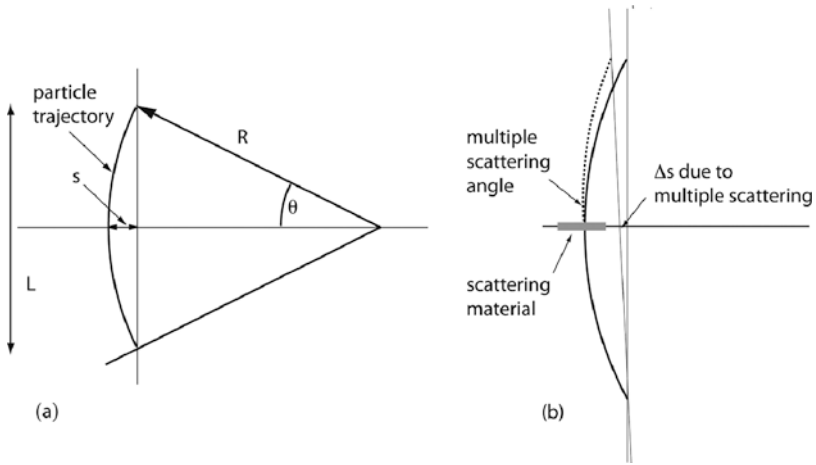**Fig. 4.16** (**a**) Relation between the curvature and the sagitta for an arc segment. (**b**) Effect of the presence of scattering material in the middle of the track. Full line, unscattered track, dotted line, scattered track

We therefore have

$$\frac{\sigma\{s\}}{s} = \frac{\sigma\{R\}}{R} = \frac{\sigma\{P\}}{P}$$

$$\frac{\sigma\{P\}}{P} = 8\sigma\{s\}\frac{Pc}{0.3\,Z\,B\,L^2}$$

In the case of only three equidistant points along the trajectory, the error on the sagitta is related to the error on the position measurement of the points σ by

$$\sigma\{s\} = \sqrt{\sigma^2 + \frac{\sigma^2}{2}} = \sqrt{\frac{3}{2}}\sigma$$

The momentum resolution due to the measurement error on the space points is therefore given by

$$\left[\frac{\sigma\{P\}}{P}\right]_{SP} = 8\sqrt{\frac{3}{2}}\frac{\{Pc\}[\text{GeV}]}{0.3\,Z\,B[\text{tesla}]\,L^2[\text{m}]}\sigma$$

It can be shown that, if there are N equidistant points along the trajectory, this expression generalises to (see [6])

$$\left[\frac{\sigma\{P\}}{P}\right]_{SP} = \sqrt{\frac{720(N-1)^3}{(N-2)N(N+1)(N+2)}}\frac{\{Pc\}[\text{GeV}]}{0.3\,Z\,B[\text{tesla}]\,L^2[\text{m}]}\sigma$$

As expected, the momentum resolution improves as $1/\sqrt{N}$ for large values of N.

The effect of multiple scattering is not always negligible. If only three equidistant points are measured along the track, an expression for the error on the momentum due to multiple scattering is obtained as follows. As a first step let us assume that all the scattering material is concentrated in the middle of the track. At this middle point the direction of the track is changed by an angle given by Eq. (2.5). We are only concerned about the angle projected on the plane perpendicular to the magnetic field and this projection gives rise to the additional factor $1/\sqrt{2}$ in the expression below.

$$\sqrt{\langle \theta_p \rangle^2} = \frac{1}{\sqrt{2}} \frac{Z}{\beta P c} (0.02\,\text{GeV}) \sqrt{\frac{L}{X_0}}$$

In this equation Pc is the momentum in units GeV, L is the total thickness of scattering material in the tracker and X_0 the radiation length of this material. If we assume that the scattering material is all concentrated in the middle of the track, the r.m.s distribution on the sagitta, caused by the presence of this scattering material, is given by (see Fig. 4.16(b)):

$$[\sigma\{s\}]_{MS} = \frac{L}{4}\sqrt{\langle \theta_p \rangle^2}$$

If the scattering material is not in the middle of the track but at a distance xL from one end, the dispersion $\sigma\{s\}$ will depend on the value of x. From the geometry of the problem, one easily finds that the r.m.s on the sagitta caused by multiple scattering is given by

$$[\sigma\{s\}]_{MS} = \frac{L}{4}\sqrt{\langle \theta_p \rangle^2} \quad f(x)$$
$$f(x) = 2x \qquad 0 < x < 0.5$$
$$f(x) = 2(1-x) \quad 0.5 > x > 1$$

If the scattering material is evenly distributed along the track, the total r.m.s squared is the sum of the squares of the contributions of all sections along the track, and we have

$$[\sigma\{s\}]_{MS} = \frac{L}{4}\sqrt{\langle \theta_p \rangle^2} \sqrt{\int_0^1 (f(x))^2\,dx} = \frac{L}{4\sqrt{3}}\sqrt{\langle \theta_p \rangle^2}$$

The contribution of the multiple scattering to the momentum resolution is hence given by

$$\left[\frac{\sigma\{P\}}{P}\right]_{MS} = \sqrt{\frac{2}{3}} \frac{0.02}{0.3} \frac{1}{L[\text{m}]B[\text{tesla}]\,\beta} \sqrt{\frac{L}{X_0}}$$

The above derivation rests on the assumption that only three equidistant points were measured along the track. However, it can be shown that the error due to multiple scattering is almost independent of the number of points measured along the track [6]. In real detectors usually the scattering material does not consist of only one type of material, moreover this material is usually not distributed homogeneously over the total length of the track. The generalisation of the above expression to this more general case is straightforward.

The total error on the momentum of a particle is the square root of the sum of the squares of the multiple scattering and the space point measurement error contributions. Notice that the error due to multiple scattering is independent of the momentum, while the error due to the position measurement is proportional to the momentum of the track.

Detectors based on gas amplification can be used to determine the trajectories of charged particles. The most straightforward way to obtain position information with proportional tubes is by using a multi-wire proportional chamber (MWPC). This detector consists of two conductive cathodes planes with a series of anode wires stretched in the middle between the cathode planes (see Fig. 4.17). The distance between the anode wires is typically 2 mm. The corresponding electric field configuration is shown in Fig. 4.18. A charged particle traversing the counter perpendicularly to the detector plane leaves a trail of ionisation behind in the gas. The electrons drift along the electric field lines to the nearest anode wires. The electric field geometry close to the anode wires is very similar to the field in a proportional tube and each individual anode wire behaves as a proportional counter. Consider an avalanche produced near one particular anode wire. This avalanche will induce a negative signal on this particular wire, and a positive signal on the neighbouring wires. This positive signal on the neighbouring wires compensates the negative signal on the same wires caused by capacitive coupling between the wires. As a result, a strong signal is only induced on the wire where the avalanche is formed. Each wire is, therefore, working as an independent counter. If each wire is equipped with its own readout electronics, the MWPC behaves as a counter giving the position in one direction. Each anode wire can handle a rate of up to 10^5 Hz/mm.

For perpendicular incident tracks, the position resolution depends on the wire spacing Δ. If we assume that the MWPC always gives the position of the wire

Fig. 4.17 Schematic representation of a multi-wire proportional chamber (MWPC). The cathode planes are at ground potential and the anode wires are at a large positive voltage, typically 3000 V. Each anode wire is equipped with its own amplifier and readout electronics. If a charged particle crosses perpendicularly to the plane of the detector, the wire nearest to the crossing point will have a signal

Fig. 4.18 Electric field geometry in a multi-wire proportional chamber. Figure from Ref. [6] in Chap. 1, with permission
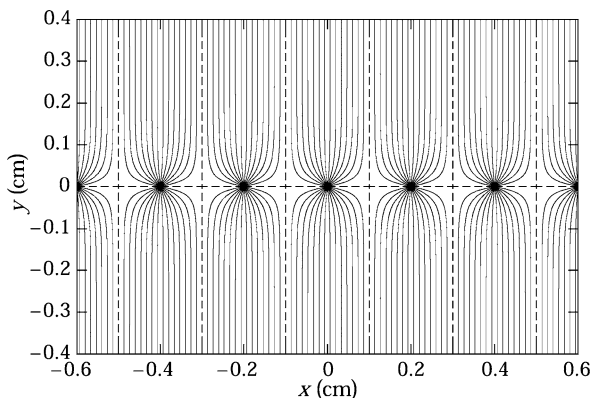
closest to the track, the r.m.s. position resolution is $\sigma = \frac{\Delta}{\sqrt{12}}$ (see Exercise 1). However, the spatial resolution depends on the readout electronics. If this electronics only lists the wires with a signal above some fixed threshold, the spatial resolution will be worse. If the electronics provides the amplitude for all the wires near to the trajectory, it is possible to calculate the centre of gravity of the signals and the spatial resolution can be significantly better than $\frac{\Delta}{\sqrt{12}}$.

Since a simple wire chamber as described above gives only one position coordinate, two superimposed chambers with crossed wire planes are needed to know both coordinates. This approach runs into trouble at high rates because, if several particles pass the two wire chambers at the same moment, it is impossible to know which x-coordinate is associated with which y-coordinate. To remove this ambiguity, more MWPC planes have to be added with the wires oriented in different directions. This is illustrated in Fig. 4.19. Another possible solution is to equip each wire with readout electronics on both sides and get an approximate position along the wire from the ratios of the signals at both ends. This method is called charge division and relies on the resistance of the wire itself.

A MWPC, in addition to the negative signals on the anodes, also has positive signals on the cathodes. The sum of all the induced negative signals equals the sum
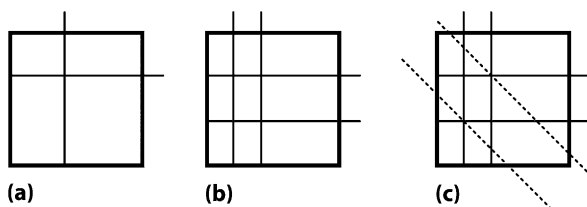

Fig. 4.19 (**a**) Two superimposed MWPC chambers with perpendicular anode wires allow the determination of the x and y coordinates of the particle track; (**b**) with two simultaneous particle tracks there are ambiguities regarding the exact position; (**c**) adding more MWPC planes under different angles allows resolving these ambiguities

Fig. 4.20 A drift chamber is a tube with a section as shown in this figure. The position of the track is derived from the time the charges need to drift to the anode. Typical voltages are anode +2000 V, cathode 0 V and field shaping electrodes, −1000 V, −2000 V and−3000 V
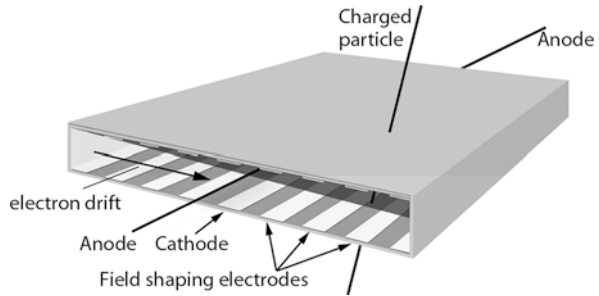

of all the induced positive signals. If, for the cathode plane we take a printed circuit board with a strip pattern orientated perpendicularly to the anode wires, one chamber can give the two coordinates of a track. In this case, the cathode signals are spread over several cathode strips, and a more elaborate electronics reading the amplitude of the signals on the strips is needed to obtain a good spatial resolution.

If the event rate in the detectors is not very large, the cost of the electronics can be considerably reduced by using a drift chamber. The principle of a drift chamber is illustrated in Fig. 4.20. Assume that a charged particle traverses the detector as indicated in the figure. A number of field shaping electrodes create an electric field pushing the electrons towards the anode wires. Close to the anode wires the field is similar to the field in a proportional tube and each electron creates an avalanche. If the time when the particle passed through the chamber is known, we can derive the distance between the trajectory of the particle and the anode from the time difference between the passage of the particle and the arrival of the corresponding anode pulse.

A few examples of applications of wire chambers are given below.

In high-energy physics collider experiments, beams of particles are made to collide at a given point in a beam pipe. A large number of particles are produced in each collision and the corresponding tracks radiate outwards from the collision point. One needs to observe all the particles emerging from the interaction point and determine the trajectories of these particles with the best possible accuracy. If there is a magnetic field, the momentum of the charged particles can be obtained from the curvature of the track. In e+ e− colliders, the rates are not very high and therefore this is often done with a large drift chamber called a time projection chamber (TPC). To explain the principle of a TPC, it is easier to consider first a simple TPC box as illustrated in Fig. 4.21(a). Assume we have a box with dimensions of the order of 1 m and we want to reconstruct the trajectories of all charged particles in this box. The box is filled with a typical wire chamber gas such as argon with 10% of isobutane. At the bottom of the box we have a plane of anode wires. The plane opposite the anode wires is brought to a large negative potential that will create a uniform electric field pushing all the electrons produced in the active volume towards the anode plane. In order to have a uniform electric field in the box, the sidewalls are covered with electrodes at intermediate potentials between the top and the bottom of the box. As the electrons reach the anode wires, the avalanches induce signals on

Fig. 4.21 (**a**) The principle of a TPC is explained with the help of this imaginary TPC box-shaped detector. The position of a charged track in the volume is obtained from the drifty time (*z*-coordinate) and from the signals induced in the cathode pads close to the anode wires (*x* and *y* coordinate). (**b**) Most TPC detectors are used at electron colliders. The TPC is a cylindrical detector surrounding the interaction region. The two halves of the cylinder form two independent detector chambers with the wire planes and readout pad planes forming the end flanges of the cylinder. Typical dimensions are given
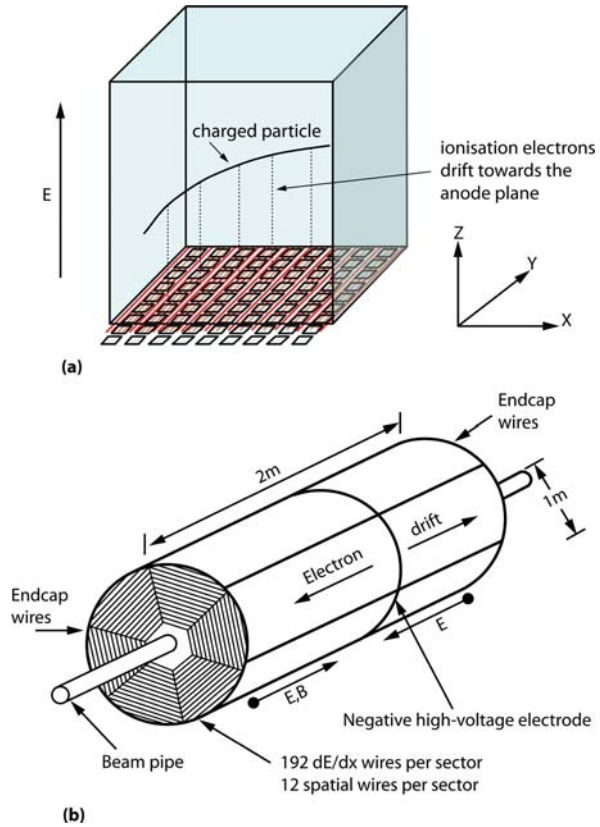
the cathode pads next to them. If the cathode pads are equipped with amplifiers and readout electronics, we can reconstruct the trajectories of all charges particles in the volume of the box. The *z*-coordinate is given by the drift time and the pad identifies the *x*- and *y*-coordinates.

In a collider experiment, owing to the overall geometry, the TPC should be a cylinder surrounding the beam pipe as shown in Fig. 4.21(b). The plane in the middle of the cylinder, just at the point of the collision, is an electrode brought to a large positive voltage. This plane divides the detector in two identical half chambers. At both ends of the cylinder there are planes of anode wires in the shape of a spider's web. Outside of these anode wires we have planes with pads. Besides the cylindrical shape, a TPC works in the same way as the TPC box described before.

In hadron colliders, such as the recently built LHC accelerator at CERN, the particle flux is too high to use conventional wire chambers. Only in the muon detection part, after a considerable amount of shielding material, is the particle flux low enough to use MWPC or drift chambers. In the CMS experiment, the muon detection in the barrel part uses layers of drift chambers similar to the ones described in
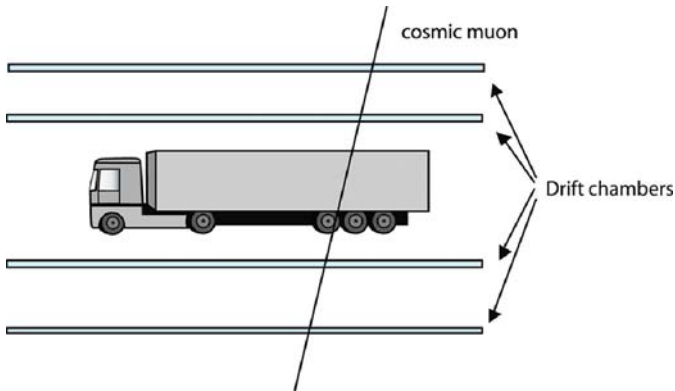
Fig. 4.22 The multiple scattering on the trajectories of cosmic ray muons can be used to detect the presence of a massive piece of lead shielding inside a lorry

Fig. 4.20 and in the end caps, where the rates are larger, it uses MWPCs with cathode strip readout. The muon system also has resistive plate chambers. The principle of this type of gas detector will be explained later in this section.

4.5.3 Applications of Gas Counters in Homeland Security

A completely different example of the use of wire chambers in charged particle tracking is the installation that is planned for use in several major US ports and that is illustrated in Fig. 4.22. The aim of the equipment is to prevent terrorists from smuggling a nuclear warhead into the US. A nuclear warhead can easily be detected by the emission of gamma rays in the MeV range. To prevent the detection of these, the terrorists could place the warhead in a massive lead box with a wall thickness of 5 cm or more. This would, indeed, stop most of the gamma rays and therefore would avoid detection. However, such a massive lead shield would reveal itself by the multiple scattering it would cause for cosmic ray muons. A complete lorry can be placed between several large detector planes made from drift tubes, each several metres long. The position along the wire is obtained from charge division. With this system, in 1 min a sufficient number of muons to detect the presence of such a massive lead shield can be observed and measured.

4.6   Recent Developments in Counters Based on Gas Amplification

An MWPC is a relatively simple and robust detector. It is most useful in a situation in which charged particle tracks have to be detected and localised over a large area. Its inherent limitations are due to the wire spacing and the ion mobility. It is difficult to build wire planes with a wire spacing of less than 1 mm. This limit is due, in part,

to the practical difficulty of constructing wire planes with smaller wire spacing and also due to the electrostatic repulsive forces between the wires. These forces push the wires above and below the plane and make the chamber unstable. This problem can only be avoided by using short wires, which removes much of the advantages of the MWPC. The low mobility of the positive ions results in a long rise-time for the pulses and in the build-up of an ion cloud around the wires. This, in turn, results in a poor time resolution and limits the maximal count rate to approximately 10^4 mm^{-1}s^{-1}.

To overcome these limitations, a number of variants of the standard MWPC or drift chamber have been developed. These devices take advantage of techniques to produce micro-patterns that have been developed over the last decades, mainly for the microelectronics industry. A few of these detectors are discussed below.

4.6.1 Micro-strip Gas Counters (MSGC)

A Micro-Strip Gas Counter (MSGC) is an ionisation chamber, where the anodes and cathodes consist of thin metallic electrodes deposited on an insulating substrate, usually a glass plate. Very narrow ($\approx 7\,\mu$m) anode strips alternate with wider cathode strips ($\approx 100\,\mu$m), making a periodic structure with a pitch of typically $200\,\mu$m (see Fig. 4.23). The anodes are at a positive potential relative to the neighbouring cathode strips.

The construction of such a fine electrode pattern is only possible by using photolithography, a technique commonly used to make masks for integrated circuits.
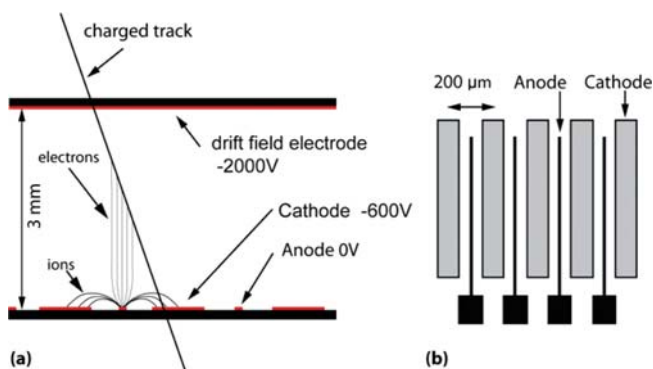
Fig. 4.23 Geometry and typical operating voltages of a Micro-Strip Gas Counter (MSGC). (**a**) The plane with the anodes and cathodes forms one wall of the gas gap; the other wall is a plane with a fully conductive surface. Both planes together create an electric field that pushes the electrons towards the cathode plane. The gas gap is typically 2 mm wide. Typical trajectories of electrons and positive ions are shown. (**b**) Structure of the plane with the anodes and cathodes. Typically, the anode strips are 7 μm wide, the cathode strips 100 μm wide and the periodicity of the structure is 200 μm

Using this technique, it is possible to produce very fine strip structures with dimensions in the micrometer range. The electrode geometry of a MSGC is illustrated in Fig. 4.23(b).

Some of the charges produced in the avalanche close to the anode will reach the surface of the insulating substrate; the surface of this substrate will become charged and this will modify the geometry of the electric field. To prevent the accumulation of positive charges on the surface of the insulator, a proper choice for the resistivity of the substrate is essential. This can be achieved either by using low-bulk resistivity glass, or by using a low-surface resistivity obtained by ion implantation or by applying diamond-like coatings. Typically, the value of the resistivity is 10^{15} or $10^{16}\,\Omega$/square. Lower values must be avoided because this would give rise to an unacceptably large dark current in the detector. The surface resistivity of normal glass is of the order of $10^{18}\,\Omega$/square or more.

The MSGC has several advantages over the MWPC. Owing to the short ion path between the anode strip and the neighbouring cathode strips, the time needed for evacuating the cloud of positive ions is considerably shorter. This results in a shorter rise-time of the pulses and in better count rate performance. Counting rates at least two orders of magnitude higher than those of a standard MWPC can be achieved. The position resolution depends, among other things, on the pitch of the structure, which in a MSGC is typically 5–10 times smaller than in a MWPC. A space resolution of 30 μm has been achieved with a structure with 200 μm pitch. Furthermore, in an MSGC, the periodicity of the structure can be maintained with very high accuracy over the entire detector surface, resulting in an identical distribution of the electric field lines and hence a homogeneous gas amplification. This means that the energy resolution of an MSGC when used as an X-ray detector will be very good.

One of the main limitations of the MSGC, compared to a MWPC, is the small value for maximum gas gain that can be achieved. This is due to the presence of a strong electric field parallel to the substrate surface. The highest gain that can be achieved is a few 10^3. This is at the limit of what is required to detect the signals from minimum ionising particles. To overcome this limitation, other geometries such as the MICROMEGAS and the GEM have been proposed. These detectors are now discussed.

4.6.2 GEM and MICROMEGAS Counters

GEM stands for 'gas electron multiplier'. In this structure, a completely different geometry is chosen to try and overcome the limitations of the proportional chambers. Figure 4.24(a) shows the structure of a GEM and typical values for the dimensions. The essential element of the GEM detector is a thin, self-supporting three-layer mesh consisting of a thin (50 μm) insulating polyimide (Kapton) foil, metal clad on both sides and with a regular pattern of small holes in it (see Fig. 4.24(b)). Typical dimensions are: holes of 70 μm diameter and a 120 μm
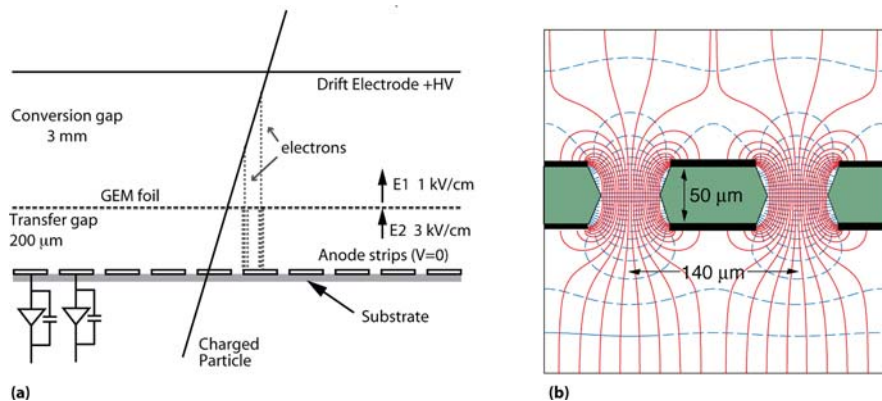
Fig. 4.24 (**a**) Schematic representation of a GEM detector with typical values for the electric field and dimensions. Electrons are liberated along the track of the charged particle and drift toward the GEM holes. Inside the holes, there is a large electric field multiplying the number of electrons. (**b**) Details of electric field lines (solid) and equipotential surfaces (dashed) in the region of the GEM holes. Electron transparencies are typically 100%. Most positive ions produced in the high-field region within a hole drift back to the GEM's top side. Figure from Ref. [6] in Chap. 1, with permission

pitch. Such a mesh can be made by conventional photolithographic methods used to produce multilayer PCBs. When applying a voltage across the metal sheets on both sides of the mesh, a very high electric field is generated in the centre of the channels (40 kV/cm is achieved with 200 V voltage difference) as shown in Fig. 4.24(b). A GEM counter has a conversion gas gap, a few mm thick, where the charged track ionises the gas. Over this conversion gap there is a voltage difference pushing the electrons towards the GEM foil. The top electrode typically is a thin metallised Mylar sheet. Electrons produced by ionisation in the gas-filled conversion gap drift into the channels and multiply in the high field present inside the channels. The electrons leave the channel and drift further towards a collecting electrode.

The ions produced during the avalanche tend to follow the field lines and are thus channelled towards the metal-clad top surface of the GEM. A stable and uniform amplification of over 2000 has been achieved with one GEM-electrode. This amplification is rather low, but it is possible to use several GEM foils with a small gap between them, resulting in an amplification that can exceed 10^6. Sometimes one GEM foil is combined with a MSGC counter.

A MICROMEGAS also has a conversion gap similar to a GEM counter, but a different structure is used for the amplification. It consists of the following components (typical dimensions are given):

– anode strips on a printed circuit or similar substrate, with a width of 100 μm and a pitch of 200 μm. The accuracy needed for this structure is much less than that required for an MSGC. Standard commercial PCB production techniques can be used.

- quartz fibres, or other spacers, with a thickness of 100 μm glued on top of the anode strips and defining the amplification gap. These spacers are necessary because of the electrostatic attraction between the mesh and the anodes.
- the micromesh: a metallic grid with 3 μm thickness, 17 μm openings and a pitch of 25 μm. It is made of nickel using the electroforming technique. Precision is better than 1 μm and the transparency is 45%.
- 3 mm thick gas-filled conversion gap.
- a drift electrode to create a field pushing the electrons created by an ionising particle towards the micromesh.
- as gas filling a typical proportional tube gases such as Ar+10% isobutane are used.

The working principle of a MICROMEGAS detector is as follows: ionisation electrons created in the conversion gap drift towards the mesh and are transferred through the micromesh to the amplification gap. In this gap between the mesh and the anode strips, there is a large electric field of the order of 10 kV/cm and this is sufficient to cause electron multiplication. Notice that in this amplification gap we have a parallel and uniform electric field, hence the amplification takes place over the full length of the gap, not only close to the anode strips. The anode strips collect the electron cloud, while the positive ions drift in the opposite direction and are collected by the micromesh. However, with this electric field geometry, the average electrons drift over an important fraction of the potential difference, therefore the electron contribution to the signal formation is much larger than with a proportional tube. Owing to this and because of the much shorter drift time of the positive ions, the signal of the MICROMEGAS is much faster than the signal of an MWPC.

An important feature of the MICROMEGAS is the fact that the micromesh is almost transparent to the electrons coming from the conversion gap and stops most of the positive ions coming from the amplification gap. It can be shown that these transparencies depend mainly on the ratio of the field strengths in the amplification and in the conversion gap. Under typical operating conditions, this ratio is large; hence, almost all electrons are transmitted from the conversion gap to the amplification gap, while only a small fraction of the positive ions is transmitted from the amplification gap to the conversion gap. Charging of the substrate between the anode strips will have little effect on the operation of the detector. A gain of 10^5 can be achieved in a MICROMEGAS chamber.

4.6.3 Resistive Plate Chambers

Another, and very different, type of detector based on the amplification in gases is the resistive plate chamber or RPC. This detector consists of a gas gap between two planar surfaces and a large voltage of 7–12 kV between the plates. The surfaces are resistive but the back side of the plates is made slightly conductive. If a charged particle traverses the gas gap an avalanche will form at this point. The

avalanche will remain localised because the resistivity is very high and the voltage drops immediately to a very low value, preventing further development of the avalanche.

The plates are made of high resistivity (10^9–10^{13} Ωcm) material, usually glass, phenol formaldehyde resin (Bakelite) or melamine laminate plates. The backside of the plates is made slightly conductive with a suitable coating to give it a surface resistivity of the order of 10^5 Ω/square. The unit Ω/square is the usual unit to express surface conductivity. It is the resistivity between two conducting lines on the surface if the distance between the lines is equal to the length of the lines. It is easy to see that this resistivity is independent of the dimension of the square. The unit Ω/square is somewhat confusing because the actual dimension of the quantity simply is Ohms.

A typical layout for an RPC is shown in Fig. 4.25(a). To make sure that the gap between the plates remains the same over the whole surface of the detector there are spacers between the plates approximately every 10 cm. Without spacers the electrostatic attraction between the plates would cause the plates to come closer together in the middle. The signal readout is via the metallic pickup strips separated by a thin insolating foil from the slightly conductive coating. The coupling is through
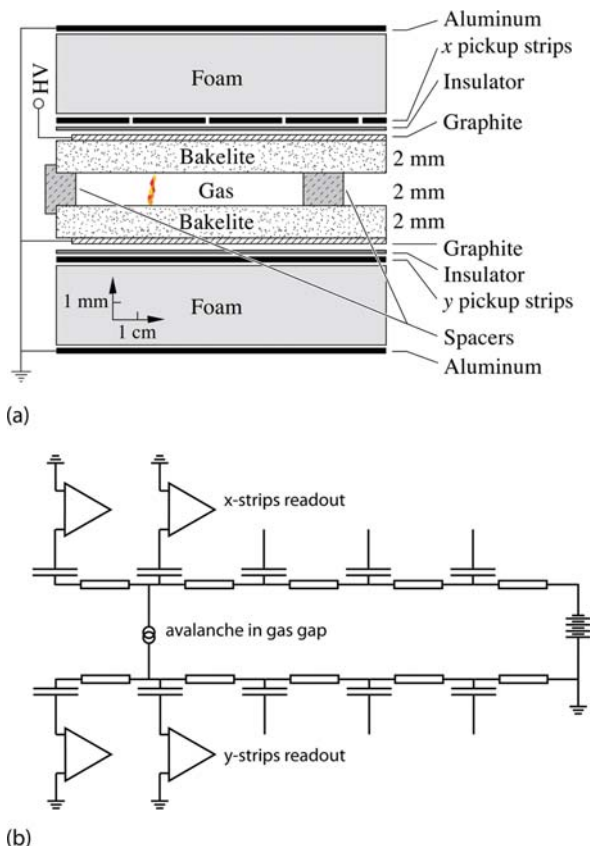

Fig. 4.25 (**a**) Typical structure of a resistive plate chamber (RPC). Figure from Ref. [6] in Chap. 1, with permission. (**b**) Equivalent electric circuit representing the readout of an RPC

the capacitance formed by the pickup strip and the semi-conductive coating and is schematically represented in Fig. 4.25(b).

RPC chambers typically have a surface of the order of 1 m^2 and are relatively inexpensive. They produce large signals, up to 300 mV, and allow good timing accuracy. The rate is limited by the time it takes to recharge the plates at the point where the discharge occurred. This time is quite long, of the order of 1 ms depending on the gain used, but this dead time is limited to the area of less than 1 cm^2 around the discharge point. The rest of the detector remains fully sensitive.

4.7 Exercises

1. Consider an MWPC with wire spacing Δ. Assume that for perpendicular tracks the signal is always on the nearest wire. Show that the r.m.s. position resolution obtainable with such a detector is given by $\sigma = \frac{\Delta}{\sqrt{12}}$.

2. You suspect that the gas in a cave is heavily contaminated by radon [$^{222}_{86}$Rn] gas. To determine the radon contamination you measure the current caused by the radon in an ionisation chamber containing one litre of air from the cave. You measure 0.1 pA. How much radiation expressed in pico Curie (pCi) per litre is there in the air of the cave? How many radon atoms per litre are there in the air of the cave?

   Radon has a half-life of 3.8 days and decays into alpha particles of 5.6 MeV nearly 100% of the time. To simplify the calculation, ignore the fact that radon decay products will also be present and will significantly contribute to the current. Also ignore the fact that often the alpha particle will hit the wall of the ionisation chamber and therefore will not use all its energy to ionise the air.

3. A GEM detector has a conversion gap of 2 mm. The gas filling is 90% Ar and 10% CH4. Cosmic ray muons are falling perpendicularly on this detector. What is the probability that a muon will be go undetected because there is no primary ionisation event in the conversion gap?

4. Calculate the mobility of nitrogen ions in nitrogen gas assuming that the cross section for the collision is 3.7×10^{-15} cm^2.

References

1. C. Powell, P. Fowler and D. Perkins, The study of elementary particles by the photographic method, Pergamon Press (1959)
2. S. Biagi, Data on electron-atom and electron–ion cross sections compiled, available from http://consult.cern.ch/writeup/magboltz/.
3. F. Sauli, Principles of operation of proportional and drift chambers, CERN 77–09 (1977). This article is reproduced in 'Experimental techniques in high energy physics', T. Ferbel, Editor, 1987, Addison-Wesly

4. W. Shockley, *J. Appl. Physics.* 9, 635 (1938)
5. R. Ramo, Proc. I.R.E., 27, 584, (1939)
6. R.L. Gluckstern, Nucl. Instr. and Meth. 24, (1963) 381, S. Tavernier, IIHE internal report 156 (1970), Statistical methods used in programs THRESH and GRIND
7. W. Blum and L. Rolandi, Particle detection with drift chambers, Springer (1994)

Chapter 5
Detectors Based on Ionisation in Semiconductor Materials

Semiconductor detectors are based on the detection of electron–hole pairs created in a semiconductor material by ionising radiation. Compared to gas-based detectors, an advantage of semiconductors is that the amount of energy needed to produce one pair of free charge carriers is about a factor 10 less. Another important difference is that the density of a semiconductor is typically 1000 larger than the density of a gas. The primary charge signal is therefore larger than in gas-based detectors. In semiconductors, however, it is much more difficult to obtain charge multiplication. All semiconductor detectors in use today are without any built-in charge amplification mechanism. As a result, the signals in semiconductor detectors are very small and extremely good low-noise electronics is essential.

Semiconductor detectors are expensive, because extremely pure starting material is required and only small detectors can be made. The most commonly used detector materials are germanium and silicon. Due to recent technological advances, cadmium–telluride and cadmium–zinc–telluride have now also been improved to the point where these materials have become useful as particle detectors.

References [1, 2, 3, 4] provide a more in-depth discussion of semiconductor radiation detectors, semiconductor devices and the fundamentals of semiconductors.

5.1 Introduction to Semiconductors

In this section, we briefly review some properties of semiconductors that are relevant to their use as detectors for ionising radiation. We always mention silicon, but similar considerations apply to germanium and other semiconductors.

The electronic configuration of silicon is the configuration of neon, plus two electrons in the 3s level and two electrons in the 3p level. These four valence electrons are involved in the chemical bonding of silicon. In a silicon crystal, all these electrons are shared between the atoms, i.e. the electrons are not bound to one particular atom and travel freely in the crystal. Each silicon atom forms four covalent bonds with four neighbouring silicon atoms. These four neighbours are located at the corners of a regular tetrahedron surrounding the atom. This gives rise to a lattice with the same structure as diamond. A silicon crystal therefore is not isotropic and

properties such as electron drift velocity will depend on the orientation of the movement relative to the lattice structure. These non-equivalent directions are designated by the Miller indices, e.g. (1,1,1), (1,0,0) etc. We refer the reader to the literature on solid state physics, for example [3], for the meaning of this notation.

In a crystal, the energy levels of individual atoms form so-called 'bands' with closely spaced energy levels. In between the allowed energy bands there are forbidden energy regions called 'band gaps'. An electron can never be in a stationary state with an energy corresponding to the band gap. The different energy levels inside a band are distinguished by the wave number k. The quantity $p^* = k\hbar$ is sometimes referred to as the 'quasi momentum', because it shares some properties with the momentum of a free particle. This analogy, however, should be used with caution. The relation between the electron energy and the wave number in the different bands is illustrated in Fig. 5.1. This diagram is called the electronic band structure, and it is obtained by solving the Schrödinger equation for the lattice. Let us assume that all the levels in the first and the second band are filled with electrons and that nearly all the levels in the third band are empty. In this case, the bands will be labelled core band, valence band and conduction band as indicated in Fig. 5.1.

In the conduction band there is one value of $p_0^* = k\hbar$, where the energy is minimum and around this minimum the energy can be parametrised as

$$E = E_c + (p^* - p_0^*)^2/2m_e^*.$$

Unlike the situation shown in Fig. 5.1, in silicon the minimum of the conduction band does not occur at the same k-value as the maximum of the valence band. If there are any electrons present in the conduction band, these electrons will sink to the bottom of the band. These electrons will behave as almost free negative particles with effective mass m_e^*. If any electrons are present in this band, they can move around freely, and for this reason this band is called the conduction band. If there is an electric field, these electrons will give rise to a net electric current.

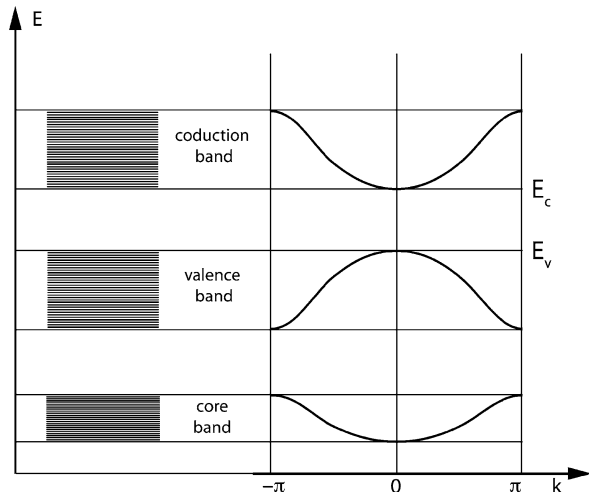
Fig. 5.1 Electronic band structure for a one-dimensional lattice. If all the levels in the first and the second band are filled with electrons, these are called the core and the valence band. The first empty band is called the conduction band

Nearly all the levels in the valence band are occupied by an electron. If all the levels are occupied by an electron, there is no net electric current. If any electrons are missing, these vacancies are called holes. The holes rise to the top of the band. Near the top of the band the energy of the levels around the maximum can be parametrised as $E = E_v - (p^* - p_0^*)^2/2m_h^*$. It can be shown that the vacancies near the top of the band behave as almost free positive particles with effective mass m_h^*. If there is an electric field present, these holes will also give rise to a net electric current.

In silicon, the four valence electrons form the valence band. The next higher electron energy level in silicon forms the conduction band. In between these two bands there is a band gap. The width of the energy gap between the valence band and the conduction band is the most important fact in determining the electrical properties of the material. If this energy gap is significantly larger than 1 eV, the material is an insulator. In a metal there is no energy gap between the two bands and the material is a conductor. If this energy gap is of the order of 1 eV, the material is a semiconductor. These different possibilities are illustrated in Fig. 5.2.

At absolute zero temperature all levels below a critical value E_F, called the Fermi energy, are occupied by electrons and all levels above the Fermi energy are empty. At any temperature different from zero, and in the condition of thermal equilibrium, the probability $f(E)$ that a particular energy level is occupied by an electron is given by the Fermi–Dirac distribution

$$f(E) = \frac{1}{e^{\frac{E-E_F}{kT}} + 1} \tag{5.1}$$

In this equation, T represents the absolute temperature and k the Boltzmann constant. The electron state where the energy of the electron is equal to E_F has 50% chance of being occupied by an electron. All the states with energy larger than the E_F have a probability of being occupied less than 50%, and all the states with energy smaller than E_F have a probability of being occupied larger than 50%. In insulators or semiconductors the Fermi energy is situated in the band gap. Except at absolute zero temperature there are always some electrons present in the conduction band and
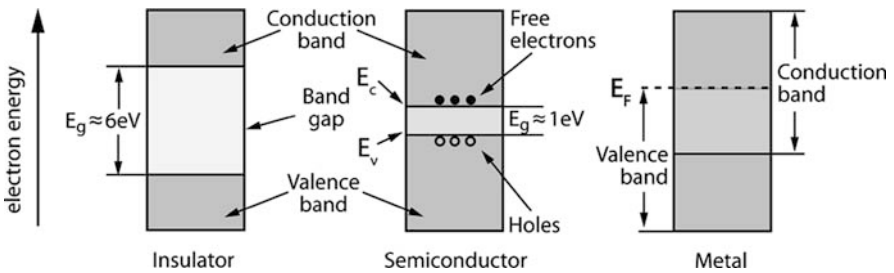

Fig. 5.2 Energy band structure of conductors, insulators and semiconductors. The vertical axis represents the electron energy, the horizontal axis the position in the lattice

some unoccupied levels, or holes, in the valence band. In an insulator, the band gap is large and the probability of having electrons in the valence band or holes in the conduction band is extremely small. The conductivity caused by such electrons or holes is negligible. In semiconductor materials, the band gap energy is much smaller, and this probability is no longer negligible. This conductivity is much smaller, however, than the typical conductivity of true conductors such as metals, hence the name semiconductors.

The density of the electrons in the conduction band of a semiconductor is given by

$$\frac{dn(E)}{dE} = \rho(E)\, f(E) \tag{5.2}$$

where $\rho(E)$ represents the density of electron states. Close to the bottom of the conduction band, this density of states is similar to the density of states of free particles enclosed in a cubic potential well. This last density is derived in Exercise 6 and is given by

$$\rho(E)\, dE = 4\pi \left(\frac{2m_e}{h^2}\right)^{3/2} \sqrt{E}\, dE$$

The density of states for electrons near the bottom of the conduction band is given by the same equation but with the energy replaced by $(E - E_c)$, where E_c is the energy of the bottom of the conduction band and the electron mass replaced by the effective electron mass '$m_e{}^*$' in the lattice.

The number of electrons per unit volume in the conduction band n_e is hence given by

$$
\begin{aligned}
n_e &= \int \rho(E)\, f(E)\, dE \\
&= 4\pi \left(\frac{2m_e^*}{h^2}\right)^{3/2} \int_{E_c}^{\infty} \sqrt{(E - E_c)} \frac{1}{e^{\frac{E - E_F}{kT}} + 1}\, dE \\
&= 4\pi \left(\frac{2m_e^*}{h^2}\right)^{3/2} e^{-\frac{E_C - E_F}{kT}} (kT)^{3/2} \int_0^{\infty} \sqrt{x}\, e^{-x}\, dx \\
&= 4\pi \left(\frac{2m_e^*}{h^2}\right)^{3/2} (kT)^{3/2} \frac{\sqrt{\pi}}{2} e^{-\frac{E_C - E_F}{kT}}
\end{aligned}
$$

Similarly, the density of hole states near the top of the valence band is given by the same equation but with the energy replaced by $(E_v - E)$, where E_v is the energy of the top of the valence band and the electron mass replaced by the effective hole mass in the lattice. In silicon, the effective electron mass and the effective hole mass are not very different from the true electron mass. This is not always the case for other semiconductors. The probability to have a hole in the valence band is given by $1 - f(E)$, where $f(E)$ is the Fermi–Dirac distribution given by Eq. (5.1). The number of holes per unit volume in the valence band n_h is therefore given by

$$n_h = 4\pi \left(\frac{2m_h^*}{h^2}\right)^{3/2} (kT)^{3/2} \frac{\sqrt{\pi}}{2} e^{-\frac{E_F - E_V}{kT}}$$

In very pure semiconductor materials the Fermi energy is halfway between the top of the valence band and the bottom of the conduction band. The density of electrons in the conduction band n_e and the density of holes in the valence band n_h are equal. This carrier density is called the intrinsic carrier density n_i and is proportional to

$$n_i \propto T^{3/2} \, e^{\left(-\frac{E_g}{2kT}\right)} \tag{5.3}$$

where E_g is the width of the band gap between the conduction and the valence band. Equation (5.3) shows that the conductivity in semiconductors will strongly increase with the temperature. In metals the conductivity is less dependent on the temperature and tends to decrease with temperature.

It is important to realise that the relation $n_i = n_e = n_h$ only holds for extremely pure and defect free materials. Let us now assume that the silicon contains trace amounts of phosphorus. Phosphorus is the next element after silicon in the periodic table and hence it has three 3p electrons. The additional 3p electron occupies an energy level that corresponds to the conduction band in silicon. Because of the higher nuclear charge Z of the phosphorus atoms, the energy of this level is a little lower than the corresponding level of silicon. If a silicon atom in the crystal is replaced by a phosphorus atom, this will create a localised level just below the conduction band and the 3p electron will occupy this level. The same holds for a number of other elements with the same electron structure as phosphorus, such as arsenic or antimony. The energy difference between these levels and the conduction band is very small and due to thermal agitation these electrons will jump to the conduction band and move freely around in the lattice, leaving the 3p phosphorus level empty most of the time. In a state of thermal equilibrium, the Fermi–Dirac Eq. (5.1) still describes the electron distribution, but the Fermi energy is now no longer in the middle of the band gap but much closer to the conduction band.

Similarly, if the silicon contains trace amounts of trivalent materials such as gallium, boron or indium, there will be empty acceptor levels just above the valence band, and these will give rise to holes in the valence band. In this case the Fermi level is close to the valence band. We see that trace amounts of impurities will strongly influence the concentration of electrons and holes in the crystal and hence the electrical properties of the material.

In the presence of impurities, the densities of electrons and holes in thermodynamical equilibrium are given by

$$n_e \propto T^{3/2} \, e^{-\frac{E_c - E_F}{kT}}$$
$$n_h \propto T^{3/2} \, e^{-\frac{E_F - E_v}{kT}}$$
$$\sqrt{n_e n_h} = n_i \propto T^{3/2} \, e^{-\frac{E_g}{2kT}}$$

If we call E_i the intrinsic energy level, i.e. the energy of the middle of the band gap, these equations can be written as

$$n_e = n_i \, e^{+\frac{E_F - E_i}{kT}}; \quad n_h = n_i \, e^{+\frac{E_i - E_F}{kT}}$$

If a pure sample of silicon is doped with a small concentration of electron donor atoms, the material is called n-type silicon. Similarly, if a silicon crystal is doped with an electron acceptor atoms, the material is called p-type silicon.

Consider a piece of n-type silicon. The position of the Fermi level is determined by the condition that the total charge in the crystal is zero. Therefore, the total number of electrons in the conduction band equals the number of holes in the valence band plus the number of empty donor levels. Usually, the number of holes in the valence band and the number of electrons left in the donor levels is negligible compared to the number of electrons in the valence band and we can write

$$n_e \approx N_D \approx n_i \, e^{+\frac{E_F - E_i}{kT}}$$
$$E_F - E_i = kT \ln \frac{N_D}{n_i}$$

The position of the Fermi level moves closer to the conduction band as the density of donors gets larger. Similarly, for p-type silicon, the position of the Fermi level moves closer to the valence band as the density of acceptor levels is larger. If both types of dopant are present, the dopant with the largest density determines the position of the Fermi level.

The electrons and holes in the crystal will move under the influence of an electric field. In the same way as charges in gases, the electrons or holes will constantly collide with 'obstacles' in the lattice and in each collision completely change direction. The time between collisions is of the order of 10^{-12} s. The thermal velocity of the electrons and holes is much larger than the drift velocity. At the macroscopic level, the drift of the electrons and holes under the influence of an electric field will look like a smooth process, but at the microscopic level the motion of the charge carriers will be completely chaotic. In analogy with the motion of charges in gases, the dependence of the electron velocity on the electric field is written as $v_e = \mu_e E$, where E is the electric field and v_e and μ_e are the electron velocity and the electron mobility, respectively. The corresponding quantities for holes are defined in the same way.

The electron and hole velocities are illustrated in Fig. 5.3. Notice that the velocities of electrons and holes are similar and that these are comparable to the electron velocities in gases. Notice also that, unlike the situation in gases, the drift velocity of electrons is only somewhat larger than the drift velocity of holes. This reflects the fact that the collision cross sections for electrons and holes are only slightly different. Because of this, it is much more difficult to achieve stable charge multiplication by the avalanche formation mechanism in semiconductors than in gases.

Table 5.1 lists some relevant physical properties of silicon and germanium. Notice the difference in radiation length and the difference in band gap energy.

To illustrate how trace impurities will strongly influence the electrical properties of a semiconductor, let us consider a cuboid of semiconductor material with a metal contact on two opposing surfaces (see Fig. 5.4).
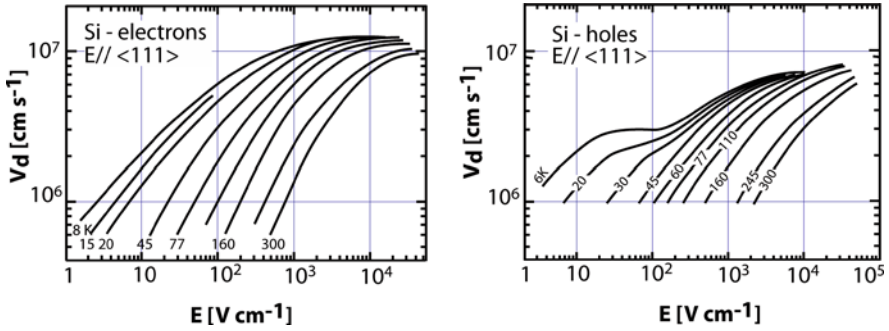
Fig. 5.3 Electron and hole drift velocities in silicon as a function of the electric field. The velocity shown here is the velocity in a plane parallel to the crystallographic <111> direction. The different curves correspond to different values of the temperature in degrees Kelvin. Figure from [5], © 1975 IEEE

From the definition of the resistivity ρ we have, where R is the resistance of this block

$$R = \frac{t}{A}\rho$$

$$i = \frac{V}{R} = \frac{A.V}{t.\rho}$$

On the other hand, the current in this block is related to the velocity of the electrons and holes by

$$i = e\,A(n_e\,v_e + n_h\,v_h)$$

Table 5.1 Some physical properties of silicon and germanium. Unless otherwise noted, data for silicon are at room temperature and data for germanium at 77 K

Physical property	Si	Ge
Atomic number Z	14	32
Atomic weight A	28.1	72.6
Density [g/cm^2]	2.33	5.32
Radiation length [mm]	93.6	23
Dielectric constant (relative)	12	16
Energy gap [eV)]	1.115	0.72
Intrinsic carrier density at 300 K [cm^{-3}]	1.5×10^{10}	2.4×10^{13}
Intrinsic resistivity at 300 K [Ω cm]	230,000	47
Electron mobility [cm^2 V^{-1} s^{-1}]	1350	36,000
Hole mobility at [cm^2 V^{-1} s^{-1}]	480	42,000
Energy/e–h pair [eV]	3.62	2.96
Fano factor	≈ 0.1	≈ 0.1
Energy loss min. ionising particles [MeV/cm]	3.87	7.29

Fig. 5.4 To measure the
resistivity of bulk silicon, a
parallelepipedal block of
silicon is equipped on two
opposing surfaces with ohmic
metal contacts
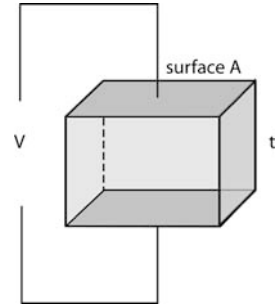
Using

$$v_e = \mu_e E = \mu_e \frac{V}{t}$$

We therefore have

$$\rho = \frac{1}{e(n_e\mu_e + n_h\mu_h)} \tag{5.4}$$

Often the number of charge carriers of one type is much larger than the number of charge carriers of the other type. In that case the resistivity of the material is entirely determined by charge carrier density of the majority charge carriers. The intrinsic carrier density and the mobilities of the charges in silicon are given in Table 5.1. From these data, one can readily calculate that the resistivity of intrinsic silicon is 230,000 Ω cm. A donor concentration as low as 0.2 ppb (2 donor atoms in 10^{10} silicon atoms) will reduce this bulk resistivity to 463 Ωcm.

This example shows that, when trying to collect the ionisation current in silicon, we will always draw a fairly large current. As is shown in Chap. 8, this gives rise to a large noise. To avoid this current we take advantage of the property of semiconductors to form diodes. If the diode is polarised in the reverse bias mode, a strong reduction of the dark current is obtained.

5.2 The Semiconductor Junction as a Detector

The operation of nearly all present-day electronic devices is based on the formation of n–p semiconductor junctions. The same principle is the basis of the use of semiconductors as detectors for ionising radiation.

An n–p junction is schematically represented in Fig. 5.5. In this junction, a region of n-type silicon is in contact with a region of p-type silicon. This should not be realised by pressing together two pieces of silicon, since the crystalline lattice should be continuous over the junction region. Both the n-type region and the p-type region are electrically neutral. However, in the n-type region there are many
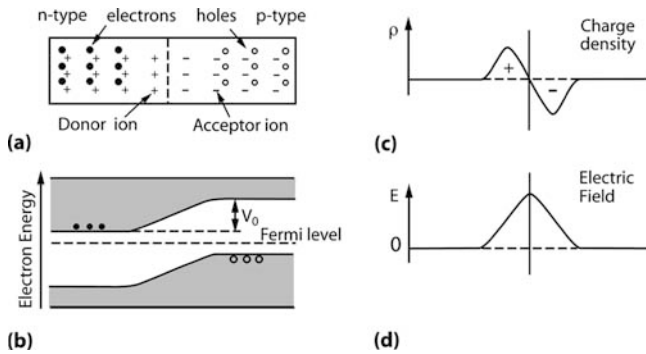
Fig. 5.5 (**a**) Schematic diagram of a n–p junction, (**b**) diagram of electron energy levels showing the creation of a contact potential V_0, (**c**) charge density, (**d**) electric field intensity

free electrons in the conduction band and almost no holes in the valence band, while in the p-type region there are many free holes in the valence band and almost no electrons in the conduction band. Owing to the thermal agitation, these charges will diffuse to the adjacent region of the other type, creating an excess of negative charge in the p-region and an excess of positive charge in the n-region. These charges will build up a potential difference and therefore an electric field over the junction region. This diffusion process will stop when the electric field generated is sufficiently large to prevent any further build-up of a charge difference. At this point the Fermi levels in the n-type silicon and in the p-type silicon are at the same level. In this contact area, the number of free charge carriers will be strongly reduced because free charges are removed by the electric field. This region with a reduced number of charge carriers is called the depletion region. This field will push any charges created in this region to the n-type side or the p-type side. Any charges created outside the depletion region will not be collected, but will simply recombine until thermal equilibrium is again reached.

It should be mentioned that this potential difference between the two sides of the silicon will not be observed as a potential difference between the two metal contacts applied to the two sides of the device. This is because at each metal–silicon contact there is also a contact potential and this will compensate the potential difference caused by the diffusion. Under condition of thermal equilibrium the Fermi levels in the silicon and in both metal contacts must be the same and therefore both metal parts must be at the same potential.

The structure represented in Fig. 5.5 will function as a diode. Let us assume we have made, so-called, ohmic metal contacts on the n-type and p-type silicon side. Such contacts only add or remove the majority charge carriers of each side and have a negligible resistance. Ohmic contacts will be discussed further at the end of this section. If, using these external metal contacts, we reduce the potential difference between the two sides, the majority charge carriers will again be able to flow to the other side and a current related to the concentration of majority charge carriers will flow. If, using these external metal contacts, we increase the potential difference

Fig. 5.6 (**a**) Profile of dopant concentration in an n–p junction. (**b**) Idealised profile of the dopant concentration in an n–p junction diode, as used in the calculation

(a)

(b)
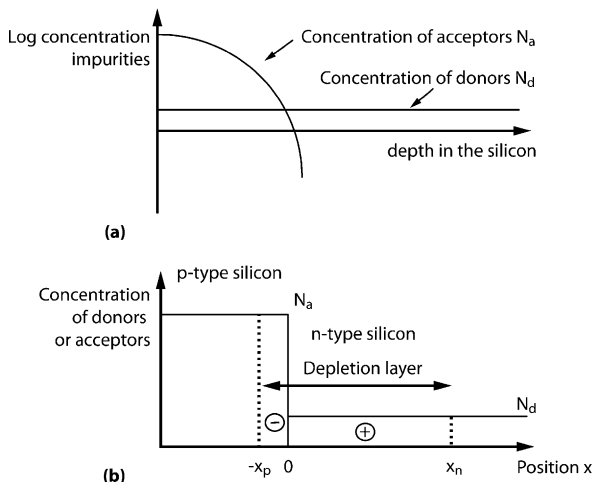
between the two sides, the majority charges will not be able to flow to the other side, and the current will be related to the concentration of minority charge carriers. That is, we have a large current if the junction is forward biased and a small current if the junction is reverse biased. As a result, the junction behaves as a diode. A diode is an electrical device that allows a large current to flow in one direction and only a small current in the other direction.

In practice, an n–p junction is usually made by starting from a homogeneous block of, say, n-type silicon. A dopant is diffused in the material from one side by exposing the silicon to a vapour of the dopant material, creating a region of p-type silicon with a finite depth. The resulting dopant concentrations in the junction are represented in Fig. 5.6(a).

Below we want to derive some essential properties of p–n junctions when used as detectors for ionising radiation. To derive these results we use a simplified model of such a p–n junction. We model the dopant concentration by assuming a concentration that is constant both in the n-type silicon and in the p-type silicon. This situation is illustrated by Fig. 5.6(b). For the calculation, we will furthermore assume the depletion layer is completely devoid of free charge carriers. The charge density in the depletion layer is eN_d and $-eN_a$ in the n-type
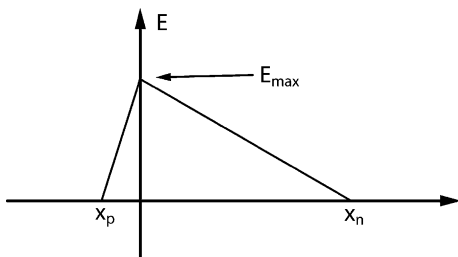

Fig. 5.7 Shape of the electric field over the junction

and p-type region, respectively, and $N_a \gg N_d$. Outside the depletion layer, the number of free electrons and holes exactly compensates the positive or negative space charge caused by the dopant atoms and, because of the presence of free charges, the electric field must be zero. The widths of the two parts of the depletion region are written as x_n and x_p, with $x = 0$ taken to be at the border between the two types of silicon. Applying the Maxwell equation, $\int D \, dv = \int q \, dv$, over a pillbox-shaped volume containing the junction, one can see that the total charge in the junction should be zero, therefore we have $N_d.x_n = N_a.x_p$, and hence $x_n \gg x_d$.

The electric potential in a piece of silicon with a junction can be described by a Poisson equation. In this equation, $\rho(x)$ represents the charge density and e is the dielectric constant of the medium.

$$\frac{d^2 V}{dx^2} = -\frac{\rho(x)}{\varepsilon}$$

The diode can be seen as made up of four regions with different charge densities. In each region, the charge density is constant and it is therefore quite easy to solve the Poisson equation in each region separately. At the boundary between the regions, the potential and the electric field should be continuous. The four regions are

$$
\begin{array}{llll}
\text{Region1,} & V_1(x): & -\infty < x < -x_p & \rho = 0 \\
\text{Region2,} & V_2(x): & -x_p < x < 0 & \rho = -eN_a \\
\text{Region3,} & V3(x): & 0 < x < x_n & \rho = eN_d \\
\text{Region4,} & V_4(x): & x_n < x < \infty & \rho = 0
\end{array}
$$

In the equations above e represents a positive number equal in magnitude to one electron charge. In region 1 and 4, there are free charge carriers and there is no electric field. We therefore have

$$\frac{dV_1}{dx} = 0, \Rightarrow V_1(x) = \text{constant}$$

$$\frac{dV_4}{dx} = 0, \Rightarrow V_4(x) = \text{constant}$$

The potential $V_3(x)$ is found by solving the Poisson equation in region 3, i.e. for $0 < x < x_n$.

$$\frac{d^2 V_3}{dx^2} = -\frac{eN_d}{\varepsilon}$$

$$\int \frac{d^2 V_3}{dx^2} dx = \frac{dV_3}{dx} = -\frac{eN_d}{\varepsilon} . x + C_1$$

$$\int \frac{dV_3}{dx} dx = V_3(x) = -\frac{eN_d}{\varepsilon} \frac{x^2}{2} + x. C_1 + C_2$$

Using the continuity condition at the boundary between region 3 and region 4, we have

$$\frac{dV_3}{dx} = 0 \text{ for } x = x_n, \text{ and therefore } C_1 = \frac{eN_d}{\varepsilon} . x_n$$

$$V_3(x) = \frac{eN_d}{\varepsilon} \left(-\frac{x^2}{2} + x . x_n \right) + C_2$$

By a similar calculation we find for the potential $V_2(x)$ in region 2:

$$V_2(x) = \frac{eN_a}{\varepsilon} \left(\frac{x^2}{2} + x . x_p \right) + C_2' \quad \text{for} \quad -x_p < x < 0$$

Using the continuity condition at $x = 0$, we have $V_2(0) = V_3(0)$ and hence $C_2 = C'_2$. The potential difference V_0 over the junction is therefore given by

$$V_0 = V_4 - V_1 = V_3(x_n) - V_2(-x_p) = \frac{eN_d}{\varepsilon} \frac{x_n^2}{2} + \frac{eN_a}{\varepsilon} \frac{x_p^2}{2} \approx \frac{eN_d}{\varepsilon} \frac{x_n^2}{2}$$

In our example where $N_a \gg N_d$, we have $x_n \gg x_p$ and the thickness d of the depletion layer is to a good approximation $d \approx x_n$. We see that the depletion region extends entirely towards the low dopant concentration region of the silicon. The thickness of the depletion layer depends on the smaller of the two dopant concentrations, in our example N_d. We obtain the following three useful relations:
Thickness of the depletion layer:

$$d = \sqrt{\frac{2\varepsilon V_0}{eN}} \tag{5.5}$$

In this and the following equations N represents the smaller of the two dopant concentrations. The capacitance of a parallel plate capacitor is given by

$$C = \frac{\varepsilon A}{d}$$

where 'A' represents the area and d the distance between the plates. We therefore have the following expression for the capacitance per unit area in a junction diode:

$$\frac{C}{A} = \frac{\varepsilon}{d} = \sqrt{\frac{\varepsilon eN}{2V_0}}$$

The capacitance of the junction is important because it represents an important source of noise (see Chap. 8). Capacitances well below 1 pF/mm^2 can be obtained.

From our calculation, it also follows that the electric field in the depletion layer has a triangular shape as shown in Fig. 5.7. The field reaches its maximum value E_{max} at the boundary between the two types of silicon. We therefore have the following expression for the value of the electric field at the point where it reaches its maximum value:

$$E_{\max} = \frac{2V_0}{d} = \frac{d\,eN}{\varepsilon} \tag{5.6}$$

Equation (5.5) relates the thickness of the depletion layer to the voltage difference over the junction and to the dopant concentration of the part of the junction with the smallest dopant concentration. The thickness of the depletion layer is important, because in a junction only the charges induced in the depletion layer are collected. The smallest dopant concentration that can be used is essentially determined by the purity of the starting silicon material. The purer the silicon, the lower the dopant concentration that can be used. Modern n-type high purity silicon typically has a resistivity of 20,000 Ω cm. From Eq. (5.4) we find that this corresponds to a donor concentration of $2.3 \times 10^{11}/cm^3$.

If no external voltage is applied, the potential difference over the depletion layer is of the order of 0.7 V and the corresponding thickness of the depletion layer is about 64 μm. To increase this thickness one can apply a reverse bias voltage over the diode. However, as the reverse bias voltage is increased, the maximum field over the junction also increases, and at some point this field is so strong that electrons and holes acquire sufficient energy to produce further electron–hole pairs, i.e. we have charge amplification. If the reverse bias voltage is increased further, eventually breakdown will occur.

From Eq. (5.6) we obtain a relation between the maximum thickness of the depletion layer and the breakdown voltage

$$d_{\max} = \frac{\varepsilon\,E_{breakdown}}{e\,N}$$

This equation shows that the maximum thickness of the depletion layer is inversely proportional to the dopant concentration. The smallest dopant concentration that can be used depends on the amount of impurities in the starting material. A thick depletion layer is only possible if extremely pure materials are used. The breakdown voltage is about 16,000 V/mm in high resistivity n-type silicon, allowing a depletion layer of up to 10 mm to be obtained in silicon. In germanium, a depletion layer of up to 10 cm can be achieved.

To have a useful radiation detector it is essential that all the metal contacts on the silicon are realised in such a way as to have a negligible resistance. These are called ohmic contacts. Figure 5.8 shows how this is realised in the case of a metal contact on n-type silicon. Over the contact layer between the metal and the silicon there is a potential difference equal to the difference in the work functions between the two materials. The work function is the energy needed to move an electron from a point inside the material at the Fermi level, to a point outside the material. A region of space charges and a depletion region in the silicon are associated with this potential difference. In a metal, the region with non-zero space charge is extremely thin. If the concentration of donors is large in the contact region of the silicon, the total depletion region is very thin and the electrons can tunnel through this potential barrier. Such a thin layer of heavily doped n-type or p-type silicon is denoted as n+ or p+ layers in the literature. Ohmic contacts allow the current to pass in both directions with a resistance that is small compared to the bulk resistance of the silicon.

Fig. 5.8 Energy levels in a metal contact on a piece of n-type silicon. The notation $e\Phi$ designates the work function. Figure (**a**) shows the levels in the absence of contact between the materials. Figure (**b**) shows the situation when there is contact
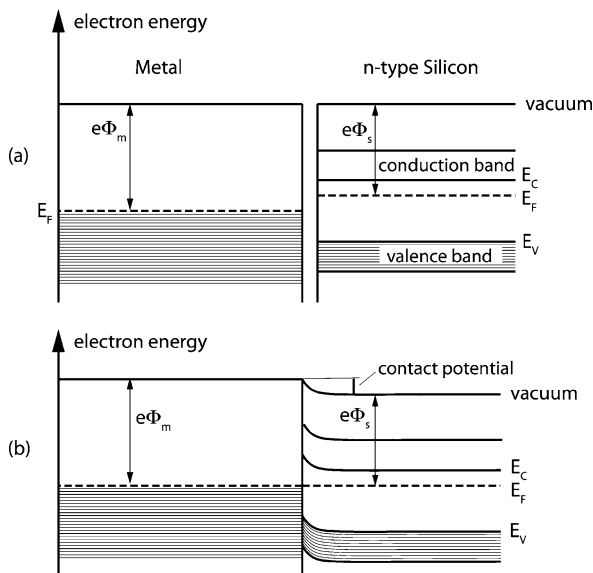

5.3 Silicon Semiconductor Detectors

Because of its rather long radiation length (93.6 mm), silicon is mainly used as a detector for charged particles. It can be used to track minimum ionising particles, and it is an almost ideal detector for alpha particles. The range of alpha particles of nuclear origin never exceeds 1 mm in silicon, and the amount of ionisation collected is therefore proportional to the energy of the alpha particle. In order to accurately measure the energy of the alpha particles it is important that the dead layer on the entrance side of the particle in the silicon is as thin as feasible. The amount of energy lost in this dead layer depends on the angle of penetration of the alpha particle and this will degrade the energy resolution. Usually, surface barrier detectors are used for alpha particle detection because this allows to obtain a very thin dead layer. In this type of detector, the junction is formed between the metal and the silicon. The resulting depletion layer behaves very much in the same way as discussed earlier. The metal contact also needs to be kept as thin as possible. Instead of using gas phase diffusion, ion implantation is often used as a method for achieving carefully controlled dopant layers. In ion implantation the surface of the silicon is exposed to a beam of ions produced by an accelerator. With this technique, entrance windows as thin as 34 nm of silicon equivalent can be achieved.

Figure 5.9 shows the energy spectrum of alpha particles recorded with a surface barrier detector. The energy resolution that can be achieved with this type of detector can be derived as follows. Let N be the number of electron–hole pairs created by the alpha particle. The energy of the alpha particle is proportional to the number of

Fig. 5.9 Alpha particle
spectrum of ^{234}U recorded by
a high-resolution surface
barrier detector. Figure from
[6] by courtesy of ORTEC
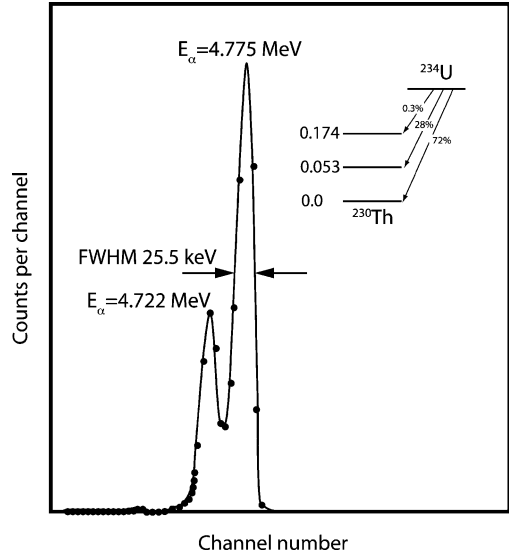
electron–hole pairs created: $E[\mathrm{eV}] = 3.62N$. The r.m.s. deviation of the measured
energy is given by

$$\sigma\{E\} = 3.62\sqrt{NF} = \sqrt{3.62\ E\ F}$$

As before, the Fano factor 'F' is due to the fact that energy conservation reduces
the fluctuations on the number of charges produced to be less than what it would
be for a Poisson distribution. The energy resolution, expressed as full width at half
maximum (FWHM) is therefore given by

$$FWHM[\mathrm{eV}] = 2.35\sqrt{3.62\ F\ E[\mathrm{eV}]}$$

The experimentally measured energy resolution is somewhat larger than what is
predicted by the formula above. For example, for an alpha particle of 5.5 MeV, the
formula predicts an energy resolution of 3.7 keV, but only about 10 keV is achieved
in the best detectors. The main reason for this discrepancy is the energy loss of alpha
particles due to elastic collisions with silicon nuclei. The recoil nuclei in these colli-
sions are usually too slow to produce any ionisation as silicon and the corresponding
energy is lost.

Silicon detectors are also commonly used to localise charged particle trajectories.
If the particles are minimum ionising particles, about 30,000 electron–hole pairs are
produced in a silicon slice of only 300 μm thick. That is a small signal but sufficient
to be detected by modern low-noise electronics. To provide particle localisation,

Fig. 5.10 Schematic representation of the structure of a silicon strip detector
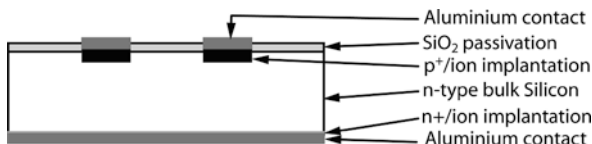
the electrode on one side of the silicon is subdivided into strips and each strip is connected to an amplifier (see Figs. 5.10 and 5.11). The distance between the strips is typically 200 μm. A detector of 300 μm thick needs a few 100 V to become fully depleted. Standard silicon wafers are round disks of silicon 5 or 8 inches (12.5 or 20 cm) in diameter. The largest detector that can be made is about 14 × 14 cm². To reduce the number of readout channels, often some kind of interpolation between the strips is used. If all the channels are equipped with electronics reading out the amplitude of the signals on all the strips, a position resolution of a few 10 μm can be achieved. The charges are collected in about 10 ns, making this indeed a very fast device.

Figure 5.12 shows a silicon strip detector with a readout allowing both the *x*- and the *y*-coordinate to be determined with the same detector. Silicon strip detectors are fast detectors since the charge is typically collected in 10 ns, but the small signal-to-noise ratio makes it difficult to obtain very good timing measurements with silicon.

Silicon strip detectors are prone to leakage currents, both in the bulk and on the surface, and particular care is needed to obtain reliable devices. These detectors are also sensitive to radiation damage.

Fig. 5.11 Charged particle tracking detector used in the CMS experiment. The detector consists of two wafers of silicon put side to side. Four amplifying chips with 512 amplifying channels each are visible in the top-left side of the picture. Each silicon wafer has strips as shown in Fig. 5.10, with a pitch of 180 μm. The r.m.s. spatial resolution is ≈25 μm. Photograph copyright CERN
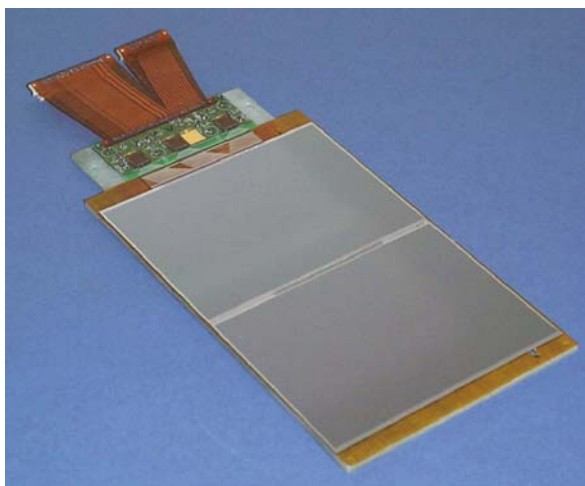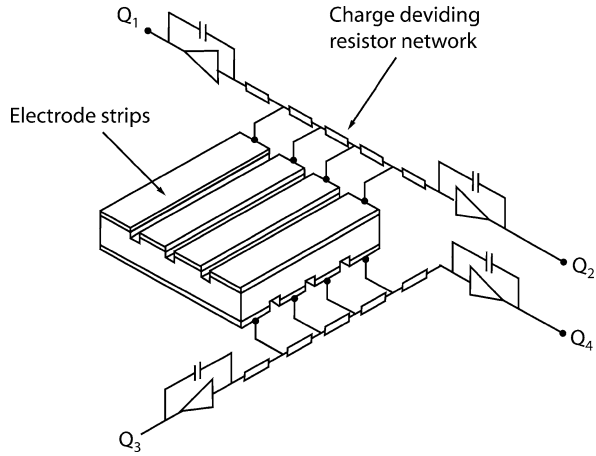
Fig. 5.12 To obtain the two coordinates *x* and *y* from one slice of silicon the detector can be equipped with mutual perpendicular strips on both sides of the silicon. In this example a resistor network was used to reduce the number of electronics channels

Q₁

Charge deviding resistor network

Electrode strips

Q₂

Q₄

Q₃

5.4 Germanium Semiconductor Detectors

The long radiation length of silicon (94 mm) and the fact that it is difficult to produce a depletion layer of much more than a few mm, makes silicon unattractive for gamma detection. Germanium has a radiation length of 23 mm, making it much more suitable for this purpose. Furthermore, modern germanium detectors are made from high purity germanium containing less than 10^{10} impurity atoms per cubic centimetre, making it possible to reach a depletion thickness of several centimetres. High purity germanium is grown in cylinder-shaped ingots. To achieve maximum use of the expensive material, germanium detectors are usually made in a cylindrical geometry as shown in Fig. 5.13(a) and Fig. 5.14.

Fig. 5.13 (**a**) Germanium detectors usually have a coaxial geometry to make optimal use of the expensive germanium material grown in cylindrical boules. (**b**) To suppress the leakage current the germanium detector must be used at liquid nitrogen temperature. It is often directly mounted on a Dewar as shown
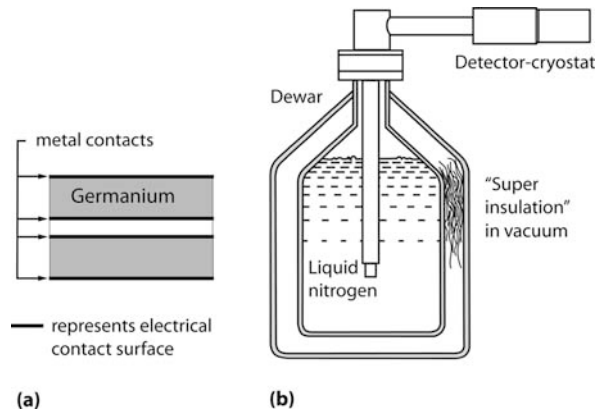
metal contacts

Germanium

— represents electrical contact surface

(**a**)

Detector-cryostat

Dewar

"Super insulation" in vacuum

Liquid nitrogen

(**b**)

Fig. 5.14 Germanium ingot
and Ge detector elements.
Figure by courtesy of
CANBERRA – an AREVA
company

Germanium has a band gap of 0.7 eV, giving rise to a large bulk leakage current
caused by thermal electron–hole pair creation in the depletion region. This current
gives rise to an unacceptable noise and therefore germanium detectors must be used
at reduced temperature, usually liquid nitrogen (77.2 K). The typical geometry of a
liquid nitrogen-cooled germanium detector is shown in Fig. 5.13(b).

Before modern high-purity germanium was available, it was common to rely on
compensated germanium, which was obtained using the lithium drift method. These
detectors must always (i.e. also when not in use) be maintained at liquid nitrogen
temperature to maintain the proper compensation.

Germanium makes a very good detector for gamma rays because of its excellent
energy resolution as is discussed below. For X-rays of the order of 10 keV, silicon is
more appropriate. However, germanium detectors are very expensive and need to be
cooled at liquid nitrogen temperatures for proper operation. Moreover, because the
signal formation depends on the drift of the charges over rather long distances, the
detector is not very fast. For all these reasons, it is often preferable to use scintillators
as is discussed in Chap. 6.

The energy resolution of a germanium detector depends on the fluctuations on
the number of charges. The number of charges produced by a gamma ray of energy
E is given by

$$N = \frac{E[\text{eV}]}{2.96}$$

The r.m.s. dispersion on this number is given by $\sqrt{NF}$ where F again represents
the Fano factor. The FWHM energy resolution is therefore given by

$$FWHM[eV] = 2.35 \cdot 2.96 \sqrt{\frac{E}{2.96} F} = 2.35 \sqrt{2.96 \, E[eV] \, F}$$

And the energy resolution FWHM in percent of the total energy is given by

$$FWHM[\%] = 235 \sqrt{\frac{2.96F}{E[eV]}}$$

For a gamma ray of 1.33 MeV, the above formula predicts an energy resolution of 1.33 eV. In practice, such a good energy resolution is never obtained. The difference is explained by incomplete charge collection.

5.5 Other Semiconductor Detector Materials

Silicon and germanium are by far the most commonly used materials in semiconductor detectors. However, there are several reasons to look for other materials. The properties of germanium make this an excellent material for making X-ray and gamma ray detectors, but the need to use the detector at the temperature of liquid nitrogen is a major complication. Moreover, detector grade germanium is expensive and the nuclear charge Z of germanium is only 32. Materials with a larger nuclear charge have a shorter radiation length and have a ratio of the photoelectric cross section over the Compton cross section that is larger. These are very desirable properties for a detector material used for detecting gamma rays.

Silicon is a good material for the tracking of charge particles, but the band gap in silicon is smaller than one would like and this causes a significant dark current and therefore a significant noise, when operating the detector at room temperature. A material with a somewhat larger band gap would be preferable. Moreover silicon is prone to radiation damage and possibly materials that perform better in this respect can be found. Furthermore, the good performance of silicon is only obtained with silicon monocrystals and these are limited in size to 5 or 8 inch. Materials that can be deposited in thin layers on large surfaces are needed in certain applications.

Table 5.2 lists a few materials that are being considered as alternatives to silicon or germanium. For all these materials the mean free path of the charge carriers before they are trapped, and therefore charge collection efficiency, is much lower than in silicon or germanium. This results in a significant reduction of the performance of the detector.

Of all the materials listed in Table 5.2 only CdTe (CT) and CdZnTe (CZT), and to less an extend HgI, have found significant applications as alternatives to germanium. CdTe suffers from the 'polarisation' of the detector material. This 'polarisation' is a change in the material caused by the previous interactions of gamma rays in the detector. This causes the detector response to become time dependent and rate dependent, a very undesirable property. The addition of a small amount of zinc reduces the dislocation density in CdTe and improves the performance. The concentration of zinc in CZT commonly used varies between 4 and 20%. CZT is less prone to polarisation and has a larger intrinsic resistivity than CdTe.

Table 5.2 Properties of some semiconductor materials for particle detection. All properties are given at room temperature, unless otherwise noted. Most of the entries in this table are taken from reference 4.5

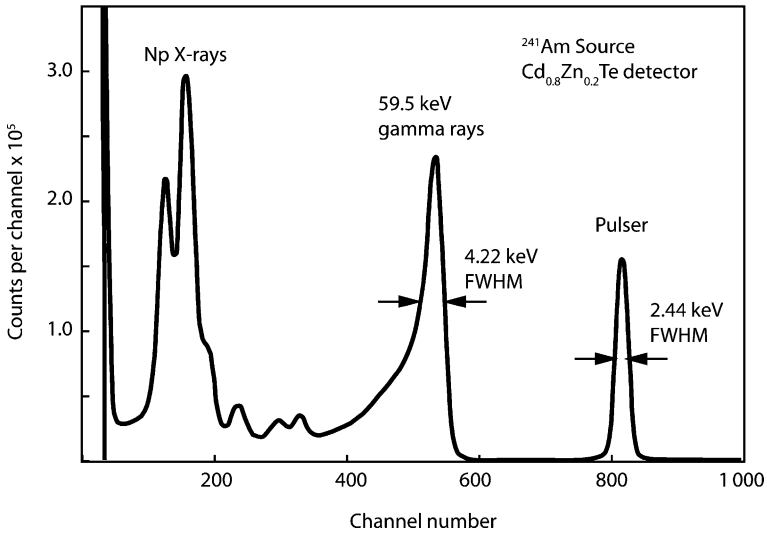| Material	Z	Density [g/cm³]	Radiation length [mm]	Bandgap [eV]	Energy per e–h pair [eV]	Intrinsic resistivity [Ωcm]	Electron mobility [cm²/(Vs)]	Hole mobility [cm²/(Vs)]	Electron lifetime [s]	Hole lifetime [s]
Si	14	2.33	93.6	1.12	3.62	320'000	1450	450	10^{-4}	10^{-4}
Ge	32	5.32	23	0.66 at 77 K	2.9 at 77 K	50	36000 at 77 K	42000 at 77 K	10^{-4}	10^{-4}
InP	49/15	4.97		1.35	4.2	$\approx 10^7$	4600	150		
GaAs (bulk)	31/33	5.32	23.5	1.424	4.2	$3.3\ 10^8$	>8000	400	10^{-8}	10^{-9}
CdTe	48/52	6.2	14.7	1.4	4.4	$\approx 10^9$	1000	80	10^{-6}	10^{-6}
$Cd_{0.8}Zn_{0.2}Te$	48/30/52	6		1.6	4.7	$\approx 10^{11}$	1350	120	10^{-6}	$2\ 10^{-7}$
HgI	80/53	6.4	11.8	2.13	4.3	$\approx 10^{13}$	100	4	$7\ 10^{-6}$	$3\ 10^{-6}$
Diamond	6	3.5	122	5.5	13	$>10^{11}$	1800	1200		
a-selenium	34	4.27	29		6–8					

Fig. 5.15 Typical pulse height spectrum taken with a 5×5×2 mm CZT detector. Data taken at 21°C and with a bias voltage of 150 V over 2 mm. Figure from [8], with permission

The production technology of CT and (CZT) has been improving steadily over the years. However, it remains difficult to produce large monocrystals and the materials suffer from poor hole collection efficiency. As a result of the hole trapping the pulse height spectra develop important 'tails', below the photopeak as shown in Fig. 5.15. These tails are due to ionisation events close to the anode. For these events the holes produced have to travel a long distance to the cathode. Since the holes are trapped before reaching the cathode, the signal is correspondingly reduced.

A number of ways have been proposed to overcome this problem. Because of the difference in drift velocity of electrons and holes, the pulse shape varies with the position of the gamma interaction relative to the readout electrodes. It is therefore possible to derive a correction for the pulse height based on the pulse shape. A different approach for reducing the sensitivity to hole trapping was developed by Luke [7]. It uses two different but coplanar anodes in such a way that the signal only depends on the movement of the electrons. This can be achieved by replacing the planar anode electrode by a set of fine parallel strips. The strips are connected alternatively to two different amplifiers, forming therefore two independent readout electrodes. One set of strips, hereafter called anode A, is brought at a slightly larger positive potential than the other set of strips, hereafter called anode B. In this way the electrons only collect on anode A. The motion of the charges at a large distance from the plane of strips induces the same signal on both anodes. The motion of the electrons close to the strips induces most of the signal on anode A only. The difference between the signals on both anodes, A and B, therefore only depends on the motion of the electrons in the last few 100 microns close to the anode.

Diamond is considered as an alternative to silicon for tracking in high-energy physics because of its good radiation hardness. Diamond of course is not a semiconductor, but the properties of diamond detectors are somewhat similar to the properties of semiconductor detectors. Its discussion is therefore included here. The radiation hardness of diamond is reported in the literature as an order of magnitude better than silicon. Detectors using true diamonds have been tried, but these are limited to very small sizes for obvious reasons. Most diamond detectors use synthetic 'diamond-like layers' obtained by chemical vapour deposition, referred to in the literature as CVD diamond. The band gap in diamond is larger than in semiconductor devices and the detectors made with CVD diamonds do not need a junction to suppress the leakage current. It is sufficient to use metal contacts on both sides of the diamond and apply an electric field. Because of the large band gap, diamond detectors can be used at elevated temperature.

Amorphous selenium (a-Se) is a good candidate for replacing silicon in applications where one would like a larger detector area than what can be obtained with silicon monocrystals. An example of such an application is the direct conversion X-ray detector for medical imaging. The simple band gap model used throughout this chapter does not really apply to amorphous selenium. For a number of reasons, the amorphous selenium layer in such X-ray detectors has to be operated at very high electric fields (up to 10 V per micron). An important issue in the use of amorphous selenium is the minimisation of dark current. Reference [9] contains an extensive discussion of the use of selenium as a detector for X-ray imaging.

5.6 Exercises

1. Calculate the dopant concentration in n-type silicon with a resistivity of 2000 Ω cm.
2. For a silicon strip detector made starting from n-type silicon with a resistivity of 2000 Ω cm, calculate the voltage to fully deplete a silicon microstrip detector of 300 μm thickness.
3. Derive an expression for the energy resolution (FWHM and in %) of silicon as an X-ray detector at room temperature. How much will the energy resolution be for X-rays of 50 keV?
4. You deposit a very thin layer of 241 Am with an activity of 2 MBq on a surface barrier alpha particle detector. This isotope emits alpha particles, see annex 6 for the characteristics of the emission. Calculate the magnitude of the pulses in number of electrons and the current in the detector.
5. Calculate the number of charges produced in a silicon strip detector of 500 μm thick by a minimum ionising particle.
6. Calculate the density of states for an electron enclosed in an infinitely deep and cubic potential well. Use the expression below for the energy levels of the electron in such a potential well. The numbers n_1, n_2 and n_3 are positive integers.

$$E = \frac{p^2}{2m} = \frac{\pi^2 (\text{h}c)^2}{a^2 2mc^2} (n_1^2 + n_2^2 + n_3^2)$$

7. Calculate the potential difference over a p–n junction if the dopant concentrations in the n-type silicon and the p-type silicon are $N_D = 10^{12}/\text{cm}^3$ and $N_A = 10^{16} \text{ cm}^3$, respectively.

References

1. B. Van Zeghbroeck, Principles of semiconductor devices, available from http://ece-www.colorado.edu/~bart
2. G. Lutz, Semiconductor radiation detectors, Springer (1999).
3. P.Y. Yu and M. Cardona, Fundamentals of semiconductors, Springer (2001).
4. C. Kittel, Introduction to solid state physics, 7th edition, John Willey & Sons (1995).
5. G. Ottaviani, C. Canali and A. Alberigi Quaranta, IEEE Trans. Nucl. Sci. NS-22(1) 192 (1975).
6. ORTEC application notes 'Alpha Spectroscopy with Surface Barrier Detectors', available from http://www.ortec-online.com/application-notes/an34/exp04.pdf.
7. P.N. Luke, Unipolar charge sensing with Coplanar electrodes – application do semiconductor Detectors, IEEE Trans. Nucl. Sci. 44(4), 207–213 (1995).
8. D.S. McGregor and H. Hermon, Room-temperature compound semiconductor radiation detectors, Nucl. Instr. meth. A395, 101–124 (1997).
9. G.S. Belev, Electrical properties of amorphous selenium based photoconductive devices for application in x-ray image detectors, Thesis, University of Saskatchewan, Saskatoon, SK, Canada (2007).

Chapter 6
Detectors Based on Scintillation

6.1 Introduction to Scintillators

When ionising radiation interacts with matter it will excite or ionise a large number of molecules. When these molecules return to the ground state, this will sometimes give rise to the emission of photons in the visible or near to the visible energy range. This phenomenon has as scientific name 'radioluminescence', but it is more commonly called scintillation. Observation of the scintillation process was one of the first techniques used for the detection of ionising radiation. Rutherford used zinc sulphide scintillating crystals in his famous scattering experiment that showed that all the positive charge in atoms was concentrated in the nucleus. Today the use of scintillators is still one of the main methods for radiation detection.

Many transparent materials will produce some small amount of scintillation light when hit by a high-energy particle or a high-energy photon, but usually this light signal is very weak. In a few materials, the conversion of the excitation energy into light is more efficient, and such materials are called scintillators. If the light emission continues for a long time after the excitation, i.e. much longer than 1 ms, this phenomenon is called phosphorescence rather than scintillation and the corresponding material is called a phosphor. Phosphors are often in the form of a thin layer of powder applied on a substrate. A different but related phenomenon is 'photoluminescence'. This is the emission of visible or near to visible light under stimulation by light of a shorter wavelength. Photoluminescent materials are also often called phosphors, but sometimes these materials are also called wavelength shifters or fluors.

Scintillation and the Cherenkov effect both are light emission effects, but the physical mechanism is completely different. Cherenkov light is only produced when the velocity of the particle is larger than the velocity of light in the medium. Also for particles travelling at a speed close to the speed of light, the intensity of the Cherenkov light emission is typically a factor 100 lower than the light output of a good scintillator.

When an ionising particle interacts in matter it produces a number of free charges. In a gas, applying a modest electric field over the gas gap is sufficient to collect these charges. In a solid, it is very difficult to collect the charges. Only in very few materials, such as silicon or germanium, is it possible to obtain efficient charge

collection. In a scintillator, the problem is avoided because the charges have only to travel to the nearest luminescence centre, often only a few atoms away. At the luminescence centre, the electrons give rise to photons. If the material is transparent, the light signal can then easily be extracted.

A scintillator should have the following properties to be useful as a detector for ionising radiation

– the material should be transparent at the wavelength of the emitted scintillation light
– the efficiency of light production should be large
– the light pulses should be as short as possible and there should be little or no delayed light emission
– the amount of light emitted should be proportional to the energy deposited by the ionising particle
– the refractive index of the material should be close to 1.5 so that light can easily be extracted from the scintillator.

Other desirable properties of the material are that it should be chemically and mechanically stable, not too difficult to produce and not too expensive.

Scintillating materials broadly speaking fall into two classes: organic and inorganic scintillators. The physics of the scintillation mechanism, the properties and the applications of both types of scintillating materials are very different. Both types of materials are discussed in this chapter. From the application point of view the important difference is that organic scintillators contain, for the most part, atoms with a small atomic charge Z, and have therefore a long radiation length. Inorganic scintillators are interesting, mainly because many of them contain a large fraction of atoms with a high atomic charge Z and therefore these materials have a short radiation length. Because of this difference in radiation length, inorganic scintillators are mainly used for X- and gamma-ray detection, while organic scintillators are mainly used for charged particle tracking. Another important application of organic scintillators is as detectors for 'fast neutrons', i.e. for neutrons with energy between ≈ 10 keV and 10 MeV. This application depends on the elastic scattering of neutrons on the hydrogen nuclei in the scintillator and is discussed in Chap. 7.

Scintillation also occurs in some inorganic gases such as nitrogen and in some inorganic liquids such as liquid xenon. The nitrogen in air emits a green glow when excited by energetic charged particles. This phenomenon causes the Northern Lights when the stream of charged particles emitted by the Sun enters the Earth's atmosphere near the poles.

6.2 Organic Scintillators

Three types of organic scintillators exist: organic crystals, organic liquids and plastic scintillators. Organic crystals such as anthracene and stilbene are efficient scintillators but, compared to plastic scintillators, they are expensive and difficult to

use. Therefore these materials have fallen in disuse and will not be discussed further in these lecture notes.

Organic liquid scintillators are obtained by dissolving an organic scintillator in an appropriate solvent. Often a wavelength shifter is added to improve the transparency of the liquid for the scintillation light or to obtain a better match between the spectral sensitivity of the scintillator and the photodetector. Liquid scintillators are less expensive than other scintillators and are therefore used when the application requires a large volume of scintillator. It is also used to count radioactive material that can easily be dissolved in the liquid. This technique is used for counting low beta activity as for example when counting ^{14}C for determining the age of archaeological samples.

The most widely used class of organic scintillators are the plastic scintillators. A plastic scintillator is made from a suitable polymerisable liquid, usually a liquid containing aromatic rings. Examples of such scintillating liquids are styrene and vinyltoluene. The base material in the plastic will scintillate in the UV, but the mean free path of the scintillation photons is only a few millimetres; therefore, a wavelength shifter, or fluor, needs to be added to the material. The fluor will absorb the primary UV scintillation light and emit it at a somewhat longer wavelength. At this longer wavelength the average mean free path of the photons is much larger. The concentration of the fluor in the scintillator is typically 1%. Sometimes, a second fluor is added at the ≈0.01% level to shift the emission to even longer wavelengths.

Plastic scintillators are often used because they are easily produced and can be shaped into whatever shape is required. In particular, they can be produced as large thin sheets or as fibres. A large number of different plastic scintillators are commercially available. Table 6.1 lists the main properties of one particular product, Kowaglass SCSN-32; these data are fairly typical of what can be achieved with plastic scintillators. An overview of the different organic scintillators available can be found in [5] in Chap. 1.

The main application of organic scintillators is as detectors for charged particles. Because of the short decay time of the scintillator, this detector can provide good timing information. To obtain position information the scintillator needs to be divided into narrow strips and each strip connected to a photodetector; for this purpose, fibres made from scintillating materials are often used. These fibres

Table 6.1 Properties of the plastic scintillator Kowaglass SCSN-32

Plastic type	Polystyrene-based scintillator
Light yield	8,000 photons/MeV, i.e. ≈16,000 photons/cm for minimum ionising particles
Decay time	3.6 ns
Emission wavelength	423 nm
Light attenuation length at 423 nm	250 cm
Optical refractive index	1.58
Density	1.08
Radiation length	30 cm

are aligned in a plane and all bundled on a suitable photodetector that allows identification of the fibres with a signal.

The typical geometry of a plastic scintillator used for detecting charged particles is shown in Fig. 6.1(a). The setup contains a sheet of scintillating plastic, a light guide and a photodetector. In many applications, one wants to read large sheets of plastic scintillator. Because of cost and for other reasons the sensitive area of the photodetector is usually much smaller than the area of the scintillator. One needs to make sure that a sufficient fraction of the scintillation photons arrives on the photodetector. The sheet of scintillator itself behaves as light guide and channels the light towards the four edges, using the effect of total internal reflection. Typically, each edge will receive about 10% of the scintillation light produced. A carefully shaped light guide brings the light from the edge of the scintillator to the photodetector using total internal reflection. The light guide is made of transparent plastic, usually polymethyl methacrylate (commonly called Plexiglas, Lucite or acrylic glass). The photodetector is usually a photomultiplier tube. It can be shown that the light guide will transmit light with good efficiency if the surface of the light guide in contact with the scintillator is the same as the surface of the light guide in contact with the photomultiplier. If the surface is smaller at the photodetector, the amount of light arriving at the photodetector is reduced in the ratio of these surfaces. This is a particular case of a general theorem in physics known as the Liouville theorem.

Because of this, the light collection efficiency can be very low if a large sheet of scintillating material is read by a small photodetector. A different approach to light collection is therefore often used and this method is shown in Fig. 6.1(b). In this method, a piece of transparent polymethyl methacrylate doped with a wavelength shifter is held against one edge of the scintillator. It is important that there is no optical contact between the scintillator and the wavelength shifter. The scintillation light is absorbed in the wavelength shifter and re-emitted at a longer wavelength.
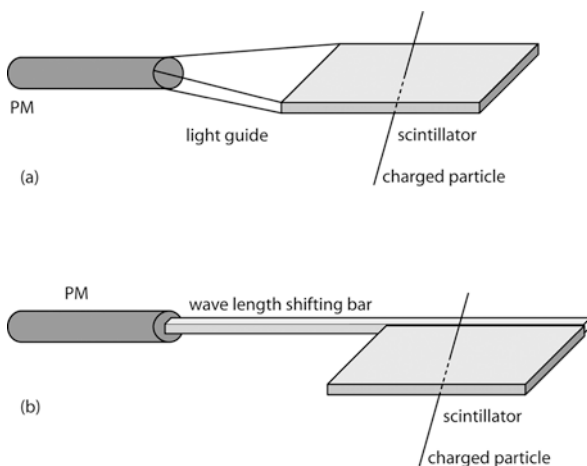

Fig. 6.1 (**a**) Plastic scintillator assembly as detector for charged particles with light guide readout. (**b**) Same with wavelength shifter readout

The direction of the photons emitted by the wavelength shifter is uncorrelated with the direction of the incoming scintillation light and again about 10% of this light is collected at each of the edges. If large sheets of scintillator are read with small photodetectors, this method of light collection is more efficient than the first method described.

6.3 Inorganic Scintillators

Inorganic scintillators are usually ionic crystals. The physical mechanism generating the light in an inorganic scintillator can best be explained with the help of the band model of a solid introduced in Chap. 6 and illustrated in Fig. 6.2. In this figure, the vertical axis represents the energy of the electron and the horizontal axis represents the position in the lattice along to one axis. The crystal has a valence band width, to first approximation, all energy levels occupied by one electron and a conduction band width to, first approximation, none of the energy levels occupied by an electron. Between the valence band and the conduction band there is an energy interval where there are no electron energy levels; an electron can never be in a stationary state with an energy corresponding to a point in this interval. If the ionic crystal is transparent to visible light, the band gap must be larger than ≈ 3 eV. In the simplest version of the model there is no way a crystal can emit light at a wavelength where the crystal is transparent. To make light emission possible we need luminescence centres that have localised levels in the crystal. Such localised levels can be intrinsically present in the material, but for the efficient scintillators these localised levels are provided by suitable dopant atoms introduced at the percent level in the host material. For the sake of definiteness, we will consider the case of a commonly used scintillator LSO. Similar considerations apply to other scintillators. Lutetium ortho-oxysilicate Lu_2SiO_5 (LSO) doped with cerium is an efficient
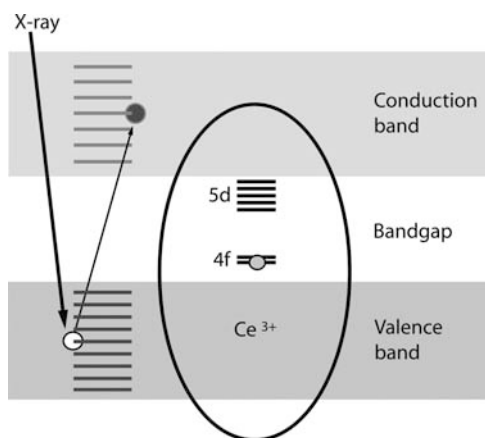

Fig. 6.2 Schematic representation of the electron structure in an LSO scintillator with a cerium 3^+ luminescence centre

scintillator. Neutral cerium has the electron configuration of xenon plus one electron in the 4f level, one electron in the 5d level and two electrons in the 6s level. When used as a dopant, cerium replaces lutetium in the lattice and will usually be in the Ce^{3+} ionisation state, hence only the 5d level is occupied by one electron. Inside an LSO host lattice the 4f and 5d levels are located near the bottom and the top of the band gap, respectively. The degeneracy of the 4f level is broken by spin-orbit coupling and the electron can have two possible energy values. The degeneracy of the 5d level is broken by the anisotropy of the host matrix and we have five possible values for the energy corresponding to the five possible values of the third component of the orbital angular momentum L_z. The Ce^{3+} ion therefore creates localised energy levels in the crystal as indicated in Fig. 6.2.

If a high-energy X-ray or one gamma ray interacts in the crystal, it will extract a deeply bound electron and raise it to an energy level in the valence band, or more often, to a significantly larger energy level. The electron will lose its extra energy by exciting further electrons and lifting them from the valence band or one of the deeper bands to the conduction band. Electrons from the outer bands fill the holes that are left in the core bands and the excess energy is again used to bring more electrons from the valence band to the conduction band. The net result is that one X-ray, or one gamma ray, produces a large number of holes in the conduction band and the same number of electrons in the valence band. These electrons sink to the bottom of the conduction band and the holes rise to the top of the valence band. All these happen in a very short time, in about 10^{-12} s. The minimum energy needed to create one electron–hole pair is equal to the band gap energy E_g, but unavoidably some of the energy of the initial X-ray or gamma ray ends up being converted into phonons, i.e. thermal energy in the crystal. At the end of the cascade of events, the number of electron–hole pairs created is given by

$$N_{e-h} = \frac{E_\gamma}{b\, E_g}$$

In this equation, E_γ is the energy of the X-ray or the gamma ray and the parameter b has a value that depends on the nature of the host matrix. For crystals of interest, this number is typically ≈ 2.

To have an efficient scintillator, the electrons and the holes should reach the cerium luminescence centres. This migration of the charges is the least understood part of the scintillation process. Indeed, the band model where electrons and holes move freely in the lattice is a gross oversimplification. There are always imperfections in the crystals and these imperfections form traps that can capture the free charges and prevent them from reaching the luminescence centre. Many of these traps are shallow traps that correspond to a small binding energy. If the temperature is sufficiently high, the thermal energy will be sufficient to liberate the charge from the trap. The retention time will strongly depend on the temperature. If the electrons and holes are sufficiently free to move around, a hole will first ionise the cerium atom and form a Ce^{4+} and subsequently an electron is trapped in the 5d level. These capture processes are efficient if the lowest 4f level and the highest 5d level are

close to the valence band and conduction band, respectively. This energy difference should be of the order of the energy of thermal agitations.

The electron captured by the 5d levels, quickly sinks to the lowest 5d level. The energy distance from this level to the 4f level is about 3 eV and this transition cannot happen by thermal agitation. The 5d → 4f transition is an allowed dipole transition and is therefore rather fast. In the case of cerium, this transition has a lifetime of the order of 40 ns, with the exact value depending on the host matrix.

The above model of the luminescence mechanism in inorganic scintillators obviously is incomplete. Indeed, it would seem that the light emitted by one luminescence centre will be absorbed by the other Ce^{3+} atoms in the crystal; hence, the crystal will not be transparent to its own scintillation light. The mechanism that avoids the above problem is called the 'Stokes shift' and it is explained in Fig. 6.3. When the electron of the Ce^{3+} final state is in the 4f state or in the 5d state, the spatial distribution of the charges is different. Therefore, the neighbouring ions have a different equilibrium position depending on the energy level of the electron.

For simplicity, I assume that the configuration of the ions round the luminescence centre can be described by a one-dimensional configuration coordinate as indicated in Fig. 6.3. Obviously, the configuration needs to be described by several variables, but with this simplification it is easier to explain the mechanism of the Stokes shift. Assume that the electron is in the 4f level. The lattice around the luminescence centre will find itself in the equilibrium position and this is the position where the energy of the electron takes the lowest value. If the electron absorbs a photon that rises it to the 5d level, at first the lattice remains in the geometry that corresponds to the electron in the 4f level. After the transition, the lattice relaxes and takes the geometry corresponding to the electron in the 5d level. This new geometry corresponds also to the minimum for the electron energy. If the electron jumps back to the ground level 4f, the energy of the emitted photon will be smaller than the energy used in the transition 4f → 5d. The energy difference between the absorbed photon and the emitted photon is called the Stokes shift. The corresponding absorption and the emission spectra of the crystal in function of the wavelength of the light are shown in Fig. 6.3(b). The distance between the edge of the absorption band and the maximum of the emission is the Stokes shift. Because of the thermal fluctuations in

Fig. 6.3 The Stokes effect allows a luminescence centre to have emission at a longer wavelength than the absorption. (**a**) The energy of the 4f and 5d levels depends on the configuration of the lattice around the luminescence centre. (**b**) Emission and absorption spectra for the electron transitions represented in (**a**)
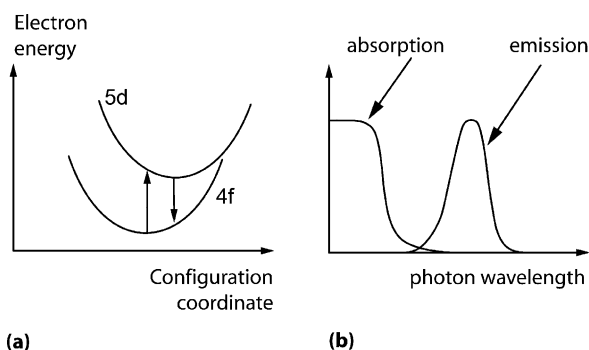
the crystal, the absorption peak is not very narrow and the band edge is not a sharp cut. Indeed, due to the thermal agitations, the configurations of the ions surrounding the luminescence centre are constantly fluctuating around the equilibrium position. The Stokes shift should be sufficiently large if the scintillator is to be transparent to its own scintillation light.

From the model described above it follows that the maximum number of photons, $N_{\max}$, a scintillator can produce is given by

$$N_{\max} = \frac{E_\gamma}{b\,E_g}$$

Let us take as an example the scintillator CsI:Tl. In this scintillator, thallium (Tl+) dopants play the role of luminescence centres. The band gap in CsI is 6.2 eV and we predict, taking $b = 2$, a maximum light yield of 80,000 photons/MeV. Experimentally it is found that CsI:Tl scintillator produces about 60,000 scintillation photons/MeV. We conclude that both the transport of the electron–hole pairs to the luminescence centres and the efficiency of the luminescence centre in CsI:Tl are rather high at room temperature.

The decay time of a scintillator is determined by the lifetime of the excited level in the luminescence centre. This lifetime also depends on the host lattice, but this is not a very large effect. All cerium-doped scintillators have a decay time of the order of 40 ns. The probability to have a transition 5d → 4f is independent of how long the electrons are already in this level, we therefore expect an exponential decay of the scintillation light. In practice, the scintillation light emission often deviates significantly from a simple exponential decay. As an example, Fig. 6.4 shows the decay spectrum of CsI:Tl. In a logarithmic plot, we clearly see two exponential decay components. There are two possible explanations for this behaviour. There can be more than one type of luminescence centre in the crystal. These can be two truly different kinds of dopant atoms, but it can also be the same type of dopant sitting in a different lattice environment. The dopant atom can for example sit next to another lattice defect and this can completely change the decay characteristics of the luminescence centre. Another possible explanation is that the transport of the charges to the luminescence centre is hindered by the presence of traps. If either the electron or the holes are trapped in a defect that retains these charges with a lifetime much larger than the decay time of the luminescence centre, we will obtain a slow component in the decay spectrum as shown in Fig. 6.4. Traps with a rise time much shorter than the luminescence centre will affect the rising edge of the scintillation signal. Instead of a very sharp rising edge, we will see a much slower increase to the maximum value.

The scintillation properties of inorganic scintillators tend to have a strong dependence on the temperature. This is illustrated in Fig. 6.5. For many materials the light output first increases with temperature, reaches a maximum and then decreases again. For BGO, the maximum is reached at a temperature below 100 K; therefore, only the decreasing part of the curve is seen. This behaviour can be understood as follows. At low temperature the light yield is low because there will always be shallow traps that capture the electrons or the holes and prevent the

Fig. 6.4 Logarithmic plot of the scintillation light intensity as a function of the time after the gamma interaction for CsI:Tl scintillation. If the light intensity decreases exponentially, this will show up as straight line in this plot. In this example, there are two components in the decay spectrum
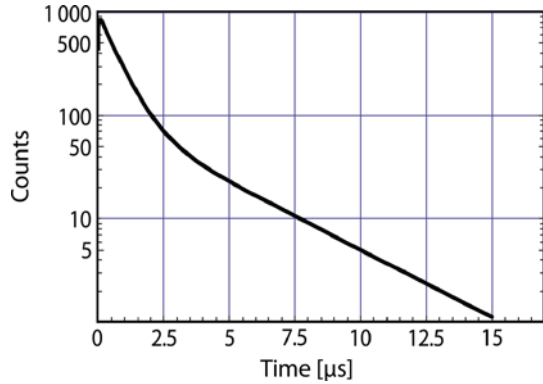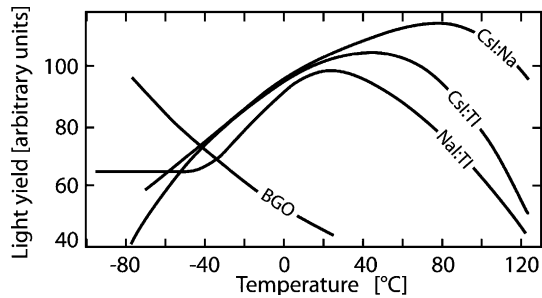
Fig. 6.5 Evolution of the light yield of some scintillators as a function of temperature. Figure adapted from [1], with permission

charges from reaching the luminescence centre. With rising temperature, these traps are no longer able to retain the charges and the light yield increases. If the temperature increases further, we have the phenomenon called thermal quenching. This can be understood with the help of Fig. 6.3(a). The energy distance between the 4f and 5d levels depends on the configuration of the host matrix around the luminescence centre. This configuration is not constant due to the permanent thermal motion of the ions around the luminescence centre. At some values of the configuration coordinate the distance between the two energy levels becomes small and the electron can jump to the lower level without emission of an optical photon, but by interacting with the phonons in the lattice. This is a non-radiative transition. As the temperature increases, the probability of this happening increases and the light yield decreases.

Inorganic scintillators are mainly used as detectors for gamma rays, usually with the aim of measuring the energy of the gamma ray. A linear relation between the light yield and the energy of the gamma ray is, therefore, a desirable property of a scintillator. For many scintillators there are significant deviations from this linear relation as is illustrated in Fig. 6.6. When a gamma ray interacts in a scintillator crystal there are a large number of different ways to deposit its energy. It can deposit all its energy in one photoelectron, but it can also undergo one or more Compton interactions before losing all its energy in a final photoelectric effect. If the response

Fig. 6.6 Ratio of the light
yield over the gamma energy
for a few commonly used
scintillators. The curves are
normalised such that this ratio
is unity if the energy of the
gamma ray is 1000 keV.
Figure from [2], © 1994
IEEE
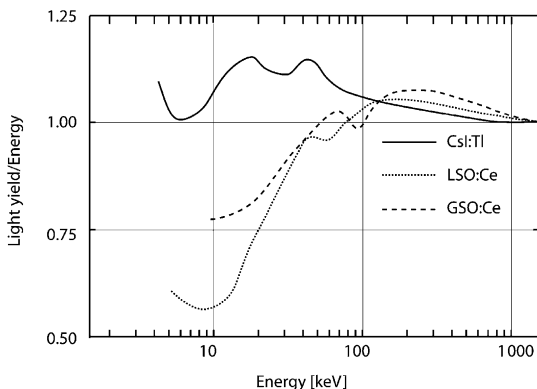
of the scintillator is not linear, this will result in a light output of the scintillator that
depends on the particular history of each gamma ray. This will degrade the energy
resolution. This effect seriously limits the energy resolution that can be achieved
with a scintillator for energies of the order of 1 MeV and less. At energies well
above 1 MeV, these effects tend to become less important.

The response of a scintillator also tends to depend strongly on the energy
loss density of the charged particles. Many scintillators have a strongly reduced
sensitivity to energy deposited by alpha particles. The light yield for the same energy
deposition by an alpha particle is usually several times smaller than for electrons.

Table 6.2 lists some commonly used inorganic scintillators. Each application has
a different set of requirements for a scintillator. In some cases, a fast decay time is

Table 6.2 Properties of some commonly used inorganic scintillators

Name*	Density	Emission λ [nm]	Light yield [photons/MeV]	Decay time τ [ns]	Radiation length [cm]
NaI:Tl	3.67	410	40,000	230	2.59
BGO	7.14	480	4000	300	1.12
BaF$_2$(fast)	4.88	215	1500	<1	2.05
BaF$_2$(slow)	4.88	310	10,000	700	2.05
CsI:Tl	4.51	565	65,000	600	1.68
CsF	4.11	390	2000	3	
PbWO$_4$	8.28	480	200	10	0.89
LSO:Ce	7.4	420	28,000	40	1.14
LuAP:Ce	8.3	360	10,000	18	
GSO:Ce	6.71	440	7500	60	1.38
LuPO$_4$	6.6	360	13,000	24	
YAP:Ce	5.37	370	16,000	25	2.7
LaBr:Ce	5.3	360	60,000	35	2.13

*The short names for the scintillators stand for the following chemical compounds: BGO =
Bi$_4$Ge$_3$O$_{12}$, GSO = Gd$_2$SiO$_5$, LSO = Lu$_2$SiO$_5$, LuAP = LuAlO$_3$, YAP = YAlO$_3$

essential, in others a large light yield and in still other applications a short radiation length is the most desired property. What is the best scintillator depends on the application.

6.4 Photodetectors

Scintillators would not be very useful as detectors for subatomic particles without some device to convert the light signal into an electrical signal. Such a device is called a photodetector. Photomultiplier tubes were invented more than 70 years ago, but are still the most commonly used photodetector for reading out scintillators. This is due to the fact that the light pulses obtained from scintillators are usually very weak, often less than a few 100 photons. A photomultiplier has a very large internal gain, therefore even a few photons are sufficient to obtain a detectable signal. An amplifier can also be used for amplifying weak signals, but as will be shown in Chap. 8, amplifiers always have noise and this noise is amplified together with the signal one wants to observe. Electrical signals that are too small are therefore not observable. In recent years, several new types of photodetectors with internal gain have been developed. Some of these are quite promising and could possibly one day replace photomultiplier tubes. These devices are briefly discussed at the end of this section.

The photomultiplier tube. The first photomultiplier tube was produced by Zworykin in 1936. Modern photomultiplier tubes have much improved performance, but the basic principle of operation is still the same. For a more in dept review of photomultipliers and their use, See [11, 12, 13]. A photomultiplier tube is a vacuum tube, usually made of glass or at least with a glass window and with a photocathode and a number of metal dynodes inside the vacuum. The structure of a photomultiplier tube is schematically represented in Fig. 6.7(a). A photocathode is a thin layer of a compound that will emit electrons when absorbing photons with a wavelength in the visible or near to the visible region. The photocathode can be deposited on a metal electrode inside the tube, but more often the
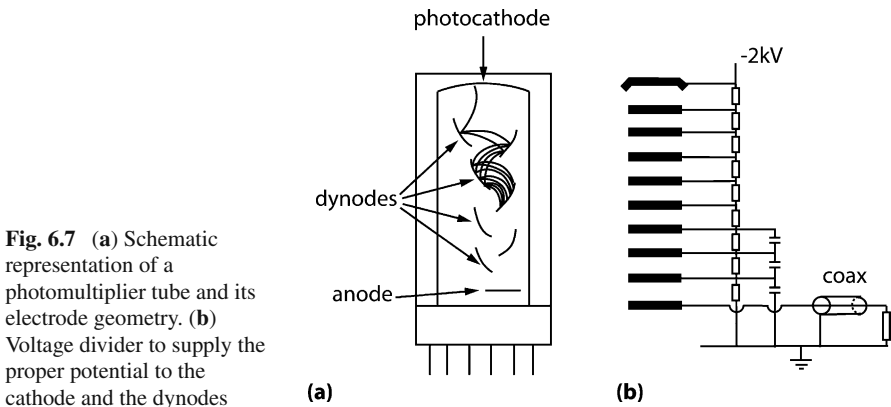

Fig. 6.7 (**a**) Schematic representation of a photomultiplier tube and its electrode geometry. (**b**) Voltage divider to supply the proper potential to the cathode and the dynodes

photocathode is deposited on the inside of the window of the photomultiplier tube. The photocathode is at some large negative potential, typically –2000 V. The dynodes are kept at potentials between –2000 V and 0 V, decreasing from –2000 V to 0 V in steps of typically 150 V. These voltages are nearly always obtained with a passive resistor chain as shown in Fig. 6.7(b). The last electrode, called the anode, is at ground potential. Sometimes one uses a different electrical layout with the cathode at ground potential and the anode at +2000 V. If a photon hits the photocathode, an electron is ejected from its surface. This electron will be attracted by the first dynode because this dynode is at a positive potential relative to the cathode and it will reach the surface of this dynode with a kinetic energy equal to the potential difference between the cathode and this first dynode. This kinetic energy is sufficient to extract several new electrons from this dynode. These few electrons, in turn, will be attracted by the next dynode and there again the number is multiplied. Of course, this multiplication process is efficient only if the shape of the dynodes is carefully optimised to provide efficient collection of the secondary electrons produced. The shape shown in Fig. 6.7 is only for illustration purposes and would not make an efficient photomultiplier tube. Figure 6.11 shows some more realistic dynode geometries. The electron multiplication step can be repeated many times, resulting in a very large total multiplication of the number of electrons, reaching 10^6 or more. At some point the total number of electrons in the cloud is so large that the electric field is strongly affected by the corresponding space charge and the structure no longer multiplies. The number of electrons in each pulse can reach up to 10^{10} before serious problems due to space charge occur. The electrical output signal is taken from the anode, where all the electrons eventually arrive. The total transit time of an electron from the cathode to the anode is typically 20 ns, depending on the size of the tube. If the tube is optimised for timing, the complete electron cloud caused by one photoelectron will arrive on the cathode in a time interval of about 10 ns.

The first important step affecting the performance by the photomultiplier tube is the conversion of the photon into a photoelectron by the photoelectric effect in the cathode. The probability for a photon to give rise to a photoelectron is called the quantum efficiency.

Many materials display the photoelectric effect for UV photons, and usually the corresponding quantum efficiency is small. However, some semiconductor materials have a large quantum efficiency for photoemission in the visible region. A thin (a few 10 nm) layer of such a material is deposited by vacuum evaporation on the inside face of the photomultiplier tube window. The quantum efficiencies that can be obtained are shown in Fig. 6.8. The largest quantum efficiencies are obtained with K_2CsSb (bialkali photo-cathodes) and are 25−30% around 400 nm. Unfortunately, this high efficiency is obtained by choosing materials that very easily emit electrons from the conduction band. As a result, electrons will also be emitted in the absence of any illumination. These materials produce between 10^2 and 10^4 thermal electrons per square centimetre at room temperature in the absence of any light stimulation. This dark current increases rapidly with temperature and can be a serious problem in certain applications. It should be mentioned that optical quality glass has a cut-off around 300 nm. Fused silica windows have a good light transmission down to 180 nm, but such PMTs are much more expensive.
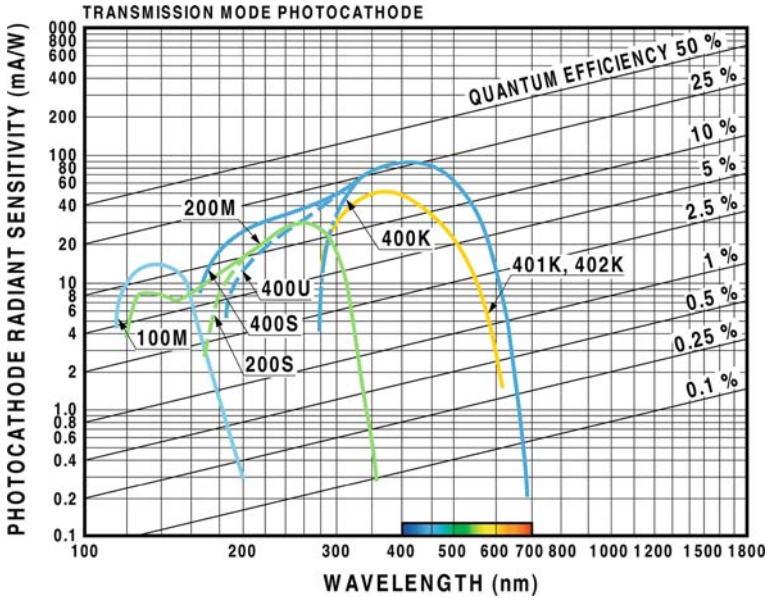
Fig. 6.8 Quantum efficiencies for a few photocathode materials as a function of the photon wavelength. Figure by courtesy of Hamamatsu

The probability for producing a secondary electron on the dynode also depends strongly on the nature of the surface of this dynode. Gallium phosphide (GaP) heavily doped with p-type material such as zinc is particularly effective. Up to 25 secondary electrons can be produced by an electron with a kinetic energy of 200 eV. This number of secondary electrons is more or less Poisson distributed. If the multiplication factor on one dynode is d and assuming all dynodes to have the same multiplication factor, the total gain is d^N, where N is the number of dynodes. The relative variance on the number of electrons produced (n_e) by one primary photo-electron can be shown to be given by $(\sigma/n_e)^2 = 1/(d-1)$ [11]. If a photomultiplier tube is illuminated with a weak light signal corresponding to only a few photo-electrons, a pulse height spectrum as shown in Fig. 6.9 is obtained. The first peak corresponds to events with one photoelectron, the second peak to events with two
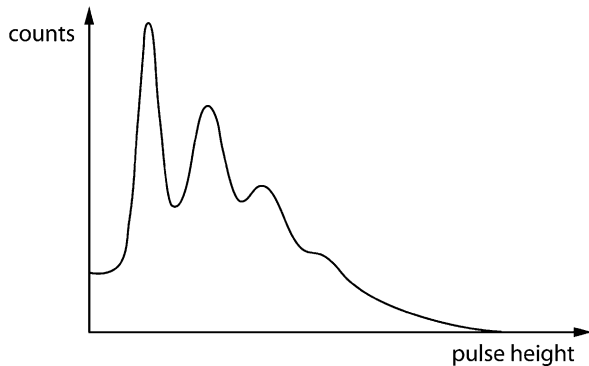
Fig. 6.9 Pulse height spectrum obtained with a photomultiplier tube where the gain of the first dynode is large. In this example, the tube is illuminated with light pulses in such a way that the average number of photoelectrons per pulse is between one and two

Fig. 6.10 (**a**) Collection of single channel photomultiplier tubes of various diameter. The largest tube in this image has a diameter of 20 cm. (**b**) Multi-anode PMT from Hamamatsu with 8×8 independent small PMTs in one single vacuum enclosure. The tube measures 25×25 mm^2

photoelectrons, etc. The peak corresponding to four photoelectrons is just visible.

Photomultiplier tubes come in many shapes and sizes. Figures 6.10 and 6.11 show a number of photomultiplier tubes and a few typical dynode geometries.
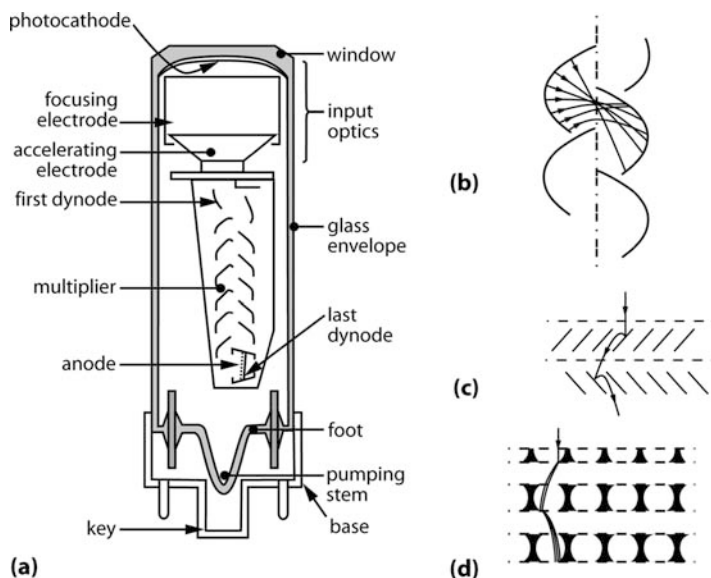

Fig. 6.11 (**a**) Internal structure of the 56AVP photomultiplier tube. To the left some other commonly used dynode structures in photomultiplier tube are shown: (**b**) linear focussing, (**c**) Venetian blind, (**d**) foils

In some applications, position information is highly desirable; therefore, position-sensitive photomultiplier tubes were developed. These are essentially of two types: multi-anode PMTs and true position-sensitive PMTs. In a multi-anode PMT, there are simply a large number of identical and independent amplifying structures, forming a collection of independent small PMTs inside one single vacuum enclosure, each PMT with its own readout anode, see Fig. 6.10(b). In a position-sensitive PMT, there are fewer anodes and the position is derived from a suitable ratio on the amplitudes of these anodes.

The photomultiplier tube is a very efficient instrument for observing and measuring weak and fast light pulses such as the light pulses caused by ionising radiation in a scintillator. However, there are a number of caveats when working with PMTs:

– A PMT has an important dark current. In addition to the thermal dark current consisting of single electron pulses, there is also after-pulsing, i.e. pulses coming a fixed time after a true pulse. These after-pulses are caused by residual gas atoms in the tube. After being ionised these atoms can drift back towards the cathode and extract a large number of electrons at the same time. Unlike the thermal dark current pulses that are single electrons pulses, these after-pulses correspond to many primary photoelectrons and can easily be confused with true signal pulses. This problem tends to increase as the tube ages but is also present in new tubes.
– PMTs, and particularly the larger ones, are extremely sensitive to magnetic fields. Even the Earth's magnetic field is sufficient to seriously affect the operation of a large photomultiplier tube. Because of this, photomultiplier tubes are usually surrounded by μ-metal cylinders working as a shield against magnetic fields. These shields are sufficient to protect against the Earth's magnetic field, but not against larger magnetic fields.
– A PMT should never be exposed to daylight when under high voltage. Also when not under high voltage, the photocathode should never be exposed to very intense light such as direct sunlight.
– When used at large gain and with fast pulses, the response of a photomultiplier tube becomes non-linear. This non-linear behaviour can have two causes: (a) too much current is drawn from some of the dynodes and their potential changes and (b) space charge effects. The potential of the dynodes can be stabilised by suitable design of the voltage divider. In Fig. 6.7(b), there are capacitors between the last three dynodes that will stabilise the potentials of these dynode to minimise this effect.

Silicon-based photodetectors. Another photodetector widely used in many applications is the silicon photodiode. In this detector, light is converted into electron–hole pairs in the depletion region. The mean free path of optical photons in silicon varies from 0.1 μm at 400 nm to 5 μm at 700 nm. The light must be able to reach the depletion region; hence, one of the electrodes, usually the p side, has to be as thin as possible and transparent to light. Silicon diodes are not expensive if the detector area is small, are insensitive to magnetic fields and have excellent quantum efficiency: around 60% at 400 nm, increasing to 80% at 800 nm. However, silicon

diodes have no internal gain, and when used for the readout of scintillators give signals too small for most applications. Figure 6.12(a) shows the internal structure of a PIN photodiode.

An avalanche photodiode (APD) is a silicon photodiode with internal gain. To make an avalanche photodiode one must change the doping profile in a diode as illustrated in Fig. 6.12. In a reach-through type APD we first have a low field region, where the photons convert into electron–hole pairs, followed by a high field region, where the field is sufficient to cause electron multiplication. The drawback of this layout is that the large dark current in the drift region is also multiplied. If we make the drift region very thin, the dark current is reduced but the detector capacitance is increased. This is solved in the reverse type APD, where we have a drift region behind the application region to decrease the capacitance.

The structure of an APD is conceptually quite straightforward but, in practice, it is very difficult to realise. The reason is that for stable avalanche multiplication, it is essential that only one type of charges is multiplied, the other type being merely collected. In silicon, the difference between the mobility of electrons and holes is small and the field necessary to start electron multiplication is very close to the field where hole multiplication starts. In order to make a good APD, it is essential to be
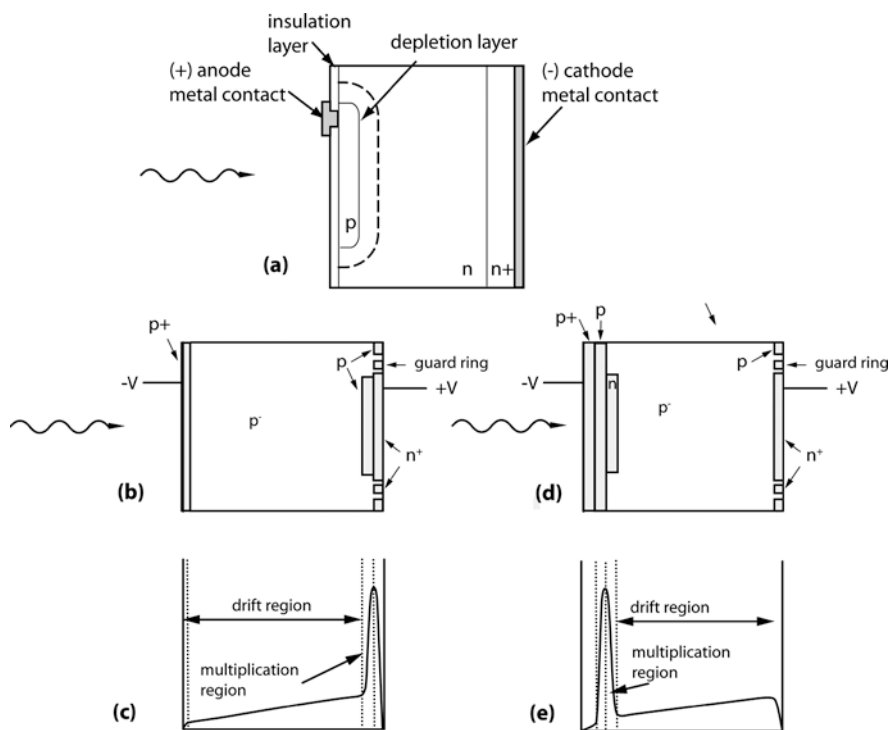

Fig. 6.12 (**a**) Structure of a PIN silicon photodiode. (**b**)+(**c**) Reach-through avalanche photodiode. (**d**)+(**e**) Reverse type avalanche photodiode. Figures (**b**) through (**e**) adapted from [3]
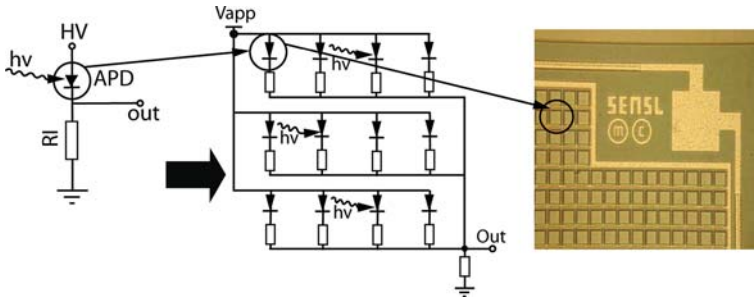
Fig. 6.13 Principle of operation of a silicon PMT. Each square cell in the figure on the left corresponds to a small diode and its quenching circuit. Figure courtesy of SENSL

able to control the fields, and hence the purity of the material and the doping profiles, extremely well. Manufacturing techniques allowing this have only recently become available. In practice, APDs can only be used at a gain of the order of 100 and can only be made in small sizes. Moreover, APDs tend to be expensive, certainly if a large photosensitive surface is required.

A very interesting recent development is the so-called silicon PMT (SiPMT), also called 'pixellised photodetector' (PPD), illustrated in Fig. 6.13. This device makes it possible to obtain a gain of 10^6 with silicon photodiodes. The idea is to subdivide the sensitive area of the photodetector in a large number of very small micro-pixels, typically measuring $50 \times 50 \ \mu m^2$. Each pixel works as a separate avalanche photodiode, but unlike in a normal APD, the electric field over the amplification region is made very large. As a result, the pixel-counter goes into discharge mode as soon as one electron–hole pair is formed.

A resistor between the micro-pixel and the power supply quenches the discharge. The separation between the individual pixels must guarantee that the discharge does not spread to the neighbouring pixels. The complete detector response remains linear with the light signal as long as only a small fraction of micro-pixels produce a signal. The quantum efficiency of a silicon PMT will certainly be less than what can be achieved with a PIN diode or an APD because of the unavoidable dead area between the micro-pixels. If the cross-talk between the pixels remains low, this type of detector will have an excess noise factor close to one. These devices are still under active development, but it seems likely that pure solid-state photodetectors with a gain of 10^6 and with a quantum efficiency similar to that which can be achieved with a photomultiplier tube will soon be available commercially.

The hybrid photomultiplier tube. Another interesting recent development is the hybrid photomultiplier tube. This photodetector owes its name to the fact that it combines ideas borrowed from photomultiplier tubes with silicon photodetectors. The hybrid photomultiplier tube is illustrated in Fig. 6.14. It consists of a vacuum tube with a photocathode. In the vacuum, facing the photocathode there is a silicon diode with a geometry very similar to that of a PIN diode. Between the cathode and the diode a very large voltage difference is applied. Any photoelectron produced
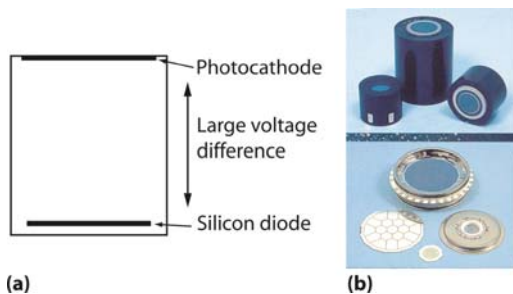
at the photocathode will be accelerated towards the diode and reach it with a large
kinetic energy equal to the voltage difference. This energetic electron will create a
number of electron–hole pairs equal to the kinetic energy divided by 3.62 eV. This
kind of photodetector is good for determining the number of photoelectrons in any
given pulse. Another advantage of this structure compared to a normal photomulti-
plier tube is that the silicon diode can easily be divided into a large number of pixels
of arbitrary shape.

The drawback of hybrid photodetectors is that a large gain can only be achieved
by using a very high voltage. To reach a gain of 10,000 it is necessary to apply
36,200 V. That is possible, but it is a significant technical complication and it makes
the system both complicated to use and expensive.

The excess noise factor. In a photodetector with internal gain, the interactions of
optical photons produce charges, and for each charge, the internal gain mechanism
multiplies the number of charges. The number of primary charges in a photodetector
is well described by the Poisson distribution law. However, all the charges are not
multiplied by the same gain factor and this is an additional source of fluctuations
in the output signal. The total signal S is given by the sum of a random number of
terms and each term is itself a random variable

$$S = \sum_i^{1...n} x_i$$

In this expression, x_i is a random variable describing the charge signal produced
by one primary charge and the integer n is a random variable with a Poisson distri-
bution of average value N, describing the number of primary charges. The quantity
$<x>$ is the internal gain factor of the photodetector. The following relations hold (see
Exercise 3, Chap. 8):

$$\langle S \rangle = N \langle x \rangle$$

$$\sigma^2 \{S\} = N \langle x^2 \rangle$$

This can also be written as

$$\sigma^2\{S\} = N\left(\sigma^2\{x\} + \langle x \rangle^2\right)$$
$$= N\langle x \rangle^2 \left[1 + \frac{\sigma^2\{x\}}{\langle x \rangle^2}\right]$$

We see that the variance of the output signal is exactly the variance that would be expected if the signal were only affected by Poisson fluctuations ($N\langle x \rangle^2$), multiplied by a factor independent of N. This factor is called the excess noise factor, and is usually denoted as F. It is unity if all the charges receive exactly the same multiplication. If the charge multiplication gives rise to an exponential distribution, $F = 2$. The noise equivalent number of photoelectrons, that is the number of photoelectrons that would give the same noise if all photoelectrons received the same gain, is given by N/F.

In comparing the merits of different photodetectors, the quantum efficiency divided by the excess noise factor is therefore the relevant quantity. The excess noise factor for photomultiplier tubes is in the range 1.2–2, depending mainly on the gain of the dynodes.

For APDs, the charge distribution of an avalanche produced by one primary charge is exponential at low gain. This is similar to the avalanche produced in a wire chamber. The excess noise factor is therefore ≈2 at moderate gain of less than 100. At higher gain the excess noise factor increases considerably. This is connected to the fact that in APDs both the electrons and the holes can contribute to the charge multiplication. SiPMs are still very new, but are expected to have a small excess noise factor. For hybrid photodetectors $F \approx 1$.

6.5 Using Scintillators in the Nuclear Energy Range

The most important application of inorganic scintillators is as detectors for X-rays or gamma rays. The present section is devoted to the discussion of a number of issues related to the use of scintillators in the nuclear energy range.

If one wants to check for the presence of some radioactive material, observing the characteristic gamma emission is often a good way of doing so. For this application, it is desirable to have a large detection efficiency for the gamma rays, accurate determination of the energy of the gamma rays and the ability to identify gamma rays, also in the presence of many other gamma rays of similar energy. In other words, one needs to measure the best possible energy spectrum of the gamma radiation present.

All scintillators in Table 6.2 have a radiation length between 0.9 cm (e.g. BGO) and 2.6 cm (e.g. NaI). For gamma rays of more than 1 MeV, the mean free path is of the order of the radiation length. The piece of scintillator should hence be a few radiation lengths thick in order to have good detection efficiency. At energies much above 1 MeV, the scintillator should be much larger and this case is discussed in Sect. 6.6.

A gamma ray detector consists of a piece of scintillator material, a photodetector, usually a photomultiplier tube, and readout electronics. The readout system will register a pulse height spectrum and if all the gamma rays deposit all their energy in the scintillator, we will see a peak at the corresponding position. To collect as much light as possible, the scintillator is wrapped in some white reflecting material and a photodetector is pressed against one of its sides. Between the photodetector and the scintillator, it is usual to add a thin layer of transparent grease with a refractive index close to the refractive index of glass. This grease will make optical contact between the scintillator and the photomultiplier window. This considerably improves the collection of the scintillation light because it avoids the internal reflection in the scintillator side facing the photodetector. For a good light collection efficiency the sensitive area of the photodetector should be sufficiently large. If the scintillator is a cube, and if the photodetector covers one complete face of the cube, the light collection efficiency will typically be ≈50%. Figure 6.15 shows some commercial PMT–scintillator assemblies.

Assume that we have a piece of LSO scintillator measuring $5 \times 5 \times 5$ cm^3, with one face covered with a 2-inch photomultiplier tube with a photocathode quantum efficiency of 25%. For a 1 MeV gamma ray completely absorbed in the scintillator we expect 3750 photoelectrons in the pulse. The photomultiplier tube behaves like a current source with a capacitance in parallel (see Fig. 6.16). The capacitance shown in this figure represents the capacitance between anode and ground and this capacitance is typically of the order of 10 pF. The product $C_a.Z_i$ gives the time constant of the readout and this time constant should be short compared to the decay time of the scintillator. In our example, this gives the condition $Z_i < 1000$ Ω. Often the signal is read directly with a 50 Ω cable. Assuming PMT gain is 10^5, the output pulse will peak at 75 mV (see Exercise 1). This pulse amplitude is large enough to be comfortably visible above the noise in any modern electronic readout system.

Fig. 6.15 Figure showing a few commercial scintillator–PMT assemblies. Photograph courtesy of Saint-Gobain crystals

Fig. 6.16 Schematic representation of the readout of a photomultiplier tube. The PMT anode behaves as a current source with a capacitance C_a. The readout amplifier has an input impedance Z_i
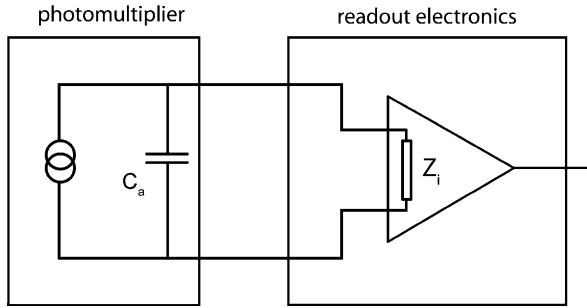
To be able to determine the energy of each peak accurately and to be able to separate nearby peaks, it is essential that the width of each peak is as narrow as possible. Obviously this requires a light yield of the scintillator, and a light collection efficiency, that are homogeneous over the volume of the scintillator. Assuming a Poisson distribution with average N_e for the number of photoelectrons, the relative energy resolution FWHM of the scintillator will be given by

$$\text{resolution   FWHM}[\%] = R_{\text{lightyield}} = \frac{100 \times 2.35}{\sqrt{N_e}} \tag{6.1}$$

In our example of an LSO scintillator, we therefore expect an energy resolution of 3.8%.

However, also under optimal conditions, it turns out that the energy resolution is much worse than Eq. (6.1) predicts. This is mainly due to the non-linear response of the scintillator to the energy deposited. The resolution of a scintillator is well described by the quadratic sum of an intrinsic, energy-independent term and a light yield dependent term given by Eq. (6.1).

$$R^2 = R_{\text{intrinsic}}^2 + R_{\text{lightyield}}^2$$

Table 6.3 gives the intrinsic energy resolutions for a few scintillator materials. At energies much above a few MeV, this equation no longer holds and this case will be considered in Sect. 6.6.

Table 6.3   Intrinsic resolutions for some scintillator materials

| Scintillator material	Intrinsic resolution $R_{\text{intrinsic}}$
BGO	$4.2 \pm 0.6\%$
CsI:Tl	$5.9 \pm 0.3\%$
LSO:Ce	$6.6 \pm 0.4\%$
YAP:Ce	$1.0 \pm 1.0\%$
LaBr:Ce	$<1.0\%$

Fig. 6.17 (**a**) Pulse height
spectrum recorded with a
LSO crystal measuring $3 \times 3 \times 10$ mm^3 and a ^{137}Cs
source. This source emits
gamma rays of 662 keV. The
photopeak, the Compton band
and the backscatter peak are
clearly visible on this plot.
The spectrum is cut at
≈ 80 keV by the trigger
threshold of the recording
electronics; (**b**)+(**c**)
explanation of the different
structures visible in the pulse
height spectrum (**a**)
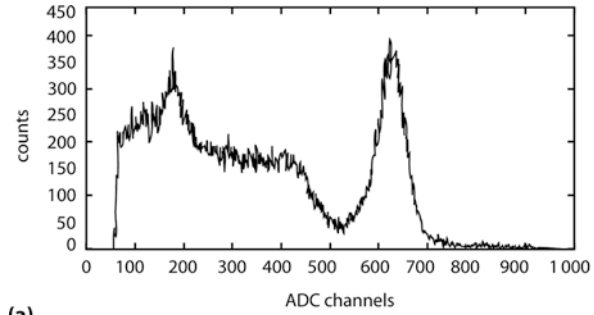
(a)
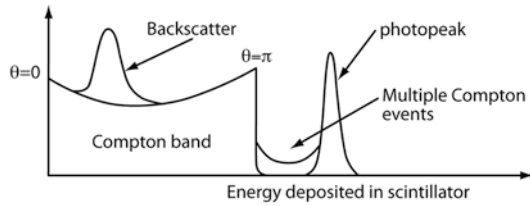
(b)
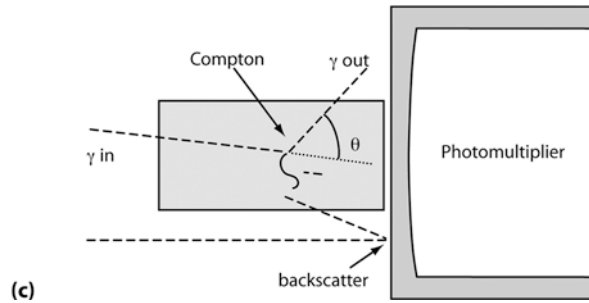
(c)

If we irradiate a scintillator with a gamma ray beam with all gamma rays of the
same energy, we not only obtain a peak corresponding to this energy, but also several
other bands and peaks as shown in Fig. 6.17. In addition to the photopeak, i.e. the
peak corresponding to events where all the energy is deposited in the scintillator,
we also have a band corresponding to events where the gamma ray has undergone
Compton scattering in the scintillator and where the scattered gamma ray escapes
from the crystal. The shape of the Compton band can be obtained from Eqs. (2.10)
and (2.11). The energy left in the scintillator is the energy of the recoil electron in
the Compton interaction. This is simply the initial gamma energy minus the energy
of the scattered photon. The maximum energy in the Compton band is obtained
using Eq. (2.10), taking $\theta = 180°$

$$\text{maximum energy compton band} = \hbar\omega - \hbar\omega'(\theta = 180°) = \hbar\omega \frac{2\hbar\omega}{2\hbar\omega + m_e c^2}$$

In Fig. 6.17, we also notice a peak at the energy of about 180 keV. The explanation of this peak is as follows. When this histogram was recorded, a ^{137}Cs source was used, and the source was next to the scintillator, on the opposite side of the PMT. Some of the 662 keV gamma rays emitted by the caesium source undergo Compton backscattering in the window of the PMT and the scintillator records the backscattered gamma rays. One can check that the value of this peak corresponds to the energy of backscattered gamma rays at 662 keV (184 keV).

Another spurious peak commonly encountered in gamma detection is the 'escape peak', see Fig. 6.18. The escape peak is caused by the following phenomenon. If a gamma ray interacts in a scintillator by photoelectric effect, the most probable electron to be involved in the interaction is the most deeply bound or K-shell electron. An outer electron quickly fills the vacancy thus left in the electron structure of the atom and the corresponding energy is emitted as an X-ray with an energy equal to the binding energy of this K-shell electron. Most of the time the scintillator immediately absorbs this X-ray, but sometimes, it can escape from the crystal. In the latter case, the total energy deposited in the crystal is the energy of the gamma ray minus the energy of the X-ray. These events will show up in the pulse height spectrum as second peak below the photoelectric peak.

Finally, gamma rays interacting in any other material present near the measurement setup can cause additional spurious peaks. If the primary gamma rays have a larger energy than 1.022 MeV, these gamma rays will create electron–positron pairs anywhere in the material surrounding the scintillator detector. When the positron annihilates, it gives rise to two gamma rays of 511 keV and these can also give rise to a peak in the pulse height spectrum. X-rays can also result from nuclear interactions of the gamma rays in the surrounding material.

All these spurious peaks can make the interpretation of a gamma ray pulse height spectrum quite complicated, certainly if gamma rays from several energies

Fig. 6.18 Escape peak: a gamma ray interacts in the scintillator and ejects a K-shell electron from one of the atoms. If the resulting K-shell X-ray escapes from the scintillator, this gives rise to a peak corresponding to the total energy minus the energy of the X-ray. In NaI, the energy of these X-rays is 49.1 keV
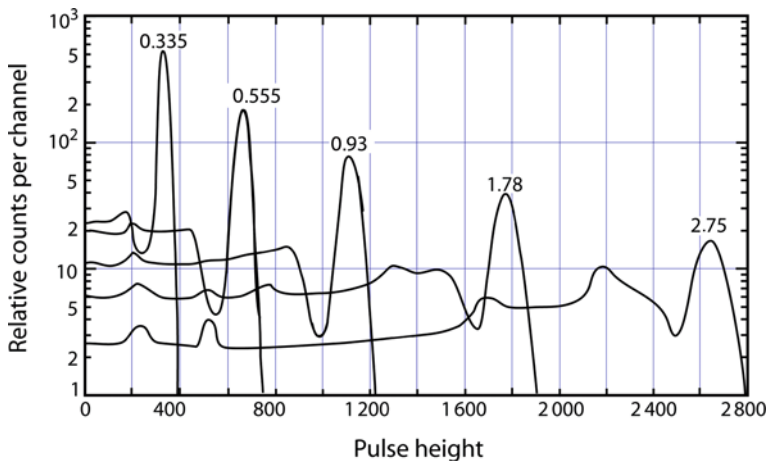
Fig. 6.19 Response of a cylindrical NaI:Tl crystal measuring 7.62 × 7.62 cm exposed to mono-energetic gamma rays of different energies between 335 keV and 2.75 MeV. Figure adapted from [4]

are present at the same time. This is illustrated in Fig. 6.19, which shows the pulse height spectrum for mono-energetic gamma rays in a NaI:Tl scintillator. An obvious way to avoid, or at least strongly suppress, most of these spurious peaks and bands is to use a well counter, i.e. a setup where the gamma source is completely surrounded by the scintillator. However, it is not always possible to put the source in a well counter, for example because the source is too large, its location is unknown, etc.

This section ends with a comparison of different methods commonly used for X-ray and gamma-ray detection. Proportional tubes with a gas filling containing mainly argon can be used to detect X-rays, but are limited to energies below 20 keV. With a gas filling containing mainly xenon, a reasonable detection efficiency can be obtained up to an energy of 100 keV. However, also for these low-energy X-rays it will take several centimetres of gas to have the same stopping power that can be achieved with only 100 μm of scintillator or germanium detector. At energies above 100 keV, only germanium detectors or scintillators have a sufficiently large stopping power to be useful for detecting gamma rays.

Comparing the energy resolution obtainable with scintillators with the energy resolution of germanium scintillators discussed in Chap. 5, we see that the energy resolution of germanium is at least a factor 10 better than even the best scintillator. This is illustrated in Fig. 6.20 showing the same gamma radiation field observed with a NaI:Tl scintillator and with a germanium detector. It may therefore seem that germanium detectors will nearly always be preferred over scintillators for gamma detection, but this is not at all the case. Germanium detectors are expensive and
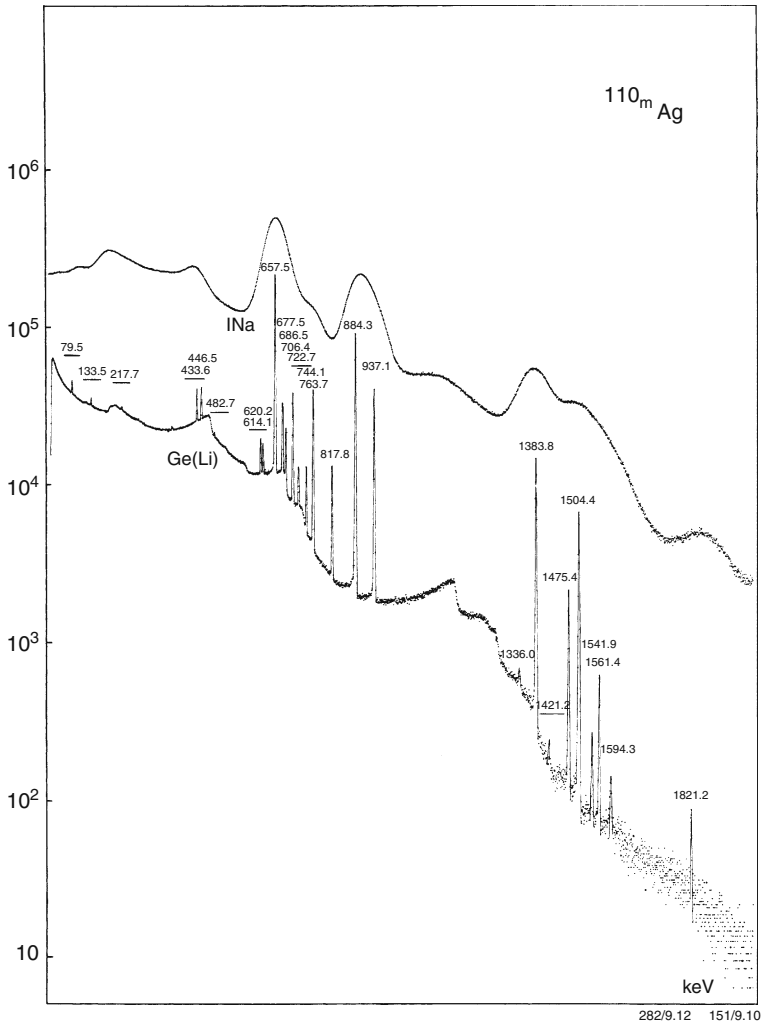
Fig. 6.20 Gamma radiation pulse height spectrum of ^{110m}Ag recorded with a germanium detector and with a NaI:Tl scintillator. Energies of the peaks are labelled in keV. Figure from [5], ©1970 IEEE

need to be cooled to the temperature of liquid nitrogen, which is a major complication. Moreover, germanium detectors are relatively slow since the signal formation requires the electrons to drift over the full length of the collection gap. Germanium detectors, therefore, cannot be used if either good time determination or large count rate is needed.

## 6.6 Applications of Scintillators in High-Energy Physics

An important type of detector for particles of high energy is the calorimeter. This type of detector measures the energy of a particle by totally absorbing the shower produced by this particle in a block of material. The energy of the particle is proportional to the amount of ionisation produced in the material. Calorimeters often use scintillators to measure the amount of ionisation. For an in-depth discussion of calorimeters, we refer the reader to [6, 7].

Electromagnetic calorimeters. Electromagnetic calorimeters are detectors for measuring the energy and the position of high-energy gamma rays or high-energy electrons and positrons. For the purpose of this discussion, high-energy means energy larger than 1 GeV. A gamma ray with this energy interacting in matter will initiate an avalanche consisting of a large number of electrons, positrons and secondary gamma rays. On average in about one radiation length the original gamma ray gives rise to an electron–positron pair. This electron and positron will create a large number of secondary gamma rays by bremsstrahlung. In one radiation length, an electron or a positron will radiate about half of its energy in this way. Many of these secondary gamma rays will again create electron–positrons pairs and these will again undergo bremsstrahlung, and so on. If the energy of the initial gamma ray is large enough, the number of particles in the shower will grow exponentially. However, at each step the average energy of the particles in the shower decreases and fewer of the secondary gamma rays have sufficient energy to produce electron–positron pairs. Moreover, below the critical energy, the electrons and positrons will lose more energy in ionisation than in bremsstrahlung and the production of additional gamma rays becomes less efficient. After a certain number of radiation lengths, the number of particles in the shower reaches a maximum and thereafter starts decreasing. Eventually, all electrons and positrons will have lost all their energy in ionisation and come to rest and all the gamma rays will be absorbed by photoelectric effect. The positrons will annihilate with electrons in two 511 keV gamma rays and these will also be absorbed.

The development of a shower is a complex process and only a detailed Monte Carlo simulation can provide a reliable quantitative description. Nevertheless, the following simple model gives a reasonable qualitative description of the shower development. The model assumes that in one radiation length a gamma ray will convert into one electron–positron pair and ignores the fluctuations on the conversion distance. The model furthermore assumes that in one radiation length an electron or a positron will emit one gamma ray with more energy than two times the electron mass. The model ignores the statistical fluctuations on the number of gamma rays and also ignores that the average number of such gamma rays depends on the energy of the electrons. This model of shower development is illustrated in Fig. 6.21(a). In this model, the total number of particles, i.e. the number of electrons, positrons and gamma rays with $E > 2m_e$, will increase with the depth in the material D as

$$N = 2^{D/X_0}$$

Fig. 6.21 (**a**) A very simple model for the development of an electromagnetic shower. (**b**) Typical geometry of a detector element in an electromagnetic calorimeter using inorganic scintillators
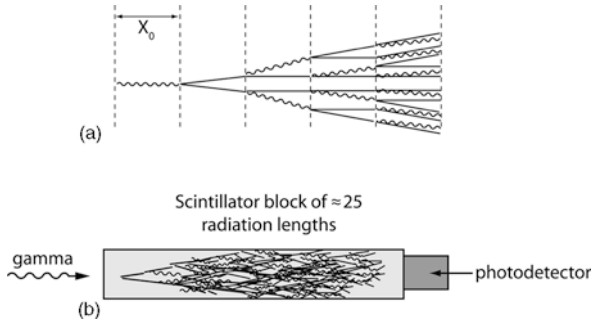
The number of particles in the shower will increase exponentially until the average energy of the electrons becomes equal to the critical energy. At this point the shower has the maximum number of particles. We thus have

$$N_{\max} = \frac{E}{E_c} = 2^{D_{\max}/X_0}$$

$$D_{\max} = \frac{X_0}{\ln 2} \ln \frac{E}{E_c}$$

After this maximum, the number of particles in the shower decreases exponentially and the attenuation length of the number of particles is again of the order of the radiation length. Since the number of the particles in the shower increases exponentially with the energy of the incoming gamma ray, the length needed to fully absorb all the particles will also increase like $\ln(E/E_c)$. Experimentally, it is found that for a gamma ray of 10 GeV the number of particles in the shower reaches a maximum after about six radiation lengths and it takes a total of ≈25 radiation lengths of heavy material to absorb 99% of the shower energy.

The lateral, or sideways, development of the shower is mainly due to the multiple scattering of the electrons and positrons and scales with the 'Molière radius'. The quantity Molière radius has the dimension of a length and it is characteristic for the medium in which the shower develops. Its value is close to [14 g/cm^2]/density for most materials. In electromagnetic showers, 90 and 95% of the energy is deposited within one and two Molière radii, respectively.

If a gamma ray enters a sufficiently large block of scintillator material, all the energy of the initial particle is deposited as ionisation in the material. The total amount of ionisation is therefore proportional to the energy of the particle and the amount of scintillation light produced will also be proportional to the energy of the initial gamma ray. It should be mentioned that a high-energy electron or a high-energy positron will initiate a shower that looks exactly the same as a shower initiated by a high-energy gamma ray. The presence or absence of the incoming charged particle track at the starting point of the shower is the only difference between an electron-initiated shower and a gamma-initiated shower.

An electromagnetic calorimeter using inorganic scintillators typically has a large number of blocks of inorganic scintillator material with a geometry similar to what is shown in Fig. 6.21(b). The length of each scintillator block is typically equal to 25 radiation lengths. The width of each block is chosen to be less than the lateral extent of the shower to allow accurate determination of the centre of the shower, and therefore the impact point of the gamma ray.

The energy resolution attainable with an electromagnetic calorimeter is usually expressed as r.m.s. energy resolution, and its energy dependence can be parameterised as

$$\frac{\sigma\{E\}}{E} = \sqrt{\frac{a^2}{E[\text{GeV}]} + b^2} \tag{6.2}$$

In this equation, $E[\text{GeV}]$ represents the energy of the initial gamma ray expressed in GeV. For a homogeneous crystal calorimeter the coefficients 'a' and 'b' are typically 0.02–0.03 and 0.005–0.01, respectively. The first term is the statistical term. One of the effects entering here is the fluctuation on the number of optical photons detected. The energy resolution due to the fluctuation on the number of detected photons is given by

$$\frac{\sigma\{E\}}{E} = \frac{1}{\sqrt{N_{p.e}}} = \sqrt{\frac{1/\varepsilon}{E[\text{GeV}]}}$$

where ε is the number of photoelectrons detected per GeV. However, there are several other effects contributing to the statistical term, e.g. leakage of a small fraction of the shower, photonuclear interactions in the shower, absorption of part of the shower in the dead material between two blocks of scintillating material.

The second term has to do with non-uniformities in the different components of the detector. Careful calibration is essential to keep this term small. To obtain the desired energy resolution, it is also essential that the signals from all the blocks containing parts of the same shower are added together.

Until now we have been considering electromagnetic calorimeters based on the use of large blocks of inorganic scintillating materials. Such detectors show excellent performance but are expensive. To reduce the cost one often uses sampling calorimeters. In a sampling calorimeter, different materials perform the function of absorbing the shower and the function of measuring the energy deposited. Such a sampling calorimeter typically is made from a large number of layers of some heavy material, usually lead, interleaved with active layers measuring the amount of ionisation present, often plastic scintillator sheets. In sampling calorimeters typically a few percent of the energy in the shower is actually sampled. As can be expected, the energy resolution that is obtained with a sampling calorimeter is significantly worse than what can be achieved with a homogeneous calorimeter. A crude estimate of the achievable energy resolution can be obtained as follows. Assume a sampling calorimeter where the thickness of the active layers is a small fraction of a radiation

length and the thickness of the absorbers is of the order of one radiation length. In such a sampling calorimeter, one is essentially counting the number of charged tracks in each sample layer. The simple model of shower development discussed before, suggest that the number of tracks reaches a sharp maximum at the point of the maximum of the shower. The error on the number of charged tracks is therefore dominated by the error on the number of charged tracks at the maximum of the shower $N_{\max} = E/E_c$. The error on the energy resolution of the sampling calorimeter is therefore approximately given by

$$\frac{\sigma\{E\}}{E} \approx \sqrt{\frac{1}{N_{\max}}} \approx \sqrt{\frac{E_c}{E}}$$

For electromagnetic calorimeters where a few percent of the shower is sampled, the statistical term in Eq. (6.2) is typically ≈ 0.1.

For sampling the number of charges in a sampling calorimeter any method for measuring the amount of ionisation can be used. Sampling calorimeters have been built with scintillators, with gas ionisation chambers and with silicon detectors. A popular type of electromagnetic calorimeter uses liquid ionisation chambers for sampling the ionisation. The basic principle of operation of a liquid-filled ionisation chamber is the same as for a gas-filled ionisation chamber. The ionising charged particles produce electron-ion pairs in the liquid and the charges drift towards the electrodes under the influence of the applied electric field. In principle, many liquids could be used for this purpose, but in practice it turns out extremely difficult to obtain a liquid that is sufficiently pure to allow efficient charge collection. Liquefied noble gases, such as argon, krypton and xenon can more easily be purified than other liquids, because the chemical properties of such atoms are very different from the chemical properties of the impurities and because these gases are used at cryogenic temperatures, where most impurities just freeze out. Among the noble gases, argon has by far the lowest cost and is therefore the preferred choice. Electron drift velocities in liquid noble gases are a few 10^5 cm/s at fields of interest, while positive ion velocities are only of the order of a few cm/s. Except if the detector were to be used at a rate of only a few Hz, the short signal integration time implies that only the electron signal will be seen.

As an example of a homogeneous electromagnetic calorimeter, I will briefly describe the electromagnetic calorimeter of the CMS detector. CMS is one of the very large detectors that were installed at the Large Hadron Collider of CERN in 2008. The main characteristics of the LHC accelerator were already presented in Sect. 3.5. Figure 6.22 shows a very schematic layout of the CMS detector. A detailed description of the CMS detector can be found in [8].

In the collision of high-energy protons, a large number of secondary particles is produced. The aim of the CMS detector is to observe as many of these particles as possible and determine their direction and their energy. The electromagnetic calorimeter of CMS is located inside a large cylindrical magnet just outside of the detector for charged particles (tracker). It is designed to measure the energy and the position of electrons, positrons and gamma rays. The layout of the device is
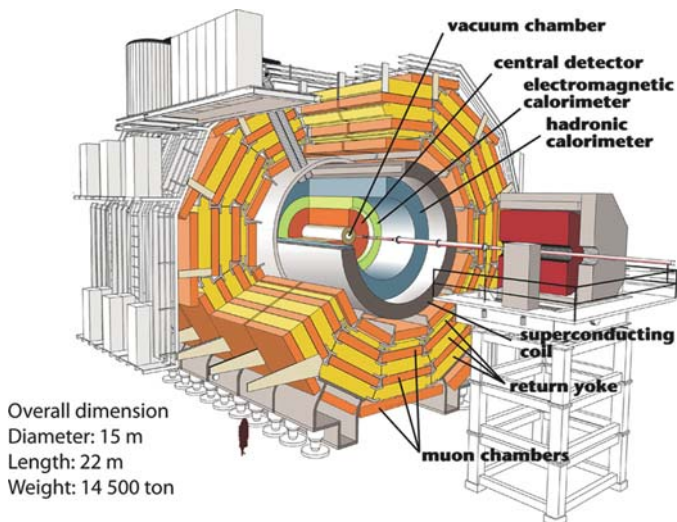
Fig. 6.22 Cut away view of the CMS detector. The beam pipe runs along the central axis of the apparatus. Starting from the centre the particles see the central detector that identifies charged particle tracks, the electromagnetic calorimeter, the hadronic calorimeter and the superconductive magnet coil. Outside the magnet coil is the muon detector. Figure copyright CERN

schematically represented in Fig. 6.23. The electromagnetic calorimeter uses blocks of lead tungstate scintillator (PWO) 230 mm long and measuring 22×22 mm^2 at the front side and 26×26 cm^2 at the backside. The scintillator blocks in the endcaps are slightly larger. A length of 230 mm of PWO corresponds to 26 radiation lengths. A total of $\approx 80,000$ such blocks of scintillator are arranged around
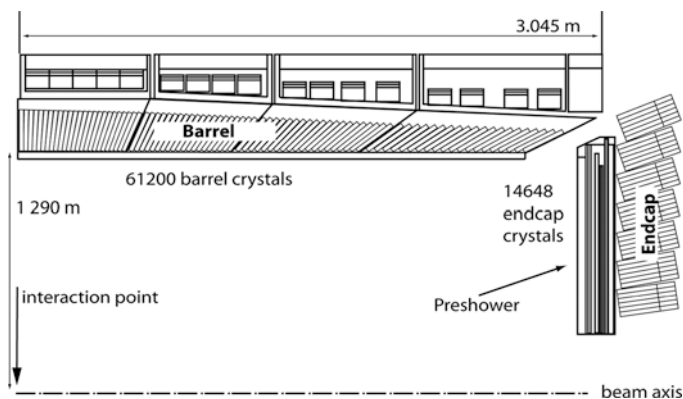

Fig. 6.23 Electromagnetic calorimeter of the CMS experiment. This figure shows one quarter of the detector. The beam–beam interaction point is in the lower left corner of the figure. The detector consists of $\approx 80,000$ blocks of BGO scintillator pointing towards the interaction point

the interaction point with the blocks always pointing with the long axis towards the interaction point. In fact, there is a small angle between the axis of the crystals and the line pointing towards the interaction point in order to prevent gamma rays from escaping detection by passing in the narrow gaps between adjacent crystals.

In CMS, PWO is chosen as a scintillator because of its short radiation length and short decay time. This scintillator has a rather low light yield, but at this high energy light yield is far less important than at lower energy. The scintillator blocks are long enough to fully contain the electromagnetic shower in the longitudinal direction, but the showers are wider than the blocks. Part of the shower will leak to the neighbouring blocks and it is necessary to sum the energy deposited in a 3×3 array of neighbouring blocks to get the correct value for total energy of the gamma ray. The spreading of the light over several blocks allows the determination the position of the gamma ray to a precision much better than the size of the blocks. In CMS, this accuracy is about 1 mm r.m.s. The light of the scintillator is read by avalanche photodiodes because the calorimeter is inside a strong magnetic field of 4 tesla, and this large field precludes the use of photomultiplier tubes. For the CMS electromagnetic calorimeter, the coefficients in the energy resolution formula (Eq. 6.2) are $a = 3\%$ and $b = 0.5\%$.

Hadronic calorimeters. For detecting neutral hadrons one uses a device somewhat similar to the electromagnetic calorimeter just discussed. A hadron interacting in a block of material will undergo strong interactions with the nuclei in the material and in the collision produce a number of secondary hadrons, mainly protons, neutrons and π-mesons. Positively charged pions, negatively charged pions and neutral pions are produced roughly in equal numbers. The neutral pions decay after a few micrometer into two gamma rays. The other secondary hadrons will again interact, producing more protons, neutrons and π-mesons and so on. The phenomenon is similar to the electromagnetic avalanche induced by a high-energy gamma ray, with a few important differences, though.

For all heavy materials, the hadronic interaction length is much longer than the radiation length. The concept of hadronic interaction length was introduced in Sect. 2.5. For iron, copper and tungsten, the hadronic interaction length is 16.8, 15 and 9.6 cm, respectively, to be compared with 1.76, 1.43 and 0.35 cm, respectively, for the radiation length. To fully absorb most of the hadronic shower produced by a particle of 100 GeV, about 8–10 hadronic interaction lengths are needed. A hadronic calorimeter made entirely of inorganic scintillator such as PWO would need a $\approx$1.5 metre thick layer of scintillator and would be prohibitively expensive. A more cost-effective solution must be used and the hadron calorimeters are always sampling calorimeters with sheets of plastic scintillators as active material. To extract the light from the scintillator, wavelength shifting rods or wavelength shifting fibres are universally used. This readout method avoids the limitation imposed by the Liouville theorem as discussed in Sect. 6.2. The energy resolution that can be obtained with a hadronic calorimeter is typically given by

$$\frac{\sigma\{E\}}{E} \approx \frac{0.6}{\sqrt{E[\text{GeV}]}}$$

Notice that the energy resolution of a hadronic calorimeter is much worse than the energy resolution of an electromagnetic calorimeter. There are several reasons for this. First, a hadronic calorimeter is always a sampling calorimeter. A very important additional cause of degradation of the performance of a hadronic calorimeter has to do with the different response of the hadron calorimeter to the hadronic part of the shower and to the electromagnetic part of the shower. The reason for this difference is that of the order of 50% of the energy of the hadrons goes into breaking up the nuclei and into the energy of nuclear fragments. The energy used for breaking up the nuclei is lost, and the kinetic energy of the nuclear fragments is converted very inefficiently into scintillation light, such that this energy is also largely lost. The excited nuclei decay with a time constant that is large compared to the integration time of the signal, and this energy is therefore also lost. In the first interaction of the shower typically one third of the energy goes into the creation of neutral pions and these decay instantly into gamma rays and in this way give rise to electromagnetic showers. In all the subsequent hadronic interactions, additional neutral pions are produced, therefore the electromagnetic fraction increases with energy. This fraction varies from $\approx$30% at 10 GeV to $\approx$60% at 1 TeV. Moreover, this electromagnetic fraction is subject to large fluctuations and together with the different response of the calorimeter to the two components in the shower this degrades the energy resolution. This, in fact, is the main effect limiting the energy resolution of hadron calorimeters. A significant improvement in the resolution of hadronic calorimeters will only be possible with a much more sophisticated design that corrects for these effects.

Cosmic air showers. A somewhat different use of the scintillation effect is the observation of cosmic air showers. If a very high-energy cosmic ray enters the atmosphere, it will initiate an avalanche of secondary particles. These particles produce light by Cherenkov effect and by scintillation of the nitrogen in the air. For energies of $\approx 10^{14}$ eV and above the scintillation light produced in such air showers is sufficient to observe the light on a moonless night with a clear sky. A detector for high-energy cosmic air showers essentially consists of a parabolic mirror and a suitable photodetector. The optical quality of the mirror does not need to be comparable of what is needed for optical astronomy. The only function of the mirrors is to collect more light.

6.7 Applications of Scintillators in Medicine

Scintillators in radiology. Inorganic scintillators are also used extensively in X-ray imaging and in nuclear medicine. Imaging the inside of the human body with X-rays, called radiology, is the oldest medical imaging technique. X-ray imaging is based on the different absorption of X-rays in different body tissues. Figure 6.24 shows the linear attenuation coefficient in water and bone. The absorption in the soft tissues of the body is similar to the absorption in water.

We see that below 20 keV the mean free path of X-rays is less than 1 cm and therefore no useful images can be made with such X-rays, except when imaging

Fig. 6.24 Linear attenuation coefficient for water, bone and CsI scintillator. The data for this figure were obtained from [9] in Chap. 1
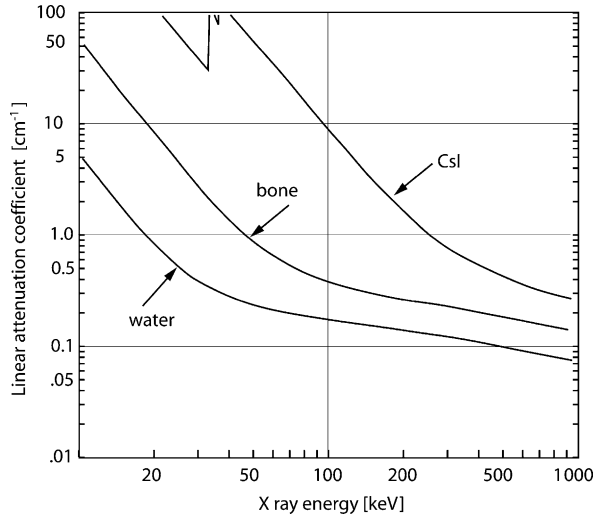
small organs such as teeth. As the energy increases, the body becomes more transparent to X-rays, but the contrast also decreases. The useful energy range for medical X-ray imaging is therefore ≈20–80 keV.

In Fig. 6.24, we also show the linear attenuation coefficient for the inorganic scintillator CsI. It can be seen that in the region of interest, ≈150 μm of scintillator will stop about half of all X-rays.

Many years ago X-ray images were made using the direct interaction of X-rays in ordinary photographic emulsions, but the detection efficiency of the photographic film for X-rays is low and results in a correspondingly larger dose to the patient. Today, this method is still used in material testing where the dose is less important. It allows a better spatial resolution in the image than any other image recording method.

Until recently, most radiological imaging was done with the film-screen method. In this technique, a layer of scintillating powder, typically 300 μm thick, is applied on a plastic support. Such a layer of scintillator is usually called a phosphor. A photographic emulsion is applied against the phosphor screen. This photographic emulsion simply records the scintillation light produced by the X-rays in the phosphor screen. After development, the film contains the image. The scintillator materials used in this application are often different from the scintillators discussed previously since speed is not important here. A scintillator often used for this purpose is Gd_2O_2S:Tb and 10 photons/keV X-ray energy will reach the photographic emulsion. The film and screen assembly used in the film-screen method is illustrated in Fig. 6.25.

The film-screen method is more and more being replaced by a technique marketed under the name 'computed radiography'. 'Storage phosphor screen' would be a more appropriate name. A storage phosphor screen looks similar to the screens

Fig. 6.25 Illustration of the film-screen method for X-ray imaging
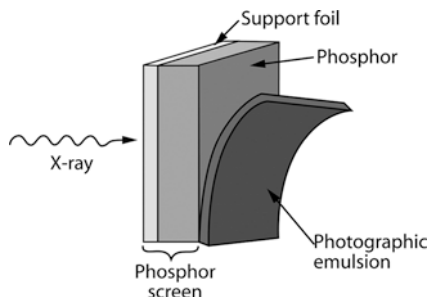
used in the film screen method, but it uses a different type of phosphor that has the capability of 'remembering' the X-ray image. The physical principle behind the image storing property is explained in Fig. 6.26.

The most commonly used scintillator in storage phosphor screens is BaFBr doped with europium 2+ ions. In this scintillator, the Eu^{2+} ion has 4f and 5d energy levels in the band gap of the host material. It acts as luminescence centre and emits at 390 nm with a decay time of ≈ 1 μs. By suitable preparation of the material, it is possible to create many vacancies of the negative ion (fluor or bromine) in the BaFBr host lattice. Such vacancies act as electron traps and are called F-centres. The binding energy of the electrons in these traps is typically in the optical region; therefore, the presence of such traps will tend to colour the material. The name F-centre derives from 'farbe-zentrum' the German name for 'colour centre'. This phenomenon causes many coloration effects in natural minerals.

In BaFBr, the binding energy of the electrons in the F-centre is ≈ 2 eV. If any X-ray interacts in the $BaFBr:Eu^{2+}$, it will create a large number of electron–hole pairs. However, most of these electrons will never reach the luminescence centre because they are trapped in the F-centre. At room temperature electrons can remain trapped almost indefinitely. If one later illuminates the phosphor plate with photons of ≈ 2 eV (≈ 630 nm), the electrons will be liberated from the F-centres, reach the Eu^{2+} luminescence centres and emit scintillation light at 390 nm.

If a point on the surface of the storage phosphor screen is stimulated with light at 630 nm, the amount of light emitted at 390 nm will be proportional to the number of
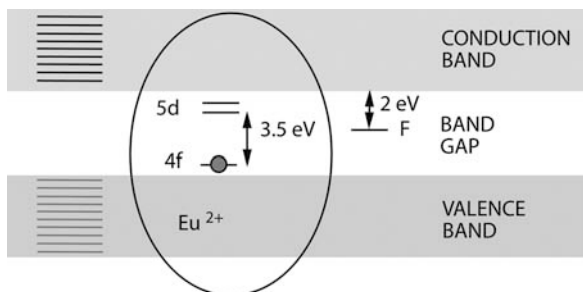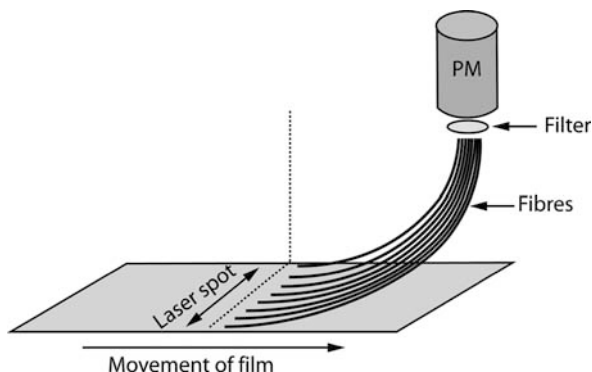

Fig. 6.26 Band structure in a BaFBr:Eu^{2+} storage phosphor screen. The F-centres are electron traps with a binding energy of about 2 eV

PM

Filter

Fibres

Laser spot

Movement of film

electrons trapped in F-centres at this point. This signal is proportional to the X-ray dose received at this point during the recording of the image.

The method to retrieve the stored image from a storage phosphor screen is illustrated in Fig. 6.27. A small light spot from a laser scans the surface of the X-ray storage phosphor image plate. An oscillating mirror causes the light spot to travel back and forth along a line. Optical fibres aligned along the line travelled by the light spot collect the photo-stimulated luminescence light emitted at 390 nm. These optical fibres guide the light to a photomultiplier tube. The stimulation light (630 nm) is stopped by means of an appropriate optical filter placed between the bundle of light guides and the photomultiplier. The scintillation light at 390 nm is not stopped by this filter and recorded by the photomultiplier tube. To scan the complete area of the plate, this plate is moving in a direction perpendicular to the laser line.

The main advantage of storage screens over the conventional film screen method is the larger dynamic range. In the film screen method, the image is recorded as variations in density (blackness), of the film. All irradiation levels above a certain value are black and can no longer be differentiated by the human eye, while all levels of irradiation below a certain level are seen as white and are also no longer differentiated. The dynamic range of this recording technique is limited to about a factor 100. With storage phosphor screens, the image is recorded electronically and a dynamic range of up to 10,000 is possible. When the image is printed on a film for inspection by a radiologist, the image processing software can adjust the density levels such that the information remains visible, also in the parts that would otherwise be too dark or too light.

It is expected that in the future the storage films will be replaced by a new technique called 'digital radiography' (DR). In DR, the X-ray image is recorded with the help of a phosphor layer in direct contact with a flat panel detector having a large number of small silicon photodetector pixels and with the circuitry allowing the readout of these pixels. The main difficulty in this technique is that for radiography it is necessary to have detectors that can record an image of 40×40 cm^2. The largest sizes of silicon monocrystals that are available have only 8 inch in diameter. The photodetectors pixels and readout circuitry must therefore be made in amorphous silicon. Amorphous silicon can be deposited by Chemical Vapour Deposition

technique over large areas. This technique is aggressively being developed for the production of low-cost solar cells. Such amorphous silicon is less suitable for making electronic components than monocrystalline silicon, but the technology has now been improved to the point where suitable large panels with silicon photodetectors and the associated readout circuitry can be made.

One of the limitations of this technique, and of all other techniques using phosphor screens to convert the X-rays to light, is the spreading of the light in the phosphor layer. The light from an infinitely narrow X-ray beam will spread by diffusion in the phosphor layer and form a light spot with a diameter of the order of the thickness of the phosphor layer. It is therefore necessary to make a compromise between the detection efficiency and the spatial resolution. A peculiar property of CsI:Tl scintillator allows to overcome this limitation. Indeed CsI can be vapour deposited in a way as to form a layer of small microcolumns or needles. These needles behave like light guides channelling the light to the photodetector, and in this way strongly reduce the lateral spread of the light. Figure 6.28 shows micrographs of such a CsI layer.

An even more ambitious road towards DR is the so-called direct conversion X-ray imaging. In this approach, one is not using a phosphor for converting the X-ray energy to light, but directly records the ionisation left in some suitable material. Silicon cannot be used for this, because the stopping power is too low. The main difficulty is the fact that one needs to cover a surface of the order of 40×40 cm^2. With monocrystals such as Ga, this surface must be made from smaller detectors, and one has to deal with artefacts at the edges between the detectors. Selenium can be deposited as amorphous selenium over large surfaces and seems the most promising material to realise this kind of X-ray imaging detectors.

All presently used detectors in radiology integrate the signal because the event rate is too large to allow counting individual X-rays. However, counting individual X-rays would have the major advantage that it would be possible to select X-rays
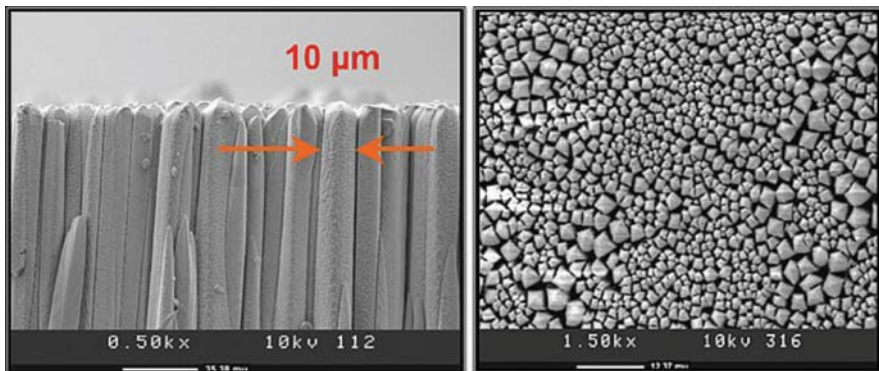

Fig. 6.28 Scanning Electron Microscope micrographs of microcolumnar CsI:Tl, Eu scintillator (cesium iodide with two added dopants, thallium and Europium) layer grown by vapour deposition. *Left figure*: seen from the side, *Right figure*: seen from above. Photographs courtesy of RMD [10]

in certain energy windows. Developing systems that count individual X-rays and therefore would really deserve the name 'digital' are presently an active research field.

   Scintillators in positron emission tomography. Another important medical application of scintillators is as gamma detectors in scintigraphy and in positron emission tomography (PET). Positron emission tomography is one of the powerful medical imaging techniques of nuclear medicine. In PET, a physiologically relevant compound is labelled with a positron-emitting isotope. The term 'labelled' means that a normally stable atom in the molecule is replaced by a radioactive isotope of the same element. The chemical properties of the labelled compound are identical to the properties of the natural one and this labelled compound will take part in all processes in the body exactly in the same way as the natural compound. At some point the radioactive isotope emits a positron and this positron has a kinetic energy of a few 100 keV. At this energy the range of the positron in the living tissue is usually less than 1 mm. After coming to rest, the positron annihilates into two nearly back-to-back gamma rays of 511 keV. The mean free path of gamma rays of 511 keV in the human body is about 10 cm. In many cases, the two gamma rays will leave the body without undergoing scattering, i.e. with their original direction unchanged. A PET scanner basically is a detector for gamma rays of 511 keV surrounding the patient. If two 511 keV gamma rays are detected at the same time, we assume that these come from the same annihilation event and we know that the annihilation and therefore the molecule containing the radioactive isotope was somewhere on the line joining the two detection points. From the observation of a large number of such annihilations it is possible to reconstruct the three-dimensional distribution in space of the annihilation events. That is the same as the three-dimensional distribution in space of the labelled molecules. PET is therefore a non-invasive technique that allows following the evolution of the labelled compound in the body. The value of PET lies in the very high sensitivity of this technique. Extremely small amounts of labelled compound are sufficient to obtain the desired information.

   Modern PET scanners nearly all use scintillators in combination with photomultiplier tubes for the readout of the scintillation light. Table 6.4 lists some scintillators

Table 6.4   Scintillators for positron emission tomography

Material (*)	Density	Att. length [mm] at 511 keV	Photo-fraction [%] at 511 keV	Light yield [ph/MeV]	Decay time [ns]	Emission [nm]
BGO	7.1	10.4	40	9000	300	480
LSO	7.4	11.4	32	26,000	40	420
GSO	6.7	14.1	25	8000	60	440
LuAP	8.3	10.5	30	11,000	18	365
LPS	6.2	14.1	29	20,000	30	380
LaBr$_3$	5.07	22.3	13.1	70,000	16	380

(*) BGO = $Bi_4Ge_3O_{12}$; LSO = Lu_2SiO_5:Ce; GSO = Gd_2SiO_5:Ce; LuAP = $LuAlO_3$:Ce; LPS = $Lu_2Si_2O_7$:Ce

that are used in PET. The most important properties of a scintillator to be used in a PET scanner are that it must allow for a good time resolution, have a large stopping power and a good energy resolution. The time resolution is important because this ensures that the two gamma rays really come from the same annihilation event rather than from two unrelated annihilation events. The stopping power is important because this ensures a large detection efficiency and therefore that a large fraction of the annihilation events will be observed. The stopping power depends on the mean free path for gamma rays of 511 keV in the scintillator material and on the photofraction. The photofraction is the ratio of the cross section for photoelectric absorption over the total gamma interaction cross section. The mean free path of the gamma rays of 511 keV in tissue is about 10 cm. The Compton scattering cross section in living tissue is much larger than the photoelectric cross section. Therefore, a very large fraction of the positron annihilations are followed by the scattering of one, or both, of the 511 keV gamma rays in the body of the patient. If such events are detected, but not recognised as scatter events, the corresponding erroneous information is included in the image. Energy resolution of the scintillator is important because it allows rejecting such scatter events.

In its simplest geometry a PET scanner consists of rings of scintillator blocks each equipped with its photodetector surrounding the patient, as shown in Fig. 6.29. If one of the 511 keV gamma rays interacts in one of the scintillator blocks, the position accuracy on this interaction point is equal to the size of the block. One therefore is lead to use a large number of small scintillation blocks, each equipped with its own photodetector. Some PET systems indeed use this scheme. However, to reduce the cost, most commercial PET scanners use a system where one uses less

Fig. 6.29 Figure illustrating the principle of a PET scanner. The white lines represent the directions of flight of the gamma rays originating from a few annihilation events. These gamma rays interact in the scintillator blocks. If two gamma rays interact simultaneously in two scintillator blocks, one assumes a positron annihilation occurred somewhere along the line joining the centres of the two blocks of scintillating material
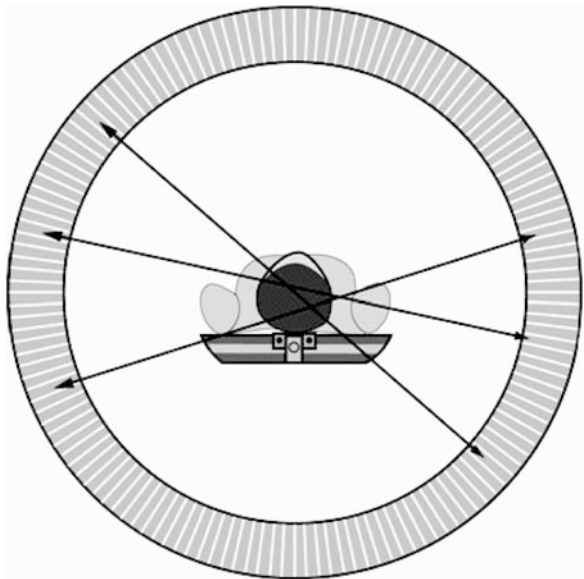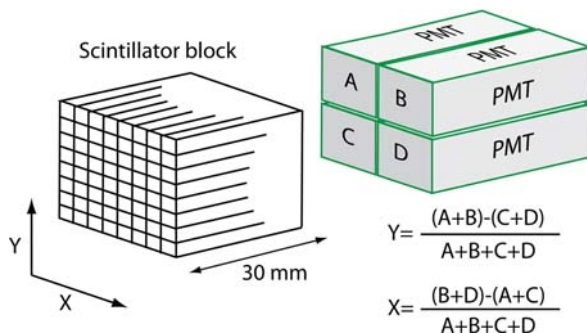
Fig. 6.30 Illustration of the principle of a block detector allowing to read a large number of small crystals with fewer photodetectors

$$Y= \frac{(A+B)-(C+D)}{A+B+C+D}$$

$$X= \frac{(B+D)-(A+C)}{A+B+C+D}$$

photodetectors than crystals. A possible realisation of this detection method is illustrated in Fig. 6.30. In this design, groves are cut in a large piece of scintillator such that it is divided in 64 individual crystals and this block of scintillator is in contact with only four photomultiplier tubes. Notice that the groves cut in the scintillator do not go all the way to the bottom of the block. In this way, the light can spread over the four photomultiplier tubes with the distribution of the light depending on the position of the crystal where the gamma ray interacted. From the values for the ratios of the PM signal amplitudes, X and Y defined in Fig. 6.30, the crystal where the gamma ray interacted, can be identified. Notice that relation between values of (X-Y) and the crystal is not linear. The mapping between the values of (X-Y) and the crystal where the gamma ray interacted must be determined experimentally.

The most commonly used positron-emitting isotopes in PET are listed in Table 6.5. It can be seen that all these isotopes have a short lifetime. This is a desirable property, because in this way the activity naturally disappears from the body of the patient after the examination. However, because of this short lifetime, it is necessary to produce the isotope and incorporate it into a molecule of interest in a very short time, just before the examination. For carbon, nitrogen and oxygen, a cyclotron and chemical synthesis equipment at the hospital is necessary. For ^{18}F a site within a few hours driving of the hospital is possible. Figure 6.31 shows an example of a commercial PET scanner.

Figure 6.32 illustrates the power of PET by showing PET scans of the brain for normal healthy volunteers. These subjects were injected with fluoro-desoxy-glucose (FDG). This is glucose where one oxygen atom was replaced by ^{18}F. The compound

Table 6.5 Most commonly used isotopes in PET

Isotope	Half-life [min]	Max e^+ energy [keV]
^{11}C	20.4	960
^{13}N	9.96	1190
^{15}O	2.05	1720
^{18}F	110.	625

Fig. 6.31 The Siemens
ECAT EXACT HR clinical
PET scanner. It contains
25′088 LSO crystals
measuring 5.9 × 2.9 ×
30 mm^2 and arranged in a
15 cm long cylinder
surrounding the patient. This
ring of detectors is hidden in
the square box shown behind
the patient's bed.
Photography courtesy of
Siemens

is metabolically very similar to normal glucose. The glucose concentrates in areas
where there is increased metabolic activity in the brain and this in turn shows up
as an increased amount of radioactivity. The right-hand side of the figure shows
four scans where the subject was given different kinds of auditory stimulation. With
only verbal stimuli (a Sherlock Holmes story), the left-hand side of the brain appears
more active; with non-verbal stimuli (a Brandenburg concerto) there is more activity
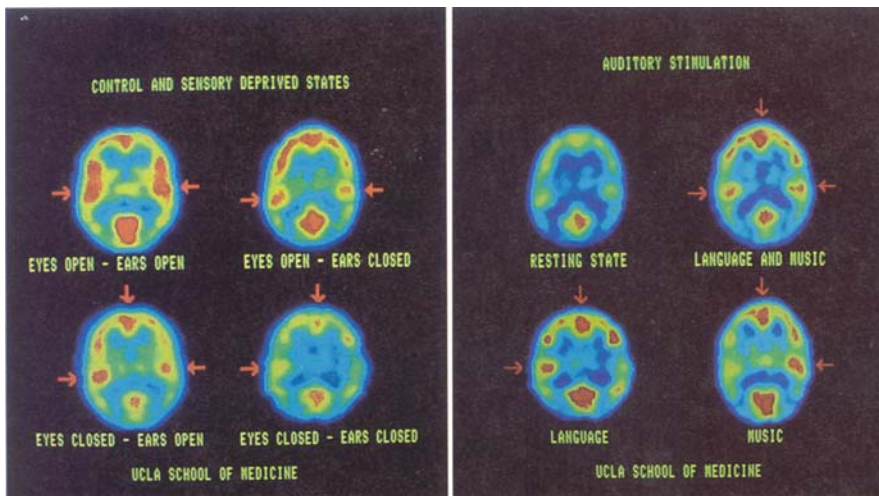in the right-hand side.

Fig. 6.32 PET scans for normal volunteers that have been injected with FDG. The different
images correspond to different conditions as indicated. The colour indicates the amount of activity
observed in a particular part of the brain. In decreasing order of activity the colour code is: *red,
yellow, green, blue, dark blue*. Figure from [9], Copyright John Wiley & Sons, Inc., 1984,
reproduced with permission

In the presently used commercial PET scanners the accuracy on the time difference between the two gamma rays is a few nanoseconds. This is sufficient for random coincidence rejection. However, if this timing accuracy could be improved to a value significantly better than 1 ns, it would be possible to use the timing information to localise the annihilation point along the line of flight of the two gamma rays. This information could be used to reduce the statistical noise in the image. Such systems are presently actively being developed.

6.8 Exercises

1. Assume that a detector for gamma rays consisting of LSO scintillator and a photomultiplier tube. The signal is taken from the anode of the PMT with a 50 Ω coax cable and brought to an oscilloscope. The input impedance of the oscilloscope is 50 Ω and the gain of the PMT is 10^5. What will be the signal amplitude when observing gamma rays of 1 MeV? Note: assume a light collection efficiency 50% and a photocathode quantum efficiency of 25%.
2. Consider a source emitting gamma rays of 511 keV. Calculate the energy where you expect the backscatter peak in the pulse height spectrum.
3. Consider a CsI:Tl scintillator. What fraction of the energy lost due to the interactions with the electrons in the material is converted to scintillation light?
4. Consider a PIN diode with a quantum efficiency of Eff = 60%. Assume that it is exposed to a light flux of 1 μW at a wavelength of 565 nm. What will be the photocurrent?
5. Consider a PET scanner with a solid angle covering around it centre point of $\Omega = 10\%$. Assume that the detection efficiency for a gamma ray of 511 keV and within the solid angle is Eff = 20%. Assume that you place a point source in its centre with an activity of 1 mCi. What will be the single count rate and the coincidence count rate?
6. In several photodetectors, the charge multiplication gives rise to an exponential pulse height distribution for single primary charges. Show that in this case the excess noise factor equals 2.

References

1. P.A. Rodnyi, Physical processes in inorganic scintillators, CRC press (1997).
2. P. Dorenbos, J.T.M. Haas and C.W.E. van Eijk, Non-linear response in the scintillation yield of Lu_2SiO_5:Ce_{3+}, IEEE Trans. Nucl. Sci. 41(4) (1994).
3. R.J. McIntyre, P.P. Webb and H. Dautet, A short-wavelength reach-through avalanche photodiode, IEEE Trans. Nucl. Sci. 43, 1341–1346 (1996).
4. R.L. Heath, Scintillation spectroscopy gamma-ray spectrum catalogue, IDO-16880, 1 and 2 (1964).
5. J.C. Philipot, IEEE Trans. Nucl. Sci. NS-17(3) 446 (1970).
6. C.W. Fabjan and F. Gianotti, Calorimetry for particle physics, Rev. Mod. Phys. 75, 1243–1284 (2003).

 7. R. Wigmans, Calorimetry, energy measurement in particle physics, Oxford University Press (2000).
 8. The CMS collaboration, The CMS experiment at the Cern LHC, Jinst 3, (2008).
 9. J.C. Mazziotta, M.E. Phelps, Human sensory stimulation and deprivation: Positron emission tomographic results and startegies, Ann. Neurol. 15(S1), 50–60 (1984).
10. RMD, Radiation Monitoring Devices, Inc., 44 Hunt Street Watertown, MA 02472.
11. J.M. Schonkeren, Photomultipliers, Philips application book series, H. Kater and L.J. Thompson (eds), Eindhoven, The Netherlands: Philips (1970).
12. M.D. Hull, Fast response photomultipliers, Philips application book series, Eindhoven, The Netherlands: Philips (1971).
13. B. Candy, Photomultiplier characteristics and practice relevant to photon counting, Rev. Sci. Instr. 56, 183 (1985).
14. S.R. Cherry, J.A. Sorenson and M.E. Phelps, Physics in nuclear medicine, Elsevier Science (2003).
15. G. Muehllehner and J. Karp, Positron emission tomography, Phys. Med. Biol. 51, R117–R137 (2006).
16. M.E. Phelps, S. Cherry and M. Dahlbom, PET, Springer (2006).
17. T. Lewellen, Recent developments in PET detector technology, Phys. Med. Biol. 53, R287–R317 (2008).
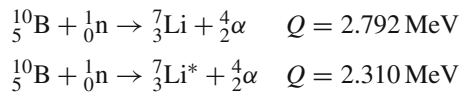
Chapter 7
Neutron Detection

A neutron has no electric charge and therefore leaves no trace in matter except when it undergoes a nuclear interaction. All detection of neutrons is therefore based on letting the neutron interact and observing the charged reaction products. The detection of high-energy neutrons was discussed in Sect. 6.6. The present chapter deals with the detection of neutrons at nuclear energies, i.e. neutrons with a kinetic energy less than a few 10 MeV. Because the neutron interaction cross sections in most materials are very strongly dependent on the energy, different methods are used for neutrons of different energies. In many applications, the neutrons have to be detected in the presence of a large gamma ray background. The issue therefore often is distinguishing the neutrons from the gamma rays.

The present chapter is for a large part based on [1] in Chap. 1 where a more extensive discussion of the subject can be found.

7.1 Slow Neutron Detection

The term 'slow neutrons' refers to neutrons with a kinetic energy of less than 0.5 eV. In this energy range, a very important reaction is the neutron capture. Some isotopes have very large neutron absorption cross sections and all slow neutron detection will be based on using one of these isotopes. The most important isotopes for this purpose are

Boron-10. The neutron absorption reaction is

$$\begin{array}{ll} {}^{10}_{5}\text{B} + {}^{1}_{0}\text{n} \rightarrow {}^{7}_{3}\text{Li} + {}^{4}_{2}\alpha & Q = 2.792\,\text{MeV} \\ {}^{10}_{5}\text{B} + {}^{1}_{0}\text{n} \rightarrow {}^{7}_{3}\text{Li}^* + {}^{4}_{2}\alpha & Q = 2.310\,\text{MeV} \end{array}$$

The symbol Q stands for the difference in mass between the particles in the initial and the final state and therefore for the total energy liberated in the reaction. The exited state ${}^{7}\text{Li}^*$ decays with emission of a gamma ray with a lifetime of $\approx 10^{-3}$ s. Usually, this gamma ray escapes unobserved from the detector. For thermal neutrons, the second reaction has a branching fraction of 96%.

Energy-momentum conservation applied to the reaction above requires

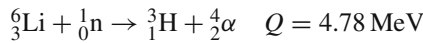$$E_{Li} + E_\alpha = Q$$
$$m_{Li} v_{Li} = m_\alpha v_\alpha$$

From this we readily get

$$E_{Li} = \frac{m_\alpha}{m_\alpha + m_{Li}} Q = 0.84 \, \text{MeV}$$

$$E_\alpha = \frac{m_{Li}}{m_\alpha + m_{Li}} Q = 1.47 \, \text{MeV}$$

Two final state objects, the ^{7}Li* nucleus and the alpha particle, are almost exactly back-to-back. The cross section for this reaction as a function of energy is shown in Fig. 7.1. The natural abundance of ^{10}B is 19.9%.

Lithium-6. The neutron absorption reaction is

$$^6_3\text{Li} + ^1_0\text{n} \rightarrow ^3_1\text{H} + ^4_2\alpha \quad Q = 4.78 \, \text{MeV}$$

The cross section for this reaction is also given in Fig. 7.1. The natural abundance of ^{6}Li is 7.4% and lithium is available in an isotopically enriched form.

Helium-3. The neutron absorption reaction is

$$^3_2\text{He} + ^1_0\text{n} \rightarrow ^3_1\text{H} + ^1_1 p \quad Q = 0.764 \, \text{MeV}$$
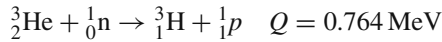
The natural abundance of ^{3}He is extremely low; only 0.000137% of all naturally occurring helium is ^{3}He! Commercial ^{3}He is made in nuclear reactions. The neutron absorption cross section of ^{3}He is shown in Fig. 7.1.
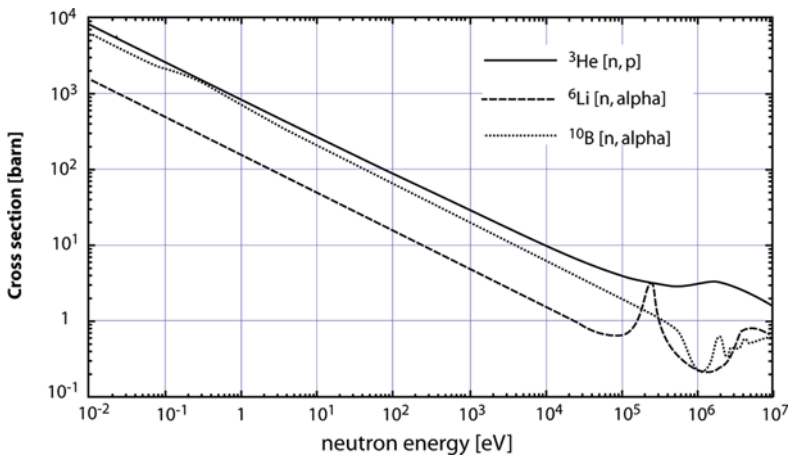

Fig. 7.1 Neutron capture cross sections versus neutron energy for ^{10}B, ^{6}Li and ^{3}He. The data for this figure were obtained from [8] in Chap. 1

Gadolinium-157. The isotope ^{157}Gd has a very large capture cross section for thermal neutrons: 255,000 barns! The natural abundance of ^{156}Gd is 15.7%. There are several final states containing gamma rays and conversion electrons. A 72 keV conversion electron is present in 39% of the neutron capture events and neutron detection is usually based on the observation of this final state electron. The gadolinium reaction will not allow easy discrimination between neutrons and gamma rays, since an electron from a gamma interaction will not easily be distinguished from the conversion electron. Gadolinium can only be used when gamma rejection is less important.

Uranium and plutonium. The fission cross sections of ^{233}U, ^{235}U and ^{239}Pu are large for slow neutrons (see Fig. 2.19). These reactions can therefore be used for slow neutron detection. The Q value in these reactions is always very large, typically 200 MeV, allowing easy discrimination between gamma rays and neutrons. However, these nuclei are all alpha emitters and this can give rise to pulses that may be confused with neutron capture events. One needs to rely on pulse height discrimination to distinguish the neutron events from alpha emissions.

Proportional tubes using one of the above neutron capture reactions are commonly used for slow neutron detection. Either a gas containing ^{10}B or ^{3}He can be used, or a tube lined with ^{10}B or with one of the fissile isotopes ^{233}U, ^{235}U or ^{239}Pu. These counters will also be sensitive to gamma rays, because the gamma rays will interact in the walls of the tube and produce electrons in the active gas volume. Such an electron will deposit about 2 keV/cm of gas and therefore a gamma ray will almost never deposit more than a few 10 keV in the counter. The neutron capture cross sections considered above give rise to very ionising alpha particles or nuclei and the range of these in the counter gas is of the order of centimetres. One or both of the fragments will usually deposit all their energy in the gas. If individual pulse heights are recorded, these detectors are therefore very efficient in discriminating between neutrons and gamma rays. This is illustrated in Fig. 7.2.

Only BF$_3$ or ^{3}He are used as active chamber gas in the proportional tubes. ^{3}He with $\approx$5% of quenching gas is a good working gas in a proportional counter. The detector will also work well at a pressure of several bar. Increasing the pressure is

Fig. 7.2 Gamma interactions and neutron interactions in a proportional counter filled with ^{3}He or BF$_3$. A neutron interaction gives rise to two very ionising tracks, while a gamma ray will usually interact in the walls of the detector and give rise to a minimum ionising track
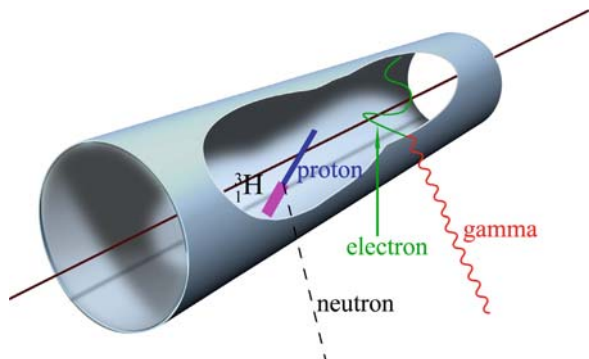
often used to increase the neutron detection efficiency. The mean free path of thermal neutrons in ^{3}He at one atmosphere is 7.3 cm (see Exercise 2). However, the lower Q value compared to boron makes distinguishing neutron interactions from gamma interactions more difficult with ^{3}He. Moreover, ^{3}He is expensive; commercial ^{3}He is made in nuclear reactions and its price was approximately 200 €/litre of gas at standard temperature and pressure in 2008.

BF$_3$ is not a very good working gas for a proportional counter, probably because BF$_3$ is slightly electronegative. Often tubes filled with BF$_3$ are used at reduced pressure where the operating conditions are more stable. BF$_3$-filled proportional tubes are also prone to ageing. Furthermore, this gas is very toxic and is also corrosive. However, the high Q value of this reaction helps in distinguishing neutron interactions from gamma interactions.

If the diameter of the tube is very large compared to the typical range of the reaction products in the gas, nearly all the interactions give rise to the same pulse height. In this case, the pulse height spectrum in ^{3}He shows one clear peak and BF$_3$ it shows two peaks corresponding to the two reaction channels. The range of the reaction products in the gas is of the order of centimetres and the proportional tubes are usually only a few centimetres in diameter. As a result, one of the reaction products will often hit the wall of the tube before it reaches the end of its range. The result is a pulse height spectrum showing a broad shoulder below the total absorption peak as illustrated in Fig. 7.3. This reduces the ability to discriminate between gamma rays and neutrons.

An alternative approach is to line the inner walls of the proportional tube with a suitable neutron-sensitive material. In this case, any proportional chamber gas, such as argon–isobutane can be used. For lining the inner wall of the detector, ^{10}B is often used. Obviously, the layer of ^{10}B should be very thin (about 1 mg/cm^2), otherwise the reaction products will be absorbed in the ^{10}B layer itself and not enter into the active gas of the counter. As a result, the detection efficiency for neutrons of this type of detector is only of the order of 1%, much lower than what can be achieved with a neutron-sensitive gas. The pulse height spectrum obtained with such a chamber will
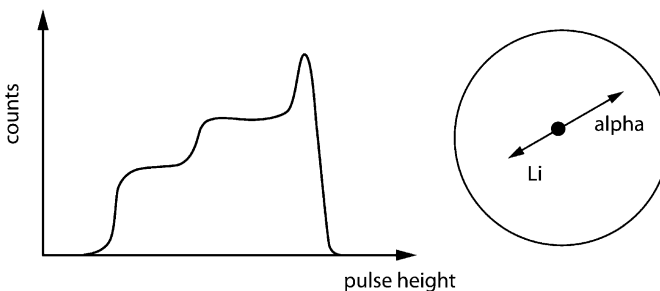

Fig. 7.3 Pulse height spectrum in a proportional tube filled with BF$_3$ exposed to a neutron flux. The step structure is due to one of the reaction products hitting the wall of the tube before it reaches the end of its range. The figure to the right shows a neutron interaction in a proportional tube where the reaction is fully contained in the active gas of the chamber. Such an event would be in the total absorption peak at the right of the pulse height spectrum

be more or less flat from zero to a maximum value corresponding to1.47 MeV. This is the maximum energy the alpha particle can deposit in the gas. Because of this flat pulse height spectrum, this type of detectors is less efficient in discriminating between neutrons and gamma rays.

One of the big limitations of proportional counters used as neutron detectors is the slow rise-time of the pulse and the concomitant poor time resolution. This rise-time is often several microseconds. A good time resolution is essential in some applications, for example for 'time of flight measurements'.

When a good time resolution is important, neutron-sensitive scintillators are preferred. A commonly used neutron-sensitive scintillator is LiI doped with $\approx 1\%$ of europium. The lithium provides the neutron sensitivity. This scintillator material is somewhat similar to NaI:Tl. It has a light yield of about 14,000 photons/MeV and a decay time of 300 ns. LiI is also very hygroscopic. Other scintillator materials that are often used are a scintillator made from fusing B_2O_3 and ZnS and a plastic scintillator based on boron-loaded plastic. Such boron-loaded plastic scintillators are commercially available with 5% of boron content and with a light yield that is 75% of a typical plastic scintillator.

7.2 Neutron Detectors for Nuclear Reactors

In nuclear reactors, most of the power is generated through fission induced by slow neutrons. It is therefore important for the reactor control to measure the slow neutron flux in the reactor. It is customary to distinguish between 'in-core' detectors and 'out-of-core' detectors. Particularly, in-core detectors must work at very high neutron fluxes, at very high temperature and need to be very small.

For reactor monitoring, gas-filled detectors are almost always used because of the low gamma sensitivity and the good radiation hardness of this type of detector.

In Sect. 7.1, we have considered proportional counters working in the pulse-readout mode. The big advantage of the pulse-readout mode is that it allows a very good rejection of the gamma rays based on pulse height. However, at event rates exceeding $\approx 10^6$ counts/s, it becomes very difficult to use the pulse-readout mode. In addition, at very large fluxes it is impossible to use detectors with gas amplification, because the space charge due to the positive ions in the tube becomes so large that it introduces strong non-linear effects. The only possibility is to use gas chambers working in the ionisation mode and measuring the ionisation current; therefore, the detector becomes much more sensitive to gamma radiation. Some gamma rejections can be obtained by measuring the fluctuations on the detector current. This is the 'Campbell technique' and is discussed in Sect. 8.6. Another way to correct for the gamma background is by using two identical ionisation chambers, one detector lined with a neutron-sensitive layer and a second identical detector but without a neutron-sensitive layer. The two chambers have the same sensitivity to gamma rays. The difference between the ionisation currents in both detectors gives the neutron-induced signal.
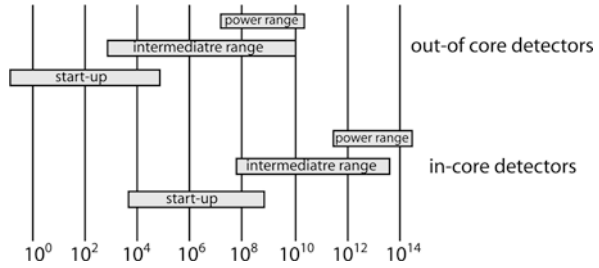
In pressurised water reactors (PWR), the routine monitoring of the neutron flux
is usually done with 'out-of-core' detectors. In addition, there are usually in-core
detectors for fuel management. The out-of-core detectors measure neutron fluxes
in the range 0–10^{10} neutrons/cm^2/s. In boiling water reactors, there are usually in-
core detectors. These work in the range 10^4–10^{14} neutrons/cm^2/s. There are usually
different detectors for the start-up of the reactors, for monitoring during full power
operation and for the intermediate regime. The neutron flux ranges for these differ-
ent detectors are illustrated in Fig. 7.4. In the start-up regime the neutron fluxes are
low, but the gamma fluxes are relatively high and good gamma rejection is essential.
Detectors working in pulse readout mode are normally used. At full power the neu-
tron fluxes are always very large and only detectors working in the current readout
mode can be used. The gamma flux in the core of a reactor working at full power is
typically 10^8 R/h (R stands for roentgen, see Sect. 3.2).

A typical in-core neutron detector for a boiling water reactor is the fission cham-
ber shown in Fig. 7.5. A small cylindrical gas volume is filled with argon at a
pressure of several atmospheres. The walls of the chamber are lined with highly
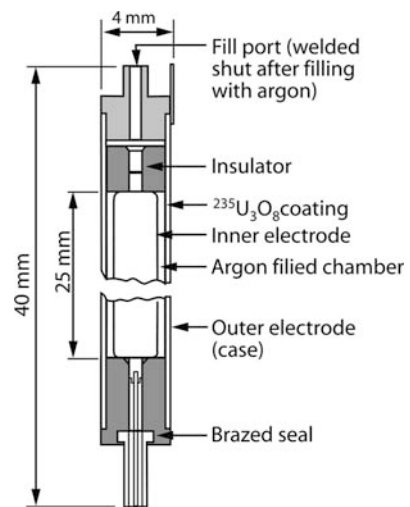

Fig. 7.5 A typical in-core
fission chamber used in BWR
neutron flux monitoring
systems

enriched $^{235}U_3O_8$. One difficulty with these devices is the burn-up of the neutron-sensitive material. One year of operation will typically correspond to an integrated neutron flux of 1.7 10^{21} and a reduction of the signal by as much as a factor of two has been reported when simple lining with $^{235}U_3O_8$ is used. One method for reducing the effects of burn-up in fission chambers is to combine fertile and fissile material in the neutron-sensitive lining of the detector. In this case, the fertile iso-tope will gradually be converted in fissile nuclei and compensate for the burn-up of the original fissile material. For this, mixtures of ^{238}U and ^{239}Pu or mixtures of ^{234}U and ^{235}U are used.

Fission chambers that have been exposed for a long time (several days) to a high neutron flux show a memory effect due to the build-up of fission products within the chamber. Immediately after exposure, a residual current of $\approx 0.1\%$ of the full current is observed. This goes to 10^{-5} after 10 days.

Another detector that is widely used for reactor monitoring is the so-called 'self-powered detector'. These devices are based on the use of a material with a relatively large cross section for neutron capture followed by a subsequent beta decay. The detector measures the current caused by the beta emissions over a very thin isolat-ing gap. The insulating material in the gap usually consists of magnesium oxide or of aluminium oxide. No voltage is needed to collect charges and this explains the name 'self-powered detector'. Figure 7.6 shows very schematically the geometry of this kind of detector. For the emitter, vanadium or rhodium are commonly used. Table 7.1 summarises some of the important properties of these materials when used as emitter in self-powered detectors. The somewhat lower cross section of vanadium turns out to be an advantage because it reduces the burn-up of the detector. Because
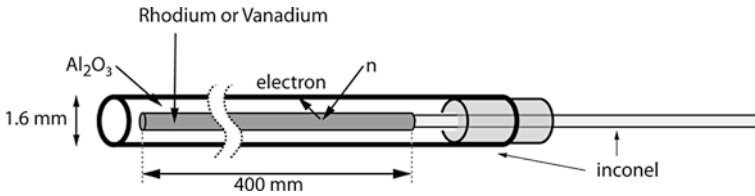

Fig. 7.6 Schematic representation of a self-powered neutron detector for in-core neutron flux monitoring

Table 7.1 Emitter materials commonly used in self-powered detectors

Emitter material	Isotope of interest and abundance	Activation cross section [barn]	Half-life of induced beta activity	Beta endpoint energy	Neutron sensitivity $\left[\dfrac{A}{n/cm^2 \cdot s}\right]$
Vanadium	$^{51}_{23}V(99.75\%)$	4.9	225 s	2.47 MeV	5×10^{-23}
Rhodium	$^{103}_{45}Rh(100\%)$	139	44 s	2.44 MeV	1×10^{-21}
		11	265 s		

of the decay time of the beta emissions, it takes several times this half-life before the detector signal fully reflects a change in neutron flux.

One of the primary disadvantages of self-powered detectors based on beta emission is this slow response time. To avoid this, sometimes, other types of self-powered detectors are used. These are based on prompt gamma emission that follows neutron capture in certain materials. Cobalt and cadmium emitters are often used for this purpose, but this type of detector tends to be less sensitive to neutrons and more sensitive to gamma rays.

7.3 Fast Neutron Detection

The detectors for slow neutrons discussed in Sects. 7.1 and 7.2 rely on the very large cross section for slow neutron capture by certain isotopes. Without the addition of a moderator, these detectors are not well suited for detecting fast neutrons. The elastic cross section for fast neutrons is large in several materials. Detecting the recoil nucleus in elastic scattering is therefore a good method for detecting fast neutrons and this forms the basis for a wide variety of neutron detectors. Most detectors for fast neutrons are based either on using neutron moderation or on using elastic scattering. These two classes of neutron detectors are discussed below.

7.3.1 Detectors for Fast Neutrons Based on Moderation

For several isotopes, the neutron capture cross sections for slow neutrons are several orders of magnitude larger than typical cross sections for fast neutrons. It is therefore possible to build efficient detectors for fast neutrons that are based on first slowing down the neutrons and then detecting them. Several useful detectors are based on this principle.

The spherical neutron dosimeter, sometimes called Bonner counter, is one of these. This instrument consists of a small detector for slow neutrons, surrounded by a sphere of moderator material, usually polyethylene or paraffin. The slow neutron detector is either a LiI scintillator or a small ^{3}He counter. The moderator sphere is typically 10–12 inch in diameter, while the LiI scintillator is only a few millimetres in size.

If a low-energy neutron enters the detector, it will quickly slow down and subsequently it has a good chance of being absorbed somewhere in the moderator before it can reach the LiI scintillator. As the energy of the neutron increases, the detection efficiency increases, because the neutron has a greater chance to reach the LiI scintillator before it is absorbed in the moderator. As the energy is increased further, the detection efficiency decreases again since the neutron has less chance to be fully thermalised in the moderator. The result is a neutron detection efficiency that first increases with energy, reaches a maximum and then decreases again. The interest in

Fig. 7.7 Fuji Electric NSN10014 neutron dosimeter. The detector consists of a ^{3}He proportional counter in a polyethylene sphere. It can be used for neutrons with energy between 0.025 eV and 8 MeV. Figure courtesy of Fuji Electric

this type of detector stems from the fact that the energy dependence of the detection efficiency mimics the energy dependence of the dose equivalent for biological damage delivered per neutron between 0 and ≈10 MeV. The similarity of the two energy dependences is purely accidental, but allows very useful neutron dosimeters to be made. For a detector consisting of a 12-inch sphere of polyethylene and a 4-mm LiI scintillator, the sensitivity is 3000 counts/mrem. The advantage of using a ^{3}He counter instead of a LiI scintillator is the reduced sensitivity to gamma rays. With a more elaborate design of the absorber sphere, it is possible to extend the sensitivity to neutrons with energy larger than 10 MeV. Neutron dosimeters based on this principle are available from several companies. Figure 7.7 shows an example of such a commercial neutron dosimeter.

Often a counter with a detection efficiency independent of the neutron energy is required. It turns out that this can be achieved by using a cylindrical geometry with a slow neutron detector in the axis and a cylinder of moderator material around it. This type of counter is often called a 'long counter' because of its shape. For the slow neutron detection, usually a ^{3}He proportional tube is used. Some additional holes and neutron absorbers are often needed to obtain a satisfactory energy response. Figure 7.8 shows an example of a simple long counter.

This detector has the desired response only if the neutrons are coming from a point situated on axis and to the right-hand side in the figure. The detection efficiency of a long counter as shown in Fig. 7.8 is only ≈0.25%, but a more elaborate design with several parallel ^{3}He detectors allows this to be increased considerably while maintaining the flat energy response.

The long counter can also provide some indication of the neutron energy spectrum if the ^{3}He counter records the longitudinal position of the neutron interaction. Figure 7.9 shows the longitudinal distribution of the neutron absorption point in a long counter exposed to neutrons of various energies. If this counter is exposed to neutron radiation of unknown energy spectrum, it is possible to derive the energy
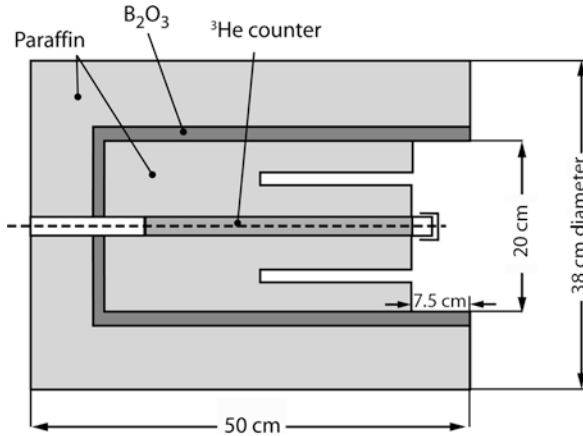
Fig. 7.8 Typical long counter. If this type of detector is exposed to a flux of neutrons coming from the right-hand side, its sensitivity is independent of the energy of the neutron for neutron energies of up to a few 10 MeV
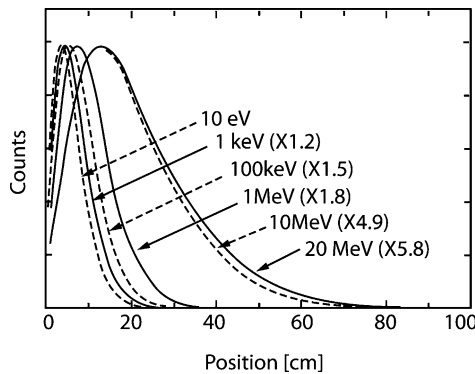

Fig. 7.9 Calculated depth distribution of the neutron absorption point for the neutrons of different energy in a long counter. In this example, we assumed a cylindrical moderator of 30 cm diameter and having a first part with 40 cm of polycarbonate followed by 60 cm polyethylene. The distributions for different neutron energies are normalised such as to all have the same height. The normalisation factors are indicated in the figure. Figure from [1], with permission

spectrum of these neutrons by fitting the shape of the observed interaction depth distribution to a superposition of response curves of neutrons with different energies.

7.3.2 Detectors Based on the Observation of the Recoil Nuclei

When detecting fast neutrons, one often wants to detect the energy of these neutrons. This is called neutron spectroscopy. The long counter discussed previously gives some limited spectroscopic information, but for this purpose usually detectors
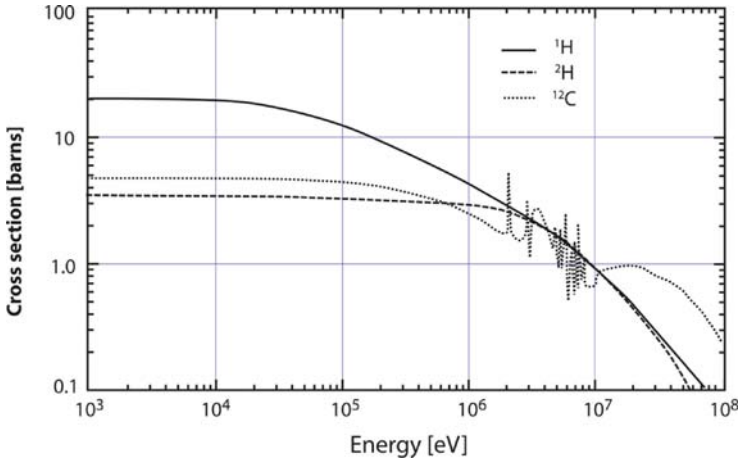
Fig. 7.10 Total elastic cross section for neutrons on hydrogen, deuterium and carbon. The energy dependence of the cross section on carbon between 2 and 10 MeV is extremely complex and is only represented in a very approximate way in the figure. The data for this figure were obtained from [4] in Chap. 1

depending on the observation of the recoil nucleus in elastic scattering are used. The most important target nucleus is hydrogen, because in neutron elastic scattering on hydrogen, the recoil nucleus can receive up to the total energy of the incoming neutron. Moreover, the elastic scattering cross section of neutrons on hydrogen is large and well known. The cross sections for elastic collisions of neutrons on a few relevant light nuclei are given in Fig. 7.10.

Let us assume that we have neutrons of a given energy interacting in a hydrogen-rich scintillator material. The pulse height spectrum of the scintillator will be equal to the energy spectrum of the recoil protons. The recoil protons have an energy spectrum that depends on the neutron energy. If we have neutrons with an unknown energy spectrum, it is possible to derive the neutron energy spectrum by fitting the shape of the observed pulse height spectrum of the scintillator to a superposition of response curves of the counter to neutrons of different energies.

If a neutron scatters on a target nucleus A, the direction and the energy of the scattered target nucleus are related by energy and momentum conservation. We are considering here neutrons with energy of at most a few 10 MeV. Since the rest mass of a neutron is 939.56 MeV, a non-relativistic approximation is sufficient. We leave it as an exercise for the student to show that we have the following relation (see Exercise 4)

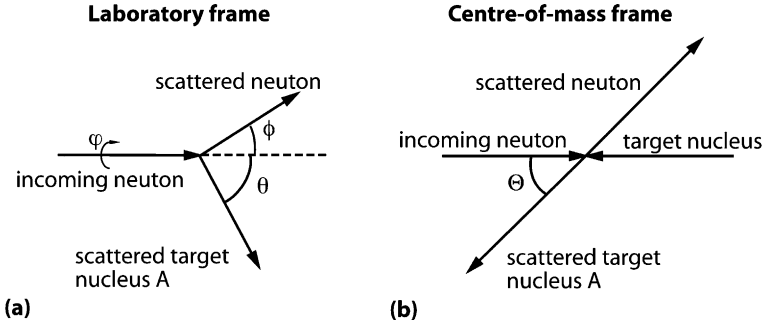$$E_{\text{recoil}} = E_{\text{neutron}} \frac{4m_A m_n}{(m_A + m_n)^2} \cos^2 \theta \tag{7.1}$$

Laboratory frame   **Centre-of-mass frame**

Fig. 7.11 Definition of the scattering angle in the centre-of-mass system and in the laboratory system. The symbol φ denotes the polar angle of the nucleus A around the direction of the incoming neutron

In this equation the notations are as follows:

E_{recoil} = energy of the recoil nucleus
$E_{neutron}$ = energy of the neutron
m_A = mass of nucleus A
m_n = mass of neutron
θ = angle of the recoil nucleus relative to the direction of the incoming neutron in the laboratory frame. See Fig. 7.11 for a definition of the angles.

Furthermore, the scattering angle in the centre-of-mass frame Θ and the scattering angle in the laboratory frame θ are related by (see Exercise 5)

$$\cos\theta = \sqrt{\frac{1 - \cos\Theta}{2}} \tag{7.2}$$

Therefore, the relation between the recoil energy and the scattering angle Θ in the centre-of-mass frame is given by

$$E_{recoil} = E_{neutron}\frac{2m_A m_n}{(m_A + m_n)^2}[1 - \cos\Theta] \tag{7.3}$$

Consider the case of a head-on collision between a neutron and a target nucleus A. In the centre-of-mass frame, the nucleus A bounces back in the direction it came from without changing its energy. In the centre-of-mass frame the angle is $\Theta = 180°$. In the laboratory frame, the nucleus A continues in the direction of the neutron and the angle of the scattered nucleus A is $\theta = 0°$. In this case, the energy of the recoil nucleus A is maximum and is given by

$$E_{recoil-max} = E_{neutron}\frac{4m_A m_n}{(m_A + m_n)^2}$$

If the collision in the centre-of-mass system is at small angle Θ, the angle in the laboratory frame is $\theta \approx 90°$. In this case, the energy of the recoil nucleus is small and in the limit $\Theta \to 0$, this energy goes to zero.

The energy spectrum of the recoil nuclei depends on the angular distribution of the neutron scattering. Let us define $\sigma(\Omega)$ as the differential cross section for elastic scattering of the neutron into the solid angle $d\Omega$, with $d\Omega = d\cos\Theta d\varphi$ and $P(\cos\Theta)$ as the probability density function for elastic scattering under an angle Θ, both quantities being defined in the centre-of-mass frame. Let us furthermore use the symbol σ_t to denote the total elastic cross section. Because of the symmetry of the problem, the cross section $\sigma(\Omega)$ does not depend on the azimuthal angle φ, but only on the polar angle Θ. Therefore we can write:

$$P(\cos\Theta)\, d\cos\Theta = \left(\int \frac{\sigma(\Omega)}{\sigma_t}\, d\varphi \right) d\cos\Theta = 2\pi \frac{\sigma(\Omega)}{\sigma_t}\, d\cos\Theta$$

If we define $P(E_r)$ as the probability distribution of the recoil energy E_r, we have

$$P(E_r) = P(\cos\Theta) \left| \frac{d\cos\Theta}{dE_r} \right|$$

And therefore, using Eq. (7.3)

$$P(E_r)\, dE_r = \frac{2\pi}{E_{neutron}} \frac{\sigma(\Omega)}{\sigma_t} \frac{(m_A + m_n)^2}{2m_A m_n} dE_r$$

There is no simple way to find the angular distribution for elastic scattering. For many nuclei, this distribution is peaked in the forward and backward direction.

In the case of hydrogen, and for the energies considered here, the elastic cross section is almost isotropic in the cm frame, i.e. $\sigma(\Omega) = \sigma_t/(4\pi)$. The recoil energy spectrum therefore is simply a constant between 0 and the maximum value! The maximum value is the neutron energy itself. See Fig. 7.12. Assume that we have a mono-energetic beam of neutrons with energy E_n and with a flux $\Phi(E_n)$. The total number of counts in the spectrum will be proportional to $\Phi(E_n)\,\sigma_t(E_n)$. The number
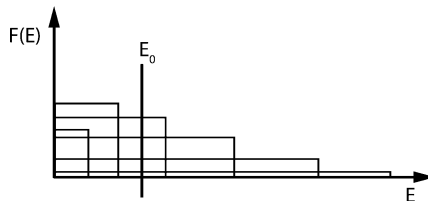

Fig. 7.12 Pulse height distribution of a scintillator exposed to a neutron beam with a mixture of neutron energies. The value of $F(E_0)$ is the sum of all recoil distributions for neutrons with an energy larger than E_0

of counts in the pulse height spectrum $F(E)$, in the interval $(E, E+\Delta E)$ will be given by

$$F(E)\Delta E = \frac{\Delta E}{E_n}\Phi(E_n)\,\sigma(E_n) \quad if\ E_n > E$$

$$F(E)\Delta E = 0 \qquad\qquad\qquad if\ E_n < E_n$$

For a beam of neutrons with spectrum $\Phi(E_n)$, the pulse height will be proportional to

$$F(E) = \int\limits_{E}^{\infty} \Phi(E')\frac{\sigma(E')}{E'}dE'$$

There is therefore a simple relation between the neutron flux and the derivative of the pulse height spectrum

$$\frac{dF(E)}{d(E)} = -\Phi(E)\frac{\sigma(E)}{E}$$

For measuring the recoil spectrum, plastic scintillators are normally used. Several plastic scintillators contain only hydrogen and carbon and are very well suited for this purpose. The neutron energy spectrum is directly related to the derivative of the pulse height spectrum! However, this simple relation can only be an approximation due to the following reasons:

– the plastic scintillator also contains other elements, at least also carbon, and elastic neutron scattering on carbon should be included in the analysis; moreover, other reactions besides elastic scattering are possible on carbon
– the recoil track sometimes is not fully contained in the scintillator; to minimise this effect one needs to use a large piece of scintillator material
– scattered neutrons can interact a second time in the scintillator; to minimise this effect one needs to use a very small piece of scintillator material
– the response of the scintillator is not linear for very ionising particles, such as the recoil protons
– the scintillator has a finite, i.e. less than perfect, energy resolution
– there usually is background due to gamma rays
– the need for an electronics threshold to cut the noise

All these factors make it non-trivial to obtain a reliable neutron energy spectrum from the pulse height spectrum observed with a scintillator. An exhaustive discussion of the solution to these problems is beyond the scope of the present text. Besides plastic scintillators, proportional tubes filled with some gas with low atomic charge Z are also used in the same way for neutron spectroscopy.
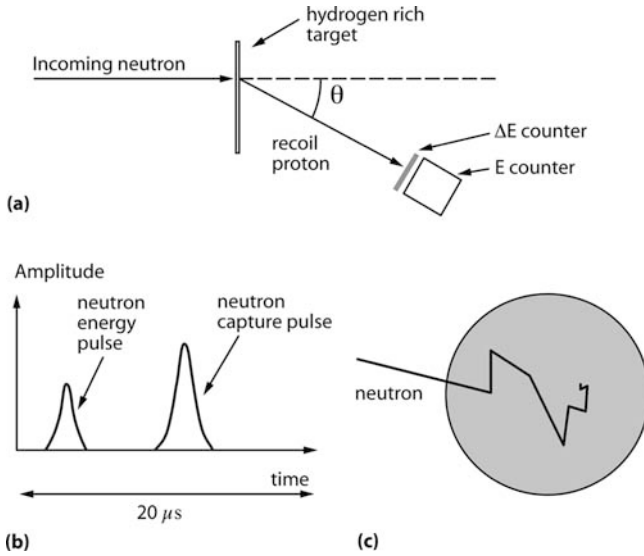
Fig. 7.13 (**a**) Schematic representation of a proton recoil telescope (**b**) Characteristic pulses in a capture-gated spectrometer (**c**) Schematic representation of a capture-gated neutron spectrometer

It is also possible to measure the neutron energy spectrum by directly measuring the scattering angle and the energy of the scattered nucleus and get the neutron energy on an event-by-event basis using Eq. (7.1). This avoids many of the difficulties discussed above, but can only be used if there is a well-collimated neutron beam. Such a proton recoil telescope is shown in Fig. 7.13(a). Neutrons are incident on a thin target foil, usually some organic polymer. The energy of the recoil proton is measured in two detectors, a thin detector measures the energy loss dE/dx and a second thick detector measures the total energy of the recoil proton. Selecting on the ratio of the pulse heights in these two detectors allows recoil particles other than protons to be eliminated and also eliminates some other background. Often these two detectors are at an angle relative to the neutron beam in order to avoid interactions of the neutrons directly in the detectors. Gas detectors, semiconductor detectors and scintillators can all be used for this purpose.

The setup is in a vacuum to prevent the recoil proton from losing too much energy in the gas. The biggest drawback of this type of detector is the low-detection probability for neutrons, typically $\approx 10^{-5}$.

Another neutron spectrometer is the 'capture-gated neutron spectrometer', described below. The principle of this type of detector is illustrated in Fig. 7.13(b). A large volume (>1 litre) of boron-loaded plastic scintillator or a boron-loaded liquid scintillator is exposed to the neutron beam. If a neutron enters the scintillator, it will slow down by elastic collisions with the hydrogen or other light atoms present in the scintillator. It will slow down in a number of steps and in each step it will lose some of its energy. All these elastic collisions happen in a short time,

typically in 50 ns. The light pulses generated in each of these elastic collisions all add up to one single pulse. After having lost all its energy, the neutron will be very slow and will continue to wander around in the scintillator until a boron atom absorbs it. The time before such an absorption takes place can be quite long, several 10 μs. In the absorption by a ^{10}B nucleus, the final ^{7}Li nucleus and the alpha particle together have a kinetic energy of 2.31 MeV and the corresponding energy is always absorbed in the scintillator. Such a neutron capture event will, therefore, have a characteristic signature consisting of a first pulse of energy proportional to the kinetic energy of the neutron followed within ≈20 μs by a second pulse with a characteristic amplitude. If the total event rate in the detector is sufficiently low, the chance association of two unrelated pulses faking a good event will be low. The big advantage of a capture-gated neutron spectrometer compared to a proton recoil spectrometer is that its detection efficiency can be of the order of ≈10% and that it does not require the neutron to come from a well-defined direction.

7.4 Exercises

1. Consider the neutron absorption reaction below, assume thermal neutrons.
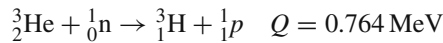
$$\frac{3}{2}\text{He} + \frac{1}{0}\text{n} \rightarrow \frac{3}{1}\text{H} + \frac{1}{1}p \quad Q = 0.764\,\text{MeV}$$

   Use energy and momentum conservation to derive the expression giving the kinetic energy for two final state particles and calculate value of these for the reaction above.
2. A commonly used detector for thermal neutrons is a proportional tube filled with ^{3}He gas. Calculate the mean free path of the neutrons in the gas if this gas is at a pressure of 5 atmospheres. If the tube has an inner diameter of 4 cm, what is the probability that a thermal neutron going through its centre will be detected?
3. Calculate the fractional decrease in sensitivity of a self-powered detector with rhodium emitter after exposure during 6 months to a flux of 3×10^{13} neutrons/cm^2/s.
4. Derive equation (7.1).
5. Derive equation (7.2).
6. Consider a proportional tube filled with ^{3}He and used as a slow neutron detector. The gas gain of the tube is 1000 and the capacitance of the anode wire is 100 pF. What will be the amplitude of a neutron pulse [in mV] be if the integration time of the pulse is very long?

Reference

1. Y. Tanimura, J. Saegusa, M. Yoshizawa, and M. Yoshida, Design of a single moderator-type neutron spectrometer with enhanced energy resolution in the energy range from a few to 100 keV, Nucl. Inst. Meth. A547, 592–600 (2005).

Chapter 8
Electronics for Particle Detectors

8.1 Introduction

In the previous chapters, we have seen that the detection of ionising radiation in the end nearly always comes down to detecting some small electrical signal. Dealing with such small signals is one of the main challenges in designing detectors for nuclear physics and particle physics. Photomultiplier tubes and gas amplification detectors such as Geiger tubes are often used because of their built-in signal amplification mechanism and therefore larger electrical pulses. However, in many detector types there is no such built-in amplification mechanism.

In the present chapter, we will explain the basics of nuclear electronics and discuss the main sources of noise. In Sect. 8.2, we will briefly discuss some important concepts of signal theory. These concepts will be needed throughout the rest of the chapter. I do assume that the reader has at least a basic knowledge of general electronics and circuit theory. In particular, I assume that she or he is familiar with the concept of complex impedance.

A detector in nuclear electronics is always some device with a large resistance. The interaction of ionising radiation induces a small electrical current. From the electrical point of view, a detector is a current source with a large internal resistance and a small capacitance. This is illustrated in Fig. 8.1. Also, in the absence of any ionising radiation there is a small current, which is called the dark current or leakage current depending on the physical mechanism causing it.

There are basically two different modes for measuring nuclear detector signals: current mode and pulse mode. In the current mode, one simply measures the total current of the detector and ignores the pulse nature of the signal. This is simple, but does not allow advantage to be taken of the timing and amplitude information that is present in the signal. In the pulse mode, one observes and counts the individual pulses generated by the particles. The pulse mode always gives superior performance but cannot be used if the rate is too large.

In many detectors the amplitude of the pulses is proportional to the initial charge signal and the arrival time of the pulse is some fixed time after the physical event. By using appropriate thresholds, one can select and count only those pulses that one wants to count. Often the 'good events' are characterised by some specific

Fig. 8.1 In the electric circuit the detector behaves like a current source with a capacitance and an internal resistance. The two intersecting circles represent a current source
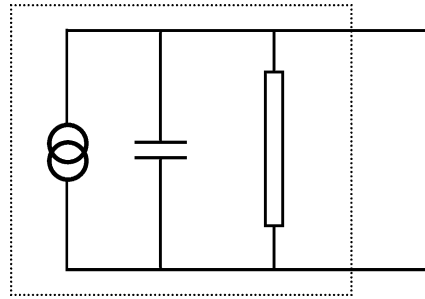
signal amplitude or by the simultaneous presence of two (or more) signals in different detectors. Sometimes also the 'good events' are characterised by the absence of some other signal. Finally, in the pulse mode, one can register a pulse height spectrum and such a spectrum contains a large amount of useful information.

The basic principle of pulse counting is illustrated in Fig. 8.2. The electronics has a threshold that should be well above the noise present in the signal. If the signal is less than the threshold, the output of the circuit is 'zero'. As soon as the signal level exceeds the threshold, the output of the circuit is 'one'. The words 'zero' and 'one' should not be understood as meaning actually zero volt and one volt, but rather as voltage levels that have the meaning 'zero' and 'one'. The number of events can
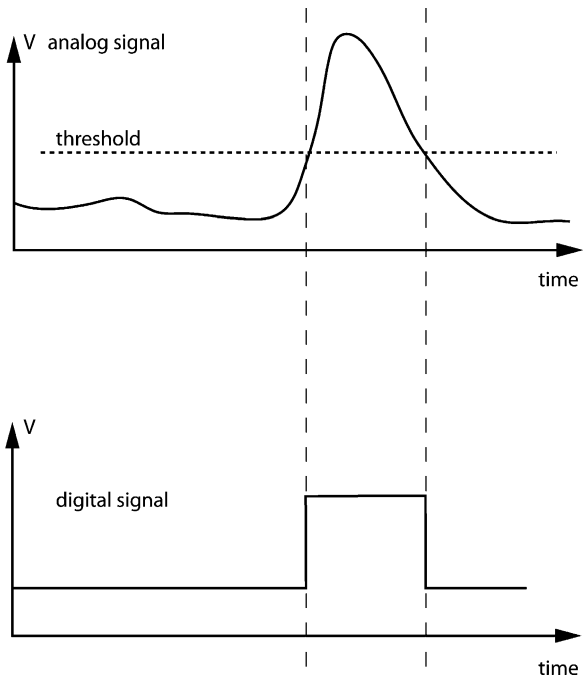

Fig. 8.2 A discrimination circuit has an analog input signal and a digital output signal. If the input signal exceeds some fixed threshold, a digital output signal is generated

now be obtained with some simple counting circuit. In this process, one should be aware that the setup could be inefficient. It means that sometimes a real event in the detector does not produce a pulse that is large enough to produce a signal exceeding the threshold, or a suitable signal was produced, but the electronics did not recognise the pulse because it was arriving at the same time as some other event. This last effect is referred to as dead time.

To obtain a pulse height spectrum the electronics will search for the maximum of the signal in some pre-defined window around the pulse and the value of this maximum is digitised and sent to a computer. The computer stores the values for the maxima of a large number of pulses and displays the result as a histogram.

To see if an event occurred simultaneously with some other event, the electronics will look for the simultaneus presence of two logical signals within some time window as is illustrated in Fig. 8.3. In coincidence counting, one should be aware of the possibility to have random coincidences. These are occurences of a concidence caused by two unrelated events arriving by chance at the same time. It is easy to see that the rate of random coincidences between two signals is proportinal to the rate of each type of signal times the duration of the coincidence window.

$$\frac{dN_{\mathrm{random}}}{dt} = \frac{dN_1}{dt}\frac{dN_2}{dt}\Delta t$$
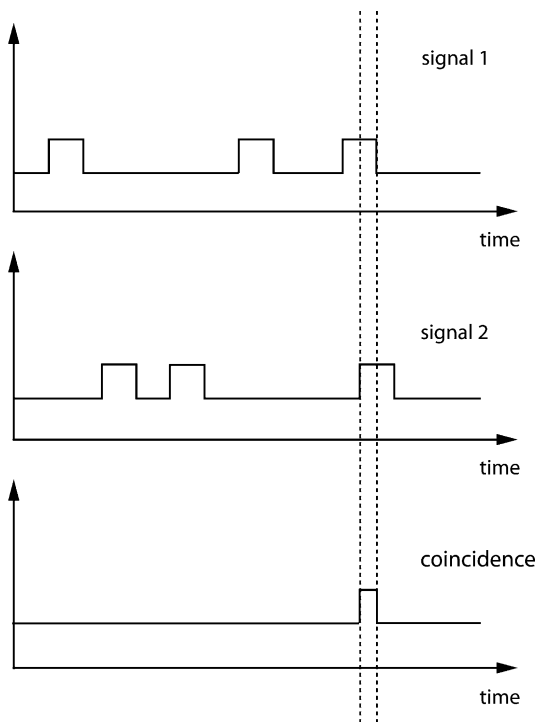

Fig. 8.3 A coincidence circuit has two digital input channels and one digital output channel. If the two input signals have some overlap in time, the two signals are said to be in coincidence and a digital output signal is generated

The main challenge in detector electronics is distinguishing the small signals from the noise. In detector systems, noise is any random signal that is not due to the physical process one intends to measure. If the noise is only present at the end of the electronics readout chain it is not a problem. One only needs to amplify until the signal is larger than the noise. However, if the noise is already present at the front-end part, at the level of the detector itself, amplification does not help since the noise is also amplified.

There are many possible causes of noise. Some of these can be reduced to arbitrarily low levels by careful design of the measurement system. A good example of such reducible noise is the pick-up noise. Some of the ubiquitous electromagnetic radiation can be captured by the front-end part of the measuring device, is amplified and is present as noise in the output signal. This noise can be due to external devices unrelated to the measuring system being used, but is often caused by the electronics of the detector itself. The digital part of the electronics and the readout computer are often sources of noise. Some level of pick-up noise is nearly always present in the measurement systems. One of the main technical difficulties in designing nuclear electronics is keeping the pick-up noise under control. The main method by which to achieve this is by enclosing the detector in a Faraday cage. A Faraday cage is simply a box made out of a good conductor, usually copper. A Faraday cage is very effective in suppressing pick-up noise. However, there are always lines entering the Faraday cage, for example, power lines or signal output lines, and particular care must be taken to avoid noise from entering the cage with such lines. Some commonly used methods to achieve this are illustrated in Fig. 8.4.

Consider the high-voltage input line shown in Fig. 8.4. One can think of the noise as an unwanted pulse travelling on this line. The problem with pick-up noise is usually with high-frequency signals and we therefore assume that we need to suppress high-frequency signals. If no protective measures are taken, a noise signal on the high-voltage line will arrive on the electronics board. Stray capacitances on the electronics board will inject a small fraction of this noise pulse into the input of the amplifier. In Fig. 8.4, we show how this can be avoided. If the high-voltage line is connected to the Faraday cage by a large capacitance, the amplitude of the noise pulse is attenuated in the ratio of the impedances. With a proper choice of the values of R and C, this strongly suppresses the noise.
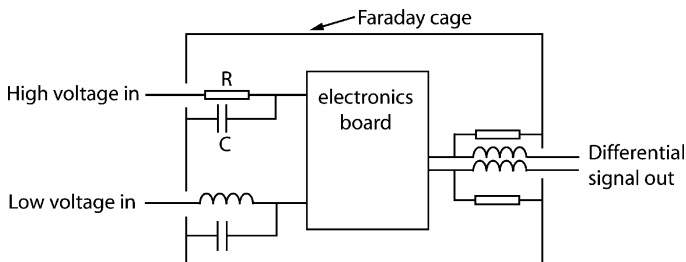

Fig. 8.4 Some methods commonly used for preventing the noise from entering a Faraday cage through signal cables or power supply lines

$$\text{Noise attenuation} = \left| \frac{\dfrac{1}{j\omega C}}{R + \dfrac{1}{j\omega C}} \right| \approx \frac{1}{\omega RC}.$$

Noise can also enter the Faraday cage through the signal output lines. One of the many possible ways to avoid this problem is illustrated in Fig. 8.4. In this solution, we use a differential output line. This means the signal and its opposite are sent on the two different lines. These two signals then go through two self-inductances that are wound together with the windings in the same direction. Because the currents in the differential line are always opposite, these coupled self-inductances have no effect on the signal. But any pick-up noise is the same on both lines, so for the noise this is seen as a real self-inductance with a large impedance. If we connect one side of the self-inductance to ground with a resistor that is not too large, we again attenuate any noise signal entering the Faraday cage. Usually the noise filters are not connected to the wall of the Faraday cage as shown on Fig. 8.4, but to the ground plane of the electronics board. This ground plane itself is connected to the Faraday cage by a low-impedance connection.

Other sources of noise can never be completely eliminated. It is necessary to understand these sources of noise and to minimise their influence on the measurement. The main sources of irreducible noise are the thermal noise of the resistors and the shot noise. The main task in designing nuclear electronics is optimising the signal-to-noise ratio and making the correct compromises for this. The designer of the detectors must also understand the implications of this to find the best detector for the problem at hand. Sections 8.4, 8.5 and 8.6 are devoted to a study of these noise effects. The discussion here follows the presentation of this subject in [1, 2].

Before we go into a more detailed analysis, I want to point out the basic reason why electronic amplification is always accompanied by noise. Consider the amplifier schematically represented in Fig. 8.5.

The detector generates some small current and with the capacitance of the detector this determines the voltage seen at the input of the amplifier. This voltage modulates the resistance of an amplifying device (usually a transistor) and this change in resistance changes the current in the output circuit and gives an output
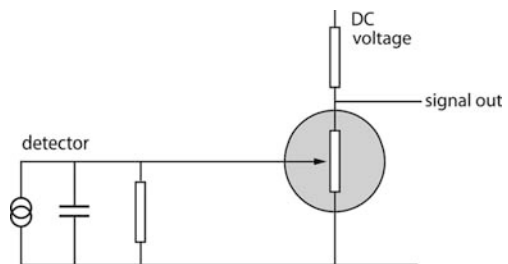

Fig. 8.5 Schematic representation of a detector and its amplifier

voltage over the load resistor. As we will show later, a resistor always has noise for fundamental physical reasons; therefore, this amplifier unavoidably introduces noise. We need to optimise things in such a way as to minimise the effect of this noise. It is immediately clear from the above that the capacitance of the detector should be kept small. The signal-to-noise ratio does not degrade inversely proportional to the capacitance as the above argument seems to suggest, but rather inversely proportional to the square root of the capacitance. The reason for this will become clear later.

8.2 Impulse Response and Transfer Function

To present a quantitative discussion of the electronic noise, we need some elements of signal theory and in particular the concepts of impulse response and transfer function. These concepts are introduced in the present section. Any amplifier, and more generally any electronic circuit, has an input impedance and an output impedance, as illustrated in Fig. 8.6.

This means that, if the input of the amplifier is part of some electronic circuit, it will behave as impedance Z_{in}. Similarly, if the output of the amplifier is part of some electronic circuit, it will behave as an impedance Z_{out} and a current or voltage source. Note that these impedances are in general complex, frequency-dependent, functions.

We now need to introduce the concept of 'linear circuit' and discuss the main properties of such circuits.

Definition of a linear circuit (see Fig. 8.7)

If V_{out1} is the output signal corresponding to the input signal V_{in1}

If V_{out2} is the output signal corresponding to the input signal V_{in2}

Fig. 8.6 The input of an amplifier behaves as an impedance and the output as a voltage source with an impedance in series
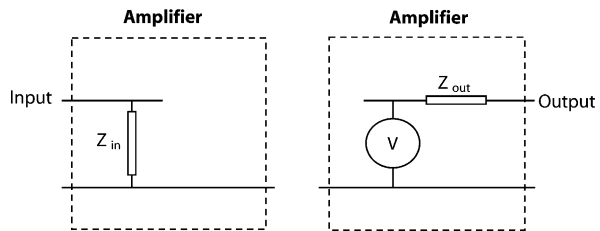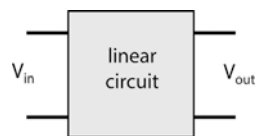

Fig. 8.7 A linear circuit has an input and an output line. It is linear if the signals satisfy the properties listed in the text

A circuit is linear if and only if the following property holds:

For any two arbitrary input signals V_{in1} and V_{in2}, to the input signal $(V_{in1} + V_{in2})$, corresponds the output signal $(V_{out1} + V_{out2})$.

Linear circuits are important, because they have simple mathematical properties. Most circuits used are therefore linear circuits. Any network of resistances, capacitances and self-inductances is a linear circuit. However, not all commonly used circuits are linear, for example a circuit with a diode is not a linear circuit.

Of course, a linear circuit is only linear in a certain range of signal amplitudes. For example, it could be linear only for positive signals of less than 5 V. However, if we make sure we only use the circuit in the linear range, we can safely apply all the results that are valid for linear circuits.

We now need to introduce the important concepts of 'impulse response' and 'transfer function'.

Impulse response. The impulse response $h(t)$ is the response of a system to a delta function like input pulse.

$$V_{in}(t) = \delta(t) \qquad V_{out}(t) = h(t)$$

Transfer function. The transfer function $H(\omega)$ is the Fourier transform of the impulse response.

$$H(\omega) = \frac{1}{\sqrt{2\pi}} \int e^{-j\omega t} h(t) \, dt$$

An amplifier, and more generally any electronic measurement system, deforms the input signal. If a delta function like voltage pulse is applied at the input of the system, the output pulse is not a delta function, but is a pulse with a finite width and usually is amplified or attenuated.

In this chapter, we follow the usual convention that 'j' is used to denote the complex number 'i'. In our notation, we hence have $j^2 = -1$. The conventions used in this text for the Fourier transform are made clear by the equations below:

$$h(t) = \frac{1}{\sqrt{2\pi}} \int e^{+j\omega t} H(\omega) d\omega$$

$$H(\omega) = \frac{1}{\sqrt{2\pi}} \int e^{-j\omega t} h(t) dt$$

With this notation, the well-known properties of the delta function are written as

$$\delta(t) = \frac{1}{2\pi} \int e^{j\omega t} d\omega \qquad f(a) = \int f(t)\delta(t-a) \, dt$$

And the inverse Fourier transform of the delta function is

$$\frac{1}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}} \int e^{-j\omega t} \delta(t) \, dt$$

It is now easy to prove the following properties of the impulse response and the transfer function for linear circuits.

(1) From the fact that $h(t)$ is real, it immediately follows that $H(\omega) = H^*(-\omega)$
(2) The transfer function $H(\omega)$ for $\omega = 0$ equals the integral of the impulse response

$$H(0) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} h(t)\, dt$$

(3) A perfect circuit without distortion and unit gain has $H(\omega) = \dfrac{1}{\sqrt{2\pi}}$. Obviously, no amplifier is perfect up to infinite frequency and above some value of the frequency the absolute value of $H(\omega)$ will drop to zero.
(4) If a pulse $V_{\text{in}}(t)$ with Fourier transform $V_{\text{in}}(\omega)$ is applied to the input of a linear circuit, the Fourier transform of the output is given by

$$V_{\text{out}}(\omega) = \sqrt{2\pi}\, V_{\text{in}}(\omega)\, H(\omega)$$

(5) For an arbitrary input signal, the output signal in the time domain can be obtained as follows

$$V_{\text{in}}(t) = \int\limits_{-\infty}^{+\infty} V_{\text{in}}(t')\delta(t' - t)\, dt'$$

Because the circuit is linear, the output signal is given by

$$V_{\text{out}}(t) = \text{gain} \int\limits_{-\infty}^{+\infty} V_{\text{in}}(t')\, h(t - t')\, dt'$$

There is a small difficulty here in that $\delta(t' - t) = \delta(t - t')$. But the two expressions

$$\int\limits_{-\infty}^{+\infty} V_{\text{in}}(t')h(t' - t)dt' \quad \text{and} \quad \int\limits_{-\infty}^{+\infty} V_{\text{in}}(t')h(t - t')dt'$$

are not the same. The impulse response $h(t)$ must be zero for all values $t<0$, otherwise there would be an output signal for an input signal that has not yet arrived. If we use the first integral, the output signal at time t depends on the part of the input signal coming after the time t. This makes no sense, therefore the second possibility should be used.

   Properties 4 and 5 allow us to obtain the response of the system to an arbitrary input function from the knowledge of the impulse response and the transfer function.
(6) If a sine wave $V_{\text{in}}(t) = \sin(\omega t)$ is applied at the input of the system, the output is given by (see Exercise 1):

$$V_{\text{out}}(t) = \sqrt{2\pi}\, |H(\omega)|\, \sin\left(\omega t + \phi(\omega)\right)$$

In this expression $|H(\omega)|$ and $\phi(\omega)$ are the modulus and the phase of the transfer function

$$H(\omega) = |H(\omega)| \, e^{j\phi(\omega)}$$

This result gives us an intuitive feeling of what the transfer function is. If a sine wave is applied at the input of a circuit, the output is also a sine wave, but the amplitude is proportional to $|H(\omega)|$ and the phase of the output sine wave relative to the input sine wave is given by $\phi(\omega)$.

(7) An ideal voltage amplifier, or operational amplifier, is an amplifier with infinite input impedance and zero output impedance. It is often represented by a triangle (see Fig. 8.8). If an electronic circuit is composed of two parts coupled by an operational amplifier, the transfer function of the complete system is given by

$$H(\omega) = \sqrt{2\pi} \, H_1(\omega) \, H_2(\omega)$$

This property is very useful when designing complex circuits. Indeed, an electronic circuit is often made up of a number of elementary sub-circuits connected by operational amplifiers.

Fig. 8.8 A shaping circuit often consistss of a succession of shaping networks connected by amplifiers
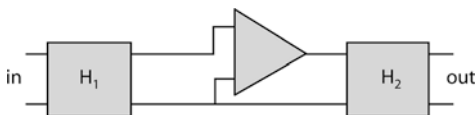
(8) Detectors in nuclear electronics are current sources. We therefore need to know the response of a system to a delta function like *current pulse*. This is the 'current impulse response' $\hat{h}(t)$. Similarly, the Fourier transform of the 'current impulse response' is the 'current transfer function' $\hat{H}(\omega)$. Where necessary we will use a '^' on top of the symbol $h(t)$ or $H(\omega)$ to distinguish the two different kinds of impulse response and transfer functions.

Assume a system with an input impedance $Z_{in}(\omega)$; if delta function like current pulse is applied at the input, the Fourier transform of the input voltage is given by

$$V_{in}(\omega) = Z_{in}(\omega) \, I(\omega) = Z_{in}(\omega) \, \frac{1}{\sqrt{2\pi}}$$

Using property 4, we immediately get the following relation between the 'transfer function' and the 'current transfer function' (see Fig. 8.9).

$$\hat{H}(\omega) = H(\omega) \, Z_{in}(\omega)$$

Fig. 8.9 Response of the
system to a delta function like
current pulse

$I_{in}=\delta(t)$ Z_{in} $V_{out}=FT^{-1}(H(\omega)\,Z_{in}(\omega))$
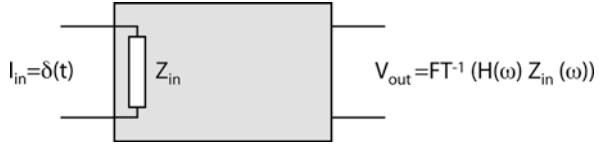
and therefore

$$V_{out}(t) = \frac{1}{\sqrt{2\pi}} \int e^{j\omega t}\, H(\omega)\, Z_{in}(\omega)\; d\omega$$

In the important case where the input impedance is a real and frequency inde-
pendent constant Z, we have $\hat{H}(\omega) = Z\, H(\omega)$ and $\hat{h}(t) = Z\, h(t)$ and the
difference between the two kinds of transfer function is only a difference in
gain. If we consider normalised impulse responses or transfer functions, the
two types of function therefore become identical.

(9) The Parseval identity is a relation between any function $f(t)$ and its Fourier
transform $H(\omega)$

$$\int\limits_{-\infty}^{+\infty} |h(t)|^2\, dt = \int\limits_{-\infty}^{+\infty} |H(\omega)|^2\, d\omega$$

From this identity we immediately obtain two useful properties of the transfer
function

$$\int\limits_{0}^{+\infty} |H(\omega)|^2 d\omega = \frac{1}{2} \int\limits_{-\infty}^{+\infty} h^2(t)dt$$

$$\int\limits_{0}^{\infty} \omega^2 |H(\omega)|^2 d\omega = \frac{1}{2} \int\limits_{-\infty}^{+\infty} \left(\frac{dh(t)}{dt}\right)^2 dt$$

The second equation is derived using Parseval's identity and the fact that the
Fourier transform of the derivative of $h(t)$ is given by $j\omega H(\omega)$.

We will now illustrate the power of these methods with two simple examples.
These examples correspond to very simple and commonly used circuits.

The integrator or low pass filter. The circuit shown in Fig. 8.10(a) is a low-
pass filter. You should imagine that this circuit is connected on the left-hand side to
a voltage source, hence to a circuit with zero impedance. In practice, this voltage
source will usually be an amplifier. You should also imagine that the output voltage
is measured with an ideal voltage meter, hence with a circuit with infinite input
impedance.

Use Ohm's law written in the frequency domain: $V(\omega) = Z(\omega)\, I(\omega)$. To calculate
the impedance of a circuit, use the familiar rules to combine resistors in series or in
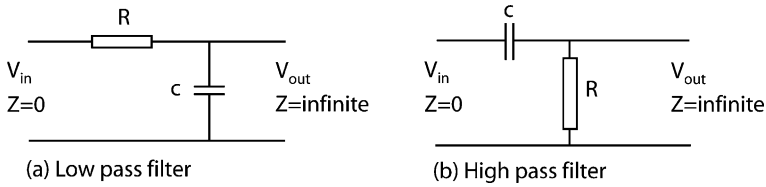
(a) Low pass filter   (b) High pass filter

Fig. 8.10 Two very simple circuits that are commonly used

parallel, but for any self-inductance L, consider it to be a complex impedance $= j\omega L$ and for any capacitance C, consider it to be a complex impedance $\dfrac{1}{j\omega C}$. In this way we readily find

$$V_{\text{in}}(\omega) = \left[R + \frac{1}{j\omega C} \right] I(\omega)$$

A second application of Ohm's law gives a relation between the Fourier transforms of the input and output voltages.

$$V_{\text{out}}(\omega) = I(\omega) \frac{1}{j\omega C} = \frac{V_{\text{in}}(\omega)}{\left[R + \dfrac{1}{j\omega C} \right]} \frac{1}{j\omega C} = V_{\text{in}}(\omega) \frac{1}{1 + j\omega RC}$$

If the input voltage is a delta function like voltage pulse, $V_{\text{in}}(\omega) = \dfrac{1}{\sqrt{2\pi}}$ and from the definition of the transfer function we immediately find

$$H(\omega) = \frac{1}{\sqrt{2\pi}} \frac{1}{1 + j\omega RC}$$

Taking the Fourier transform of the above transfer function, one finds for the impulse response

$$\begin{cases} h(t) = \dfrac{1}{RC} e^{\frac{t}{RC}} & t \geq 0 \\ h(t) = 0 & t < 0 \end{cases}$$

The actual calculation of the Fourier transform is rather involved, but this result can be found in any handbook with tables of Fourier transforms. In addition, it is not necessary to go through the Fourier transform calculation. The same result can be obtained quite simply by approximating the delta function by a square pulse with $\Delta V \times \Delta t = 1$. Taking the limit for $\Delta t \rightarrow 0$, one readily obtains the result above. Indeed, the voltage pulse of amplitude ΔV and duration Δt first charges the capacitor and immediately after the pulse the voltage over the capacitance is $1/RC$. This voltage decays, because the voltage source at the input side has zero output impedance. This decay is exponential with decay constant RC.

To obtain the response of the low-pass filter to an arbitrary and time-dependent input voltage, we use property (5) of the transfer functions

$$V_{\text{out}}(t) = \int\limits_{-\infty}^{+\infty} V_{\text{in}}(t')h(t - t')dt'$$

$$V_{\text{out}}(t) = \int\limits_{-\infty}^{t} V_{\text{in}}(t')\frac{e^{-\frac{t-t'}{RC}}}{RC}\,dt'$$

For a more intuitive derivation of this last result, consider the situation illustrated in Fig. 8.11. The input pulse can be seen as a sum of square pulses of duration Δt and amplitude $V_{\text{in}}(t')$. The output pulse corresponding to each of these input pulses is just the impulse response with a weight $[V_{\text{in}}(t')\,\Delta t]$. The total output at some point in time t is just the sum of all the *preceding* small pulses. In the limit $\Delta t \to 0$, this sum becomes an integral and we recover the result above.

If the duration of the pulse is short compared to the time constant RC, we have

$$V_{\text{out}}(t) = \frac{1}{RC} \int\limits_{-\infty}^{t} V_{\text{in}}(t')\,dt'$$

This explains the name integrator.

The differentiator or high-pass filter. The differentiator or high-pass filter is illustrated in Fig. 8.10(b). Repeating the same calculation as above for the high-pass filter, one finds

$$H(\omega) = \frac{1}{\sqrt{2\pi}}\frac{j\omega RC}{1 + j\omega RC} = \frac{1}{\sqrt{2\pi}}\left(1 - \frac{1}{1 + j\omega RC}\right)$$

Fig. 8.11 This figure illustrates the relation between the input voltage and the output voltage for an integrator circuit. The input signal can be seen as a sum of short square pulses, each causing the impulse response as an output signal. The total output signal is the sum of all prior input pulses
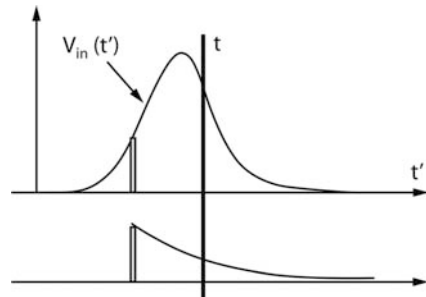
Notice that this transfer function simply is the constant $\dfrac{1}{\sqrt{2\pi}}$ minus the transfer function of the low-pass filter. Therefore, its Fourier transform is readily obtained from the previous calculation. It is a delta function minus the impulse response of the low-pass filter. The impulse response of a high-pass filter is therefore given by

$$\begin{cases} h(t) = \delta(t) - \dfrac{1}{RC}e^{-\frac{t}{RC}} & t > 0 \\[2mm] h(t) = \delta(t) & t < 0 \end{cases}$$

(a) Impulse response h(t) for a low pass filter with RC=1

(b) Transfer function H(ω) for a low pass filter

(c) Impulse response h(t) for a high pass filter with RC=1

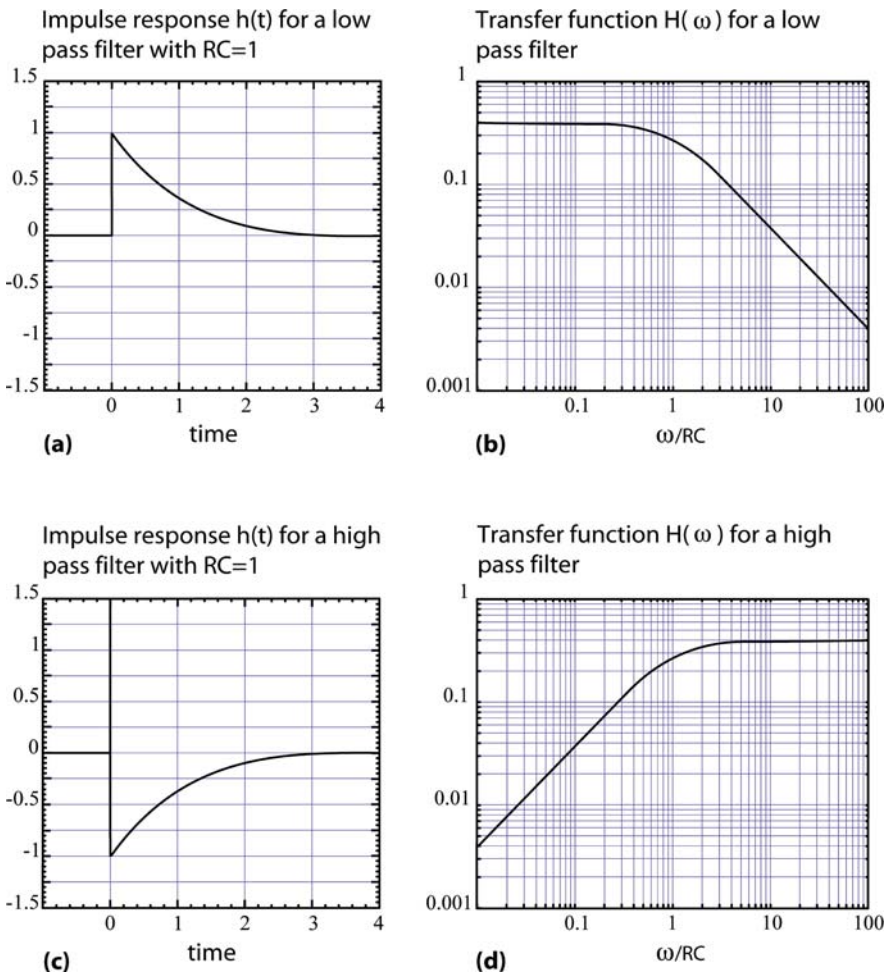(d) Transfer function H(ω) for a high pass filter

Fig. 8.12 Impulse response and transfer function for a low-pass filter and a high-pass filter

And the response to a voltage pulse is given by

$$V_{\text{out}}(t) = V_{\text{in}}(t) - \frac{1}{RC} \int\limits_{-\infty}^{t} V_{\text{in}}(t') e^{-\frac{t-t'}{RC}} \, dt'$$

If the time constant RC is much shorter than the duration of the pulse, only the values of $V_{\text{in}}(t')$ in the vicinity of $t' = t$ contribute to the integral and we can use a Taylor expansion of the function $V_{\text{in}}(t')$ around $t' = t$, using as expansion parameter $(t' - t)$.

$$V_{\text{out}}(t) = V_{\text{in}}(t) - \frac{1}{RC} \int\limits_{-\infty}^{t} \cdot \left[V_{\text{in}}(t) + \frac{dV_{\text{in}}(t)}{dt}(t' - t) + \ldots \right] e^{-\frac{t-t'}{RC}} \, dt'$$

With the following change of variables: $u = \dfrac{t - t'}{RC}$, this becomes:

$$V_{\text{out}}(t) = V_{\text{in}}(t) - V_{\text{in}}(t) \int\limits_{0}^{\infty} e^{-u} \, du + \frac{dV_{\text{in}}(t)}{dt} RC \int\limits_{0}^{\infty} u e^{-u} \, du + \ldots$$

$$V_{\text{out}}(t) \approx RC \frac{dV_{\text{in}}(t)}{dt}$$

This explains the name 'differentiator'. Figure 8.12 shows the impulse response and the transfer function for the low-pass filter and the high-pass filter.

8.3 Amplifiers for Particle Detectors

An amplifier contains a number of components arranged in such a way as to amplify a voltage at its input. Basically an amplifier contains a succession of circuits similar to the one shown in Fig. 8.13.

A bare amplifier would be almost useless. Its gain would be extremely sensitive to variations in the supply voltage and the temperature. In order to stabilise the
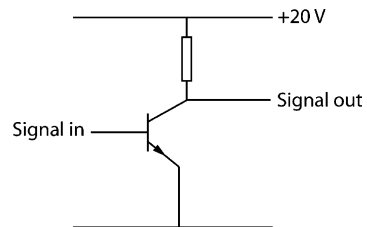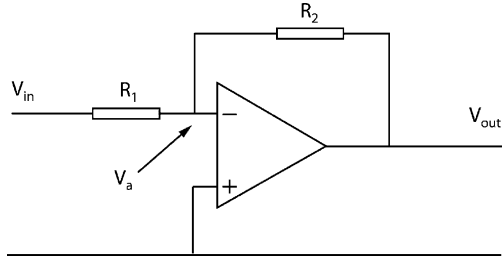

Fig. 8.13 A transistor is the basic element in an amplifier

Fig. 8.14 In a voltage
amplifier, the feedback
resistances R_1 and R_2
stabilise the gain

amplifier, it is necessary to have a feedback mechanism. A typical amplifier with
feedback is illustrated in Fig. 8.14.

In this figure, the triangle represents an amplifier without feedback. It has a large
gain, a high-input impedance and a low-output impedance. It has two inputs and it
amplifies the difference between the two voltages at these inputs. Usually one input
is connected to the ground and the other input receives the signal to be amplified.
Notice the symbols '+' and '−' (minus) at the two inputs. These indicate that the
amplifier is used in a reversing mode when connected as shown in Fig. 8.14. A pos-
itive input signal will give rise to a negative output signal. Therefore, this amplifier
has negative feedback.

If the input impedance of the open loop amplifier is very large compared to R_1
and R_2, the current flowing through these two resistances must be the same.

$$\begin{cases} I = \dfrac{V_{\text{in}} - V_a}{R_1} = \dfrac{V_a - V_{\text{out}}}{R_2} \\ V_{\text{out}} = -GV_a \end{cases} \tag{8.1}$$

For the time being, we assume that G is a real and positive number.
Eliminating V_a and I from Eq. (8.1), we find

$$V_{\text{out}} = \frac{R_2\, V_{\text{in}}}{\left[R_1 + \dfrac{(R_1 + R_2)}{G} \right]}$$

$$V_{\text{out}} \approx -\frac{R_2}{R_1}\, V_{\text{in}} \qquad \text{if } \frac{R_2}{R_1} << G$$

Eliminating V_a and V_{out}, we find

$$V_{\text{in}} = I\left(R_1 + \frac{R_2}{G+1} \right)$$

$$V_{\text{in}} = I\, R_1 \qquad \text{if } \frac{R_2}{R_1} = G$$

These equations show that if the open loop gain is large compared to the ratio
R_2/R_1, the gain of the amplifier with feedback is simply given by the ratio of the
two resistances and that the input impedance is equal to R_1. The amplifier with

feedback is very stable and its gain is hardly influenced by moderate changes in the supply voltage or the temperature.

However, the amplifier just described is not a useful amplifier for nuclear electronics. The main problem is the presence of resistances connected to input. As will be shown later in this chapter, such resistances give rise to noise and should be avoided. In addition, the gain is in general a complex function of ω. Moreover, this amplifier is well adapted for amplifying voltage signals and we are looking for an amplifier to be connected to a pulsed current source.

A possible way to avoid the problem of the resistor is to use a capacitor as the feedback element. However, at each pulse this capacitor will charge up and quickly the circuit will be out of the linear range. This needs a mechanism to reset the amplifier after each pulse. Special resetting mechanisms are sometimes used in very low-noise amplifier designs, thus completely avoiding the use of a feedback resistance. The most commonly used method for resetting the capacitor is to have a feedback resistor in parallel with the feedback capacitor as shown in Fig. 8.15. The resistor can have a very large value because it only needs to reset the capacitor after each pulse. This is called a charge-integrating amplifier.

Because of the presence of a capacitor, we need to use the Fourier transforms of the currents and voltages to calculate the gain and the input impedance for the amplifier shown in Fig. 8.15. We also need to take into account the fact that the naked amplifier itself also has a transfer function and this transfer function depends on the frequency and has a frequency-dependent phase. The constant open loop gain in the previous calculation has therefore to be replaced by a complex and frequency-dependent gain $G(\omega)$. The gain of a realistic amplifier can often be approximated by

$$G(\omega) = \frac{G_0}{1 + j\dfrac{\omega}{\omega_k}} \tag{8.2}$$

For frequencies below ω_k, the gain is real and constant. Around $\omega = \omega_k$ the phase turns by 90° and above ω_k the gain decreases like $1/\omega$. For the charge-integrating amplifier, we do a similar calculation as we did for the voltage amplifier shown in Fig. 8.14, but now working with the Fourier transforms of the voltages and currents.
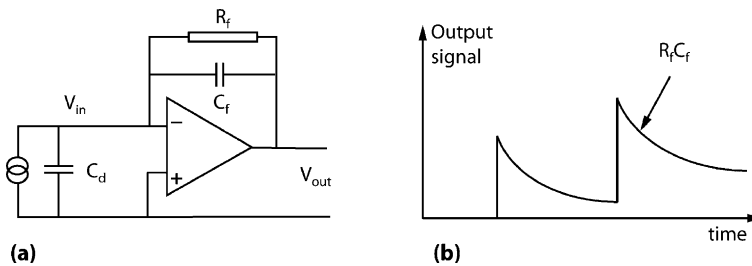

Fig. 8.15 (**a**) A charge-integrating amplifier. (**b**) Output of a charge-integrating amplifier if the input is a series of short current pulses

As before, we assume that the absolute value of the gain and the input impedance of the naked amplifier, are very large. We have the following two equations

$$\begin{cases} V_{\text{out}}(\omega) = -G(\omega)V_{\text{in}}(\omega) \\ V_{\text{in}}(\omega) - V_{\text{out}}(\omega) = \left[\dfrac{1}{R_f} + j\omega C_f\right]^{-1} \cdot I_{\text{in}}(\omega) \end{cases} \tag{8.3}$$

Eliminating V_{in} from the two equations above gives

$$V_{\text{out}}(\omega) = -\frac{G(\omega)}{G(\omega)+1} \cdot \frac{R_f}{1+j\omega C_f R_f} \cdot I_{\text{in}}(\omega) \approx -\frac{R_f}{1+j\omega C_f R_f} I_{\text{in}}(\omega)$$

The last equation gives us the relation between the output voltage and the input current. If the input current is a delta pulse, $I_{\text{in}}(\omega) = \dfrac{1}{\sqrt{2\pi}}$, then the resulting output voltage is given by

$$V_{\text{out}} = \hat{H}(\omega) = \frac{-R_f}{\sqrt{2\pi}} \frac{1}{1+j\omega C_f R_f}$$

We notice that the current transfer function for a charge-integrating amplifier is the same as the transfer function of a low-pass filter multiplied by a factor $-R_f$! The response of a charge-integrating amplifier to a delta function like current pulse is therefore given by

$$\begin{cases} h(t) = -\dfrac{1}{C_f} e^{-\frac{t}{R_f C_f}} & t \geq 0 \\ h(t) = 0 & t < 0 \end{cases}$$

A charge-integrating amplifier will behave as an amplifier with a 'current to voltage gain' at low frequency equal to R_f. Its response to a current pulse will be a sharp rising edge at the moment the pulse arrives, followed by an exponential decay with time constant $R_f C_f$.

The input impedance of the charge-integrating amplifier is obtained by eliminating V_{out} from the two equations (8.3):

$$V_{\text{in}}(\omega) = \frac{1}{G(\omega)+1} \frac{R_f}{1+j\omega C_f R_f} \cdot I_{\text{in}}(\omega) \approx \frac{1}{G(\omega)} \frac{R_f}{1+j\omega C_f R_f} \cdot I_{\text{in}}(\omega)$$

Using this result together with Eq. (8.2), we get

$$Z_{in}(\omega) = \frac{V_{in}(\omega)}{I_{in}(\omega)} = \frac{R_f}{G_0} \frac{\left(1 + j\dfrac{\omega}{\omega_k}\right)}{\left(1 + j\omega R_f C_f\right)}$$

$$Z_{in}(\omega) = \frac{R_f}{G_0} \frac{\left(1 + j\dfrac{\omega}{\omega_k}\right)\left(1 - j\omega R_f C_f\right)}{\left(1 + \omega^2 R_f^2 C_f^2\right)}$$

$$Z_{in}(\omega) = \frac{R_f}{G_0} \frac{\left(1 + j\dfrac{\omega}{\omega_k} - j\omega R_f C_f + \dfrac{\omega^2}{\omega_k} R_f C_f\right)}{\left(1 + \omega^2 R_f^2 C_f^2\right)}$$

For low frequencies this impedance is a real number; it behaves like a pure resistor. However, as the frequency increases the impedance becomes self-like or capacity-like, depending on the value of the parameters. If we choose the value of the feedback capacitance and feedback resistor such that $R_f C_f = (1/\omega_k)$, the two imaginary parts cancel and the impedance becomes a real number for all values of ω. In fact, this condition assures that the feedback is negative for all frequencies and is a necessary condition for the system to be stable. With this condition the input impedance of the system simply becomes

$$Z_{in}(\omega) = \frac{R_f}{G_0}$$

The input impedance is a real and frequency independent constant. Moreover, this impedance will not be very large. These are highly desirable properties for an amplifier for particle detection. The input impedance of the amplifier must be large compared to the internal impedance of the detector itself.

In Sect. 8.5, we will show that the feedback resistance is a source of noise. In order to minimise this noise it is essential that the feedback resistance is as large as possible. On the other hand, the feedback capacitance cannot be made arbitrarily small. It should always be sufficiently large compared to any stray capacitances that are unavoidably present in the system. In practice, it is difficult to have this capacitance less than about 1 pF. The result is that $R_f C_f$ must be large if we want to have low noise. Typically, this time constant will be several 100 µs.

The output pulses of the amplifier shown in Fig. 8.15 are very long pulses and this will severely limit the count rate capability of the counter. We need a way to make the output pulses short while keeping the product $R_f C_f$ large. This can be achieved with a shaping stage after the charge-integrating amplifier as shown in Fig. 8.16. The total transfer function of this amplifier is simply given by the product of the transfer functions of a charge-integrating amplifier and the transfer functions of a high-pass and a low-pass filter.
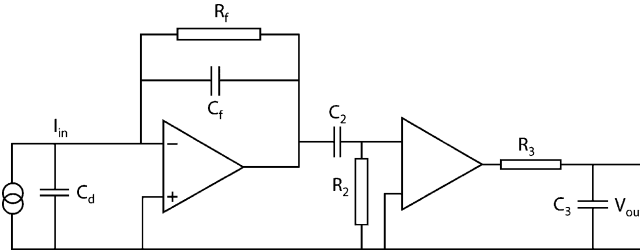
Fig. 8.16 Charge-integrating amplifier with *CR–RC* shaping

$$\hat{H}(\omega) = \frac{-R_f}{\sqrt{2\pi}} \frac{1}{1 + j\omega R_f C_f} \frac{j\omega R_2 C_2}{1 + j\omega R_2 C_2} \frac{1}{1 + j\omega R_3 C_3}$$

If we chose the values of the different capacitances and resistances such that

$$\tau = C_2 R_2 = C_3 R_3 \text{ and } \tau << \tau_f = C_f R_f$$

Omitting the factor '$-R_f$', this transfer function can be written as

$$\hat{H}(\omega) = \frac{1}{\sqrt{2\pi}} \frac{1}{1 + j\omega \tau_f} \frac{j\omega \tau}{(1 + j\omega \tau)^2} \tag{8.4}$$

Taking the Fourier transform of the above expression (as can be obtained with some calculations starting from a table of Fourier transforms, see Exercise 2) one gets

$$\begin{cases} \hat{h}(t) = -\dfrac{1}{\tau(\tau_f - \tau)^2} \left[(\tau^2 + t(\tau_f - \tau)) e^{-t/\tau} - \tau^2 e^{-t/t_f} \right] & t \geq 0 \\ \hat{h}(t) = 0 & t < 0 \end{cases}$$

If $\tau << \tau_f$ and for values of t of the order of τ, we have

$$\hat{h}(t) \approx \frac{t}{\tau \tau_f} e^{-t/\tau}$$

The output pulses of this amplifier are shown in Fig. 8.17(a). We have indeed managed to keep the feedback resistor R_f large, while at the same time producing short output pulses. This $RC–CR$ shaping is often used because of its simplicity. Its biggest drawback is the long negative tail after the main pulse. At high rates this can cause a baseline shift. In particular, if one wants to use this circuit for measuring pulse heights, it is totally unacceptable because it will broaden all peaks and in this way ruin the energy resolution.
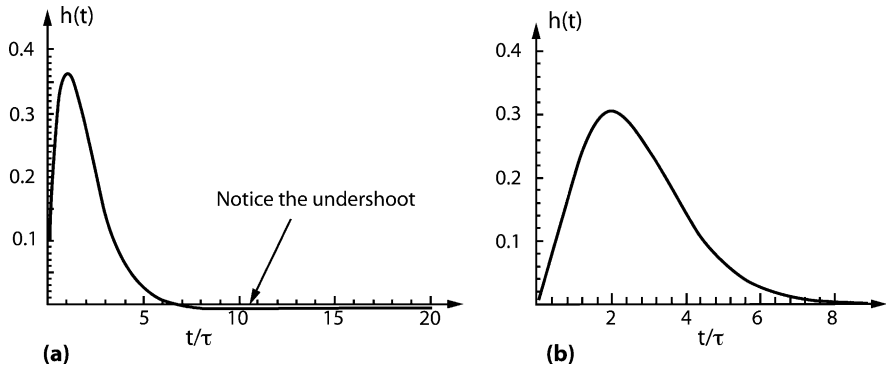
Fig. 8.17 (a) Impulse response of a charge-integrating amplifier with a simple $CR-RC$ shaping stage. The impulse response in this figure was multiplied by τ_f/τ, such that the integral over the positive part of the function approximately equals one. (b) Output pulse of a charge-integrating amplifier with pole zero cancellation for a delta function like input pulse

From the expression of the Fourier transform, one can see that this undershoot is caused by the factor $\left(\dfrac{1}{1+j\omega\tau_f}\right)$ in the transfer function. This introduces a pole close to $\omega = 0$. If we find a way to cancel this factor, the undershoot will be removed. This can be achieved with a somewhat more complicated amplifier design and is referred to in the literature as 'pole zero cancellation'. Figure 8.18 shows a charge-integrating amplifier with 'pole zero cancellation' and a simple shaping stage. The total transfer function of this amplifier is again simply the product of the contributions of the three parts. With proper choice of the values for the components, the pole in the factor $\dfrac{R_f}{1+j\omega\,C_f R_f}$ is exactly cancelled by the factor $(1 + j\omega\,R_0 C_1)$. In this way, the undershoot is removed. The last factor represents the shaping circuit. With this particular choice of the shaping stage and with a proper choice of the values of the components, the resulting transfer function can be written as

$$\hat{H}(\omega) = \frac{1}{1+j\omega\tau}\cdot\frac{1}{1+(5/3)j\omega\tau - \omega^2\tau^2}$$

The Fourier transform of this last expression is shown in Fig. 8.17(b). This last figure was obtained with a numerical Fourier transform and the program Mathematica.

Many different shaping circuits can be used and this is only one particular example. A popular shaping circuit is the $CR-(RC)^4$ shaping. This is a CR circuit followed by 4 RC circuits connected through amplifiers. This produces a nearly Gaussian output pulse with a time structure given by

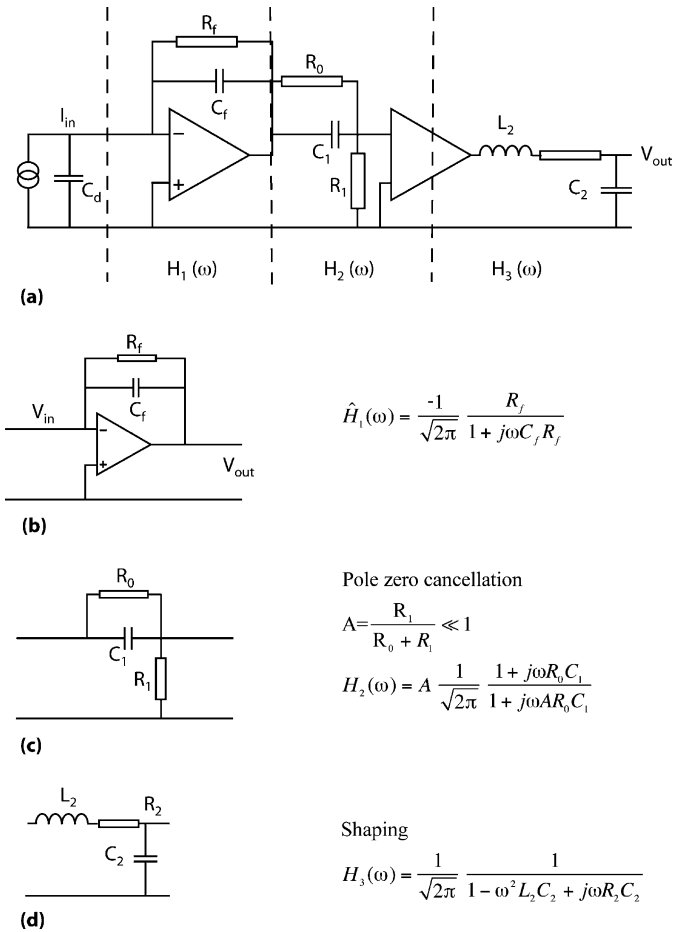$$V_{\text{out}}(t) \propto \left(\frac{t}{\tau}\right)^4 e^{-t/\tau}$$

Fig. 8.18 (a) Charge-integrating amplifier with pole zero cancellation and a shaping stage. This figure shows one among many possibilities for the shaping stage. The transfer function for this amplifier is the product of the three parts shown in (a), (b), (c). With the proper choice of components the pole zero will be cancelled

The results obtained above are valid provided the open loop gain is well represented by Eq. (8.2). Obviously, above some large angular frequency, ω_{max}, the gain will drop faster than expression (8.2) and the results above no longer apply. In particular, it is not possible to make the output pulse shorter than $\approx 2\pi/\omega_{max}$, therefore limiting the maximum rate the amplifier can handle.

The amplifier we have just described is very well suited for measuring weak and fast electrical pulses. It has a small and real input impedance and the feedback resistor can be chosen very large, minimising the noise.

8.4 The Thermal Noise of a Resistor

Any piece of matter is made up of electrons and nuclei. These charges are constantly in motion owing to the thermal agitation. It is obvious that these motions will induce small voltage and current fluctuations in any piece of material and in particular in any resistor.

A priori we do not know what these noise signals will look like, but the Norton and Thevenin theorems tell us what the electrical equivalent of the noise sources will be like. The Thevenin theorem states that any two-terminal network of resistors and voltage sources is equivalent to a single resistor in series with a single ideal voltage source. The Norton theorem states that it is also equivalent to a single resistor in parallel with ideal current source. An ideal voltage source has zero internal resistance and an ideal current source has infinite internal resistance. Obviously, the value of the resistor is the same and the current and voltage sources are related by $V = RI$.

From the Thevenin and Norton theorems, we expect that the thermal noise will be equivalent to a small current source or a small voltage source in parallel or in series with the resistor as indicated in Fig. 8.19. The average value of these voltages or currents will be zero, but at any particular instant in time we expect to measure a small but non-zero value of the voltage or current. The mean square deviation (r.m.s.) will be different from zero.

$$\left\langle V_{\text{noise}}^2 \right\rangle \neq 0; \ \left\langle I_{\text{noise}}^2 \right\rangle \neq 0$$

As usual the square brackets $\langle x \rangle$ denote the average of x. From the central limit theorem, we also know that these fluctuations will have a Gaussian distribution.

The voltage noise or current noise is completely characterised by its mean square deviation. The first important point to be made is that this noise does not depend on the nature of the resistor, it only depends on the value of the resistance. To see this consider a thought experiment illustrated in Fig. 8.20.

Two resistors with the same value of the resistance but made from different materials are connected as shown in Fig. 8.20. Each resistor is enclosed in a thermally isolated box. Assume that for a moment the resistor in the left box has noise characterised by a given $\left\langle I_{\text{noise}}^2 \right\rangle \neq 0$ and the resistor in the right box has no noise. The current generator in the left box will induce equal currents in the two resistances and in this way induce heat in each resistor. The energy to produce this heat is
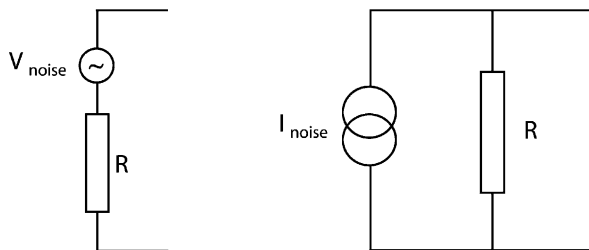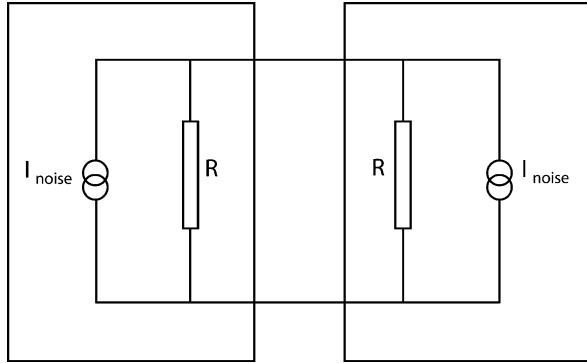

Fig. 8.19 Equivalent networks for the thermal noise in a resistor. The two diagrams above will be equivalent if $V_{\text{noise}} = R\, I_{\text{noise}}$

Fig. 8.20 The second law of thermodynamics requires that the thermal noise of all resistors with the same value of the resistance, is same

extracted from the left box by the noise current generator so, with time, the left box will become cold and the right box will become hot. This situation would be in total contradiction to the second law of thermodynamics. The only way out is to assume that both resistors have exactly the same noise. Following the same argument but considering two different resistors, one readily finds that $\langle I^2_{noise}\rangle \propto \dfrac{1}{R}$ and $\langle V^2_{noise}\rangle \propto R$.

We are now only left with the task of finding the proportionality constant. If there is one particular resistor for which we can calculate the thermal noise, we have solved the problem. There is indeed such a device, namely the ideal transmission line. We will show that an ideal transmission line will behave like a purely ohmic resistor with a resistance equal to its characteristic impedance. Moreover, we will show that it is possible to calculate the thermal noise in a transmission line.

A transmission line is a very important concept in fast electronics. We tend to think of a connecting wire in an electronics circuit as something that has negligible capacitance and negligible self-inductance and where any voltage applied at one end of the wire is immediately present at the other end. As faster and faster signals are used, this assumption becomes less and less valid. A practical rule of thumb is the following: we can think of connecting lines in an electronics circuit in the conventional way as long as the length of the wires is less than 2% of the rise time of the signal multiplied by the speed of light. In nuclear detectors, electronics pulses as short as 10 ns are common. For such pulses the maximum allowable wire length is 6 cm! If any longer wires are used the behaviour of the circuit will be totally unpredictable. To transport a signal over a longer distance one needs to use a transmission line!

Transmission lines come in many variants, but as far as nuclear electronics is concerned the most common form of a transmission line is the shielded coaxial cable. A coaxial cable is shown schematically in Fig. 8.21.

This geometry serves the following purposes: the braided shield protects the inner conductor carrying the signal from any pick-up noise and the structure of the wire makes it behave as a transmission line.

For the purpose of the present argument we will consider an ideal transmission line with a capacitance per unit length C and a self-inductance per unit length *of*
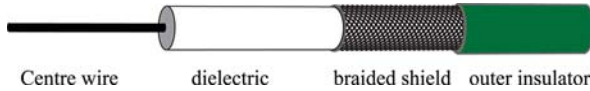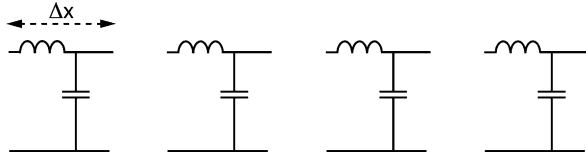
Fig. 8.21   A typical coaxial cable. The braided shield is made from tightly woven fine wires such as to allow the cable to remain flexible

Fig. 8.22   A transmission line can be seen as a succession of identical infinitesimal networks like the ones shown here

L. Capacitance and self-inductance are constant along the line. Figure 8.22 shows a transmission line represented as a set of discrete components. An ideal transmission line has negligible ohmic resistance in the conductors and no leakage currents. While leakage currents are usually indeed negligible, the resistance of the wire is usually not negligible. This resistance causes an attenuation of the signal. The exact value of this attenuation depends on the details of the structure of the line, but an attenuation length of the order of 100 m is typical at a frequency of $\approx$100 MHz. An ideal transmission line without resistance does not exist. It is nevertheless useful for the purpose of the present argument.

A transmission line can be viewed as a succession of a large number of sections of length Δx, each with a capacitance $C\Delta x$ and a self-inductance $L\Delta x$. In the limit $\Delta x \to 0$ this will behave like a real transmission line.

Consider one small section of the line with length Δx. A capacitor and a self-inductance in the time domain is described by a first order differential equation. Over this section the change in voltage and the change in current are given by

$$\begin{cases} \Delta V = -L\Delta x \dfrac{dI}{dt} \\[2mm] \Delta I = -C\Delta x \dfrac{dV}{dt} \end{cases}$$

Taking the limit for $\Delta x \to 0$, we obtain the following two differential equations for an ideal transmission line.

$$\begin{cases} \dfrac{dV(x,t)}{dx} = -L\dfrac{dI(x,t)}{dt} \\[3mm] \dfrac{dI(x,t)}{dx} = -C\dfrac{dV(x,t)}{dt} \end{cases}$$

These equations hold at any point along the line. From these two first order equations we obtain the following two second order equations:

$$\begin{cases} \dfrac{d^2\, V(x,t)}{dx^2} = LC\dfrac{d^2\, V(x,t)}{dt^2} \\[4mm] \dfrac{d^2 I(x,t)}{dx^2} = LC\dfrac{d^2 I(x,t)}{dt^2} \end{cases}$$

The current and the voltage at each point along the line satisfy the string equation! One readily verifies that any function of the variable $(x - t/\sqrt{LC})$ or of the variable $(x+t/\sqrt{LC})$ is a solution to the string equation. The most general solution takes the form

$$V(x,t) = f_1\left(x - \frac{t}{\sqrt{LC}}\right) + f_2\left(x + \frac{t}{\sqrt{LC}}\right)$$

$$I(x,t) = \sqrt{\frac{C}{L}}f_1\left(x - \frac{t}{\sqrt{LC}}\right) - \sqrt{\frac{C}{L}}f_2\left(x + \frac{t}{\sqrt{LC}}\right)$$

In these equations f_1 and f_2 are arbitrary functions of one variable. This solution represents a sum of two waves travelling at a velocity $v_0 = \dfrac{1}{\sqrt{L.C}}$, one wave travelling towards increasing values of x, the other wave travelling towards deceasing values of x. The actual shape of these waves will be determined by the boundary conditions. If we apply a variable voltage to one end of the line, this signal will travel along the line with a velocity v_0. If only one wave is present, the current and the voltage at any point along the line are related by

$$V = \sqrt{\frac{L}{C}}\, I.$$

This will also apply at the end of the line. Hence, if we apply a voltage V at the end of the line, we will induce a current in the line given by the equation above; therefore, the line will behave as a resistor with resistance equal to the characteristic impedance of the line

$$Z_0 = \sqrt{\frac{L}{C}}.$$

This is a remarkable result because this impedance is a real number; this impedance is purely ohmic! Of course this is only true if the only wave present in the transmission line is the wave travelling away from the measurement point. After some time this wave will reach the other end of the line and be reflected back. When

this reflected wave reaches the measurement point the relation no longer holds. Only an infinitely long transmission line will behave like a true resistor.

Because of the arguments at the beginning of this section, the thermal noise of the line will be the same as the thermal noise of a resistor with the same value of the resistance. But for this unusual resistor it is possible to calculate the thermal fluctuations! The voltage $V(x,t)$ and the current $I(x,t)$ satisfy the string equation. A string in thermal equilibrium with its surroundings will vibrate at all the stationary vibration modes of this string. Assume a transmission line with length D and with open ends. These stationary solutions are of the form $f(x).g(t)$ and the boundary conditions are that the current at the beginning and the end of the line should be zero: $I(x = 0,t) = I(x = D,t) = 0$.

We now look for stationary solutions of the form:

$V(x,t) = V'(x) \cdot V''(t)$

$I(x,t) = I'(x) \cdot I''(t)$

The stationary solutions can be found to be (see Exercise 6):

$$
\begin{cases}
I_n(x,t) = I_n \sin\left(\dfrac{n\pi x}{D}\right) \sin\left(\dfrac{n\pi v_0 t}{D} + \varphi_n\right) \\
V_n(x,t) = I_n Z_0 \cos\left(\dfrac{n\pi x}{D}\right) \cos\left(\dfrac{n\pi v_0 t}{D} + \varphi_n\right) \quad n = 1, ..\infty
\end{cases}
$$

These are the well-known stationary vibration modes of the string, I_n and φ_n are arbitrary constants to be determined by the boundary conditions. Figure 8.23 shows the solutions for $n = 1, 2$ and 3.

The energy contained in a wave of wave number n is given by

$$
E_n = \frac{1}{2} \int_0^D (CV^2 + LI^2)\, dx = \frac{I_n^2}{4} DL
$$

This relation allows us to express the amplitude of the wave as a function of the energy of the wave. The voltage that will be observed at the end of the transmission line is given by

$$
V(x = 0) = \sum_n V_n = \sum_n \sqrt{\frac{4E_n}{DC}} \cos\left(\frac{n\pi v_0 t}{D} + \varphi_n\right)
$$

The average of this voltage is zero and the average square voltage is obtained by averaging these elementary noise waves over the amplitudes and over time. Let us first consider the time averaging only
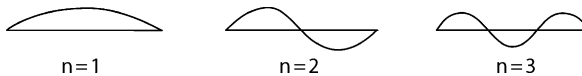

<center>n=1   n=2   n=3</center>

Fig. 8.23 The fundamental frequency and the first two harmonics of a vibrating string. The value of the current in an ideal transmission line with open ends follows the same pattern

$$\left\langle V^2(x = 0)\right\rangle = \frac{1}{T} \int\limits_{-T/2}^{+T/2} V^2(x = 0)\, dt = \frac{1}{T} \int\limits_{-T/2}^{+T/2} \sum_n \sum_{n'} V_n V_{n'}\, dt$$

The integral is to be taken in the limit that T goes to infinity. In calculating this time average, all the terms where n is different from n' vanish. To see this just remember that $\cos(\omega_1 t)\cos(\omega_2 t) = \cos[(\omega_1 + \omega_2)t] + \cos[(\omega_1 - \omega_2)t]$. The sum reduces to

$$\left\langle V^2(x = 0)\right\rangle = \frac{1}{T} \int\limits_{-T/2}^{+T/2} \sum_n V_n^2\, dt$$

$$= \sum_n \frac{4E_n}{DC} \frac{1}{T} \int\limits_{-T/2}^{+T/2} \cos^2\left(\frac{n\pi v_0 t}{D} + \varphi_n\right) dt$$

Using $\cos^2(x) = \frac{1}{2}[1 + \cos(2x)]$ this becomes

$$\left\langle V^2(x = 0)\right\rangle = \sum_n \frac{2E_n}{DC}$$

If the transmission line is in thermal equilibrium with its surroundings, each mode of vibration will be present with a random amplitude and a random phase. According to the equipartition theorem, each quadratic term in the Hamiltonian contributes $kT/2$ to the energy of the system. The Hamiltonian per unit volume for the electromagnetic wave is given by $\frac{1}{4\pi}(E^2 + B^2)$. Therefore, each mode of vibration will acquire an average energy $<E_n> = kT$, where k is the Boltzmann constant and T is the absolute temperature. The average noise voltage that will be observed at the end of the line is therefore given by

$$\left\langle V^2(x = 0)\right\rangle = \sum_n \frac{2kT}{DC}$$

The sum runs over all frequencies that are present in the system, each contributing the same amount to the total noise.

We can now calculate the noise contribution of all the waves with angular frequency in the interval $(\omega, \omega + d\omega)$. Since $\omega = \frac{n\pi v_0}{D}$, the number of waves in the interval $d\omega$ is given by $dn = \frac{D}{\pi v_o}d\omega$; this noise contribution is given by

$$\left\langle V^2(x = 0)\right\rangle = \frac{2}{\pi}kTZ_0\, d\omega$$

Notice that the length of the line D has disappeared from the equation and this equation also applies for an infinitely long transmission line. However, this noise is observed by an instrument that is characterised by a transfer function $H(\omega)$ and from property 6 of the transfer functions we know that the observer sees, for each wave with angular frequency ω, an amplitude $\sqrt{2\pi}|H(\omega)|$.

Hence the total 'observed' noise, integrated over all frequencies is given by

$$\left\langle V_{\text{noise}}^2 \right\rangle = 4kTZ_0 \int\limits_0^{+\infty} |H(\omega)|^2 \, d\omega$$

At the end of the open transmission line we will measure a noise voltage given by the formula above. We can hence conclude that for any resistor the noise voltage and the noise current are characterised by

$$\begin{cases} \left\langle V_{\text{noise}}^2 \right\rangle = 4kTR \int\limits_0^{+\infty} |H(\omega)|^2 \, d\omega & (8.5) \\[2ex] \left\langle I_{\text{noise}}^2 \right\rangle = \dfrac{4kT}{R} \int\limits_0^{+\infty} |H(\omega)|^2 \, d\omega. & (8.6) \end{cases}$$

In this equation k is the Boltzmann constant and T is the absolute temperature. Equations (8.5) and (8.6) are in fact integrals over the current impulse response, but this distinction is usually not made in the literature, and will also not be made here.

Equations (8.5) and (8.6) should be understood as follows. The voltage over a resistor has a fluctuating value. If this voltage is measured at random moments, every time a different value will be obtained. Since this noise originates from a large number of random fluctuations, the voltages will have a Gaussian distribution. The average of this distribution is zero and the expression above represents the variance of this distribution. The square root of this quantity is the standard deviation (or r.m.s.) of the current or voltage noise. This situation is illustrated in Fig. 8.24.

Each frequency interval contributes in the same way to the noise. This is so-called 'white noise'. If the measurement system has a unit gain up to a maximum frequency, f_{max}, and then has a sharp cutoff, we have

$$\int\limits_0^{+\infty} |H(\omega)|^2 \, d\omega = f_{\text{max}} = \text{Band width of the circuit}$$
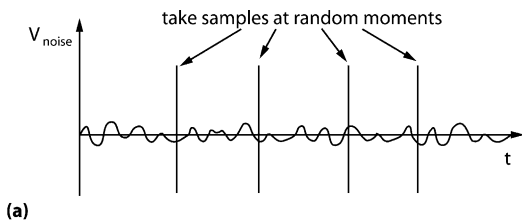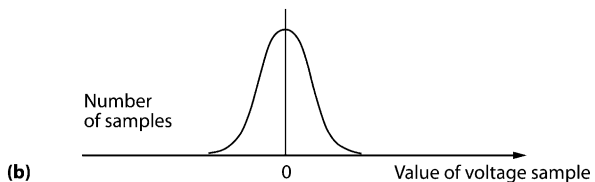

Fig. 8.24 Pulse samples taken at random times have a Gaussian amplitude distribution. The r.m.s. of this distribution is a measure of the noise present in the system

The mean square noise is proportional to the bandwidth and hence the r.m.s. noise is proportional to the square root of the bandwidth.

Let us illustrate this with a numerical example: at room temperature ($\approx$293 K) $4kT = 1.62 \times 10^{-20}$[V^2/(Hz.$\Omega$)]. The thermal noise of a 1 $M\Omega$ resistor at room temperature measured with a voltmeter with a bandwidth of 100 MHz has an r.m.s. voltage noise of 1.27 mV. Assume a detector with a fairly typical capacitance of 30 pF. It needs about 2×10^5 electron charges to produce a voltage of 1 mV over this detector. This example makes it abundantly clear that the thermal noise is going to be an essential consideration in nuclear electronics.

Using property 9 of the transfer functions we can also write the noise as a function of the impulse response in the time domain:

$$\begin{cases} \langle V_{\text{noise}}^2 \rangle = 2kTR \int\limits_{-\infty}^{+\infty} h(t)^2 \, dt \\[2em] \langle I_{\text{noise}}^2 \rangle = \dfrac{2kT}{R} \int\limits_{-\infty}^{+\infty} h(t)^2 \, dt \end{cases}$$

Equation (8.5) seems to imply that, as the value of the resistor goes to infinity, the voltage noise over this resistor will also go to infinity. While this is strictly speaking correct, one will never measure an infinite noise voltage because the input impedance of the measuring device will always have some large but finite value. If R_{inp} denotes the value of the input impedance of the voltage meter, the measured voltage will be given by

$$\sqrt{\langle V_{\text{noise}}^2 \rangle_{\text{measured}}} = \sqrt{\langle V_{\text{noise}}^2 \rangle} \frac{R_{\text{inp}}}{R + R_{\text{inp}}} = \frac{R_{\text{inp}}}{R + R_{\text{inp}}} \sqrt{4kTR \int\limits_{0}^{+\infty} |H(\omega)|^2 \, d\omega}$$

Taking now the limit R to infinity, one sees that the *measured* noise voltage goes to zero. Similarly, if one measures the noise current over a zero resistance, the measured value will be zero.

8.5 Resistor and Transistor Noise in Amplifiers

Any resistor in the detector readout electronics contributes to the noise, but obviously, the resistors in the front-end part, close to the detector itself, will make the biggest contribution to the noise. We have already seen in previous sections that there are at least two unavoidable resistances in the front-end part of the readout electronics. First, there is the resistance of the detector itself and there is the feedback resistance. This feedback resistance is connected on one side to the input of the amplifier and on the other side to the output of the amplifier. Since the amplifier has a low output impedance, from the noise point of view, this is the same as a resistor between the input and the ground. In both cases this behaves as a resistor that

is in parallel with the detector itself. These are collectively referred to as a parallel resistor. Below we shall study what the noise effect of such a parallel resistor will be. We also consider the effect of a resistor between the detector and the input of the amplifier. There are often technical reasons to have such a resistor. Such a series resistor has a different noise effect, as will become clear below. Finally, we discuss the noise contribution stemming from the resistance of the front-end transistor itself.

8.5.1 Noise Contribution of a Parallel Resistor or a Series Resistor

Consider a charge amplifier with a resistor in parallel to the detector element, as illustrated in Fig. 8.25. The amplifier has a real and frequency independent input impedance and this impedance is small compared to the internal resistance of the detector. We also consider the output signal normalised to unit gain. Therefore, it is not necessary to distinguish the current and voltage impulse response or transfer functions. The parallel resistor R_p will generate a noise current given by

$$\left\langle I_{\text{noise}}^2 \right\rangle = \frac{4kT}{R_p} \int\limits_{0}^{+\infty} |H(\omega)|^2 \, d\omega$$

What matters is how the noise generated by the resistor compares to the detector signal. The detector is a current source and the signal has a certain amount of charge, usually expressed as a certain number of electrons. If we assume that this charge signal is generated in a short time, the response of the amplifier to a signal is equal to the impulse response. The noise is usually expressed by the quantity 'equivalent noise charge' (ENC), which is defined as a hypothetical charge produced in the detector that gives a peak output response equal to the r.m.s. of the noise. The concept of 'equivalent noise charge' is illustrated in Fig. 8.27.

If the current impulse response of the amplifier to a unit current pulse is given by $h(t)$, the response to a charge Q is given by $Qh(t)$ and the maximum value reached by this pulse is given by $Qh_{\max}(t)$. From the definition of ENC we therefore have
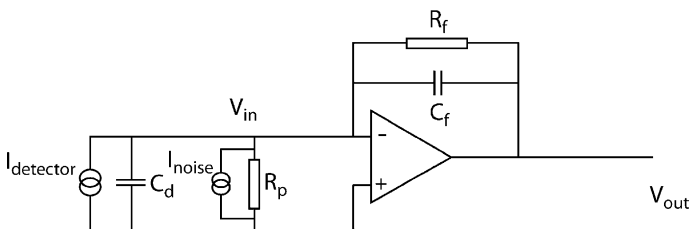

Fig. 8.25   A charge-integrating amplifier with a resistor in parallel with the detector

$$(ENC_p \, h_{max})^2 = \langle I^2_{noise} \rangle$$

$$ENC_p^2 = \frac{4kT}{R_p} \frac{1}{h_{max}^2} \int_0^\infty |H(\omega)|^2 \, d\omega$$

This equation is not elegant since it mixes quantities in the time domain and quantities in the frequency domain. Using property 9 of the transfer functions this can be written as

$$ENC_p^2 = \frac{4kT}{R_p} \frac{1}{2h_{max}^2} \int_{-\infty}^\infty [h(t)]^2 \, dt$$

The equivalent noise charge is usually expressed as a number of electron charges, rather than as a number of Coulombs.

Let us now consider the effect of a resistor R_s between the detector and the amplifier. Such a resistor is in series with the detector. The noise current source associated with this resistor is not equivalent to a noise source in parallel with the detector and therefore cannot be compared directly with a detector signal. We should calculate the noise spectrum of an imaginary current source in parallel with the detector, which will generate exactly the same noise currents as the noise of the series resistance R_s.

For this calculation, it is convenient to start from the representation of the noise of the resistor R_s as a noise voltage source in series with the resistor as shown in Fig. 8.26. Consider the circuit loop formed by the detector, the resistor R_s and the input of the amplifier. In any realistic set-up, the impedance of the detector capacitance will by far be the largest impedance in the loop. At this point we have to remember that the noise voltage of the resistor is due to a sum of a large number of elementary noise voltage signals. In the frequency interval $(\omega, \omega+d\omega)$, the elementary noise voltage signals are given by

$$V_{noise}(t) = \sqrt{2}a \sin(\omega t + \varphi)$$

Fig. 8.26 A charge-integrating amplifier with a resistor in series with the detector
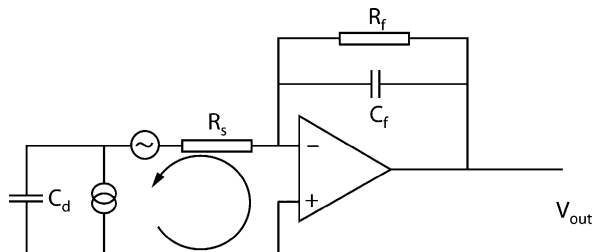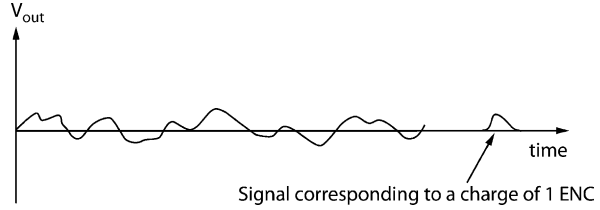
Fig. 8.27 Output signal of
the amplifier with noise. The
output pulse corresponding to
a charge pulse equal to one
ENC is also shown

In this expression 'a' is a real and positive number representing the amplitude
for a particular elementary noise wave. The phase φ has a random probability dis-
tribution with each value of the phase being equally likely. The amplitudes 'a' also
have a probability distribution, but we do not need to know this distribution. We
only need to require that this distribution is such that the Eq. (8.5) is satisfied. This
means that the probability distribution of 'a' has to satisfy the condition

$$\left\langle V_{\text{noise}}^2 \right\rangle = \left\langle a^2 \right\rangle = 4kTR_s \frac{d\omega}{2\pi}$$

Each of these elementary noise waves will cause an elementary noise current
in the detector. The noise current generated by each elementary noise signal is
given by

$$I_{\text{noise}} = \omega\, C_d\, \sqrt{2}a\, \sin\left(\omega t + \varphi + \pi/2\right)$$

In calculating the average square noise current the phase factor $\pi/2$ is unimpor-
tant, since all values of the phase are equally likely. The series resistor therefore
induces a noise current identical to the current induced by a current source in par-
allel with the detector and with an average square noise current given below. As
before we have to remember that this sine wave will be observed by some electron-
ics characterised y a transfer function $H(\omega)$ and that therefore the amplitude of the
wave is multiplied by $\sqrt{2}\omega\, |H(\omega|$

$$\langle I_{\text{noise}}^2 \rangle = C_d^2\, 4kT\, R_s \int_0^\infty \omega^2\, |H(\omega)|^2\, d\omega$$

The equivalent noise charge caused by a series resistor is therefore given by

$$(ENC_{\text{series}}\, h_{\text{max}})^2 = C_d^2\, 4kT\, R_s \int_0^\infty \omega^2\, |H(\omega)|^2\, d\omega$$

$$ENC_{\text{series}}^2 = \frac{4kT}{h_{\text{max}}^2}\, C_d^2\, R_s \int_0^\infty \omega^2\, |H(\omega)|^2\, d\omega$$

This equation is not elegant since it mixes quantities in the time domain and
quantities in the frequency domain. Using property 9 of the transfer functions this
can be written as

$$ENC^2_{series} = \frac{4kT}{2h^2_{max}} C_d^2 R_s \int_{-\infty}^{+\infty} \left(\frac{dh(t)}{dt} \right)^2 dt$$

The expressions for the *ENC* of a series resistor and a parallel resistor are usually written in a slightly different way

$$ENC^2 = \frac{4kT}{R_p} \tau a_1 + 4kT \, C_d^2 R_s \frac{a_2}{\tau}$$

$$a_1 = \frac{1}{2\tau \hat{h}^2_{max}} \int_{-\infty}^{+\infty} h^2(t) \, dt \qquad (8.7)$$

$$a_2 = \frac{\tau}{2\hat{h}^2_{max}} \int_{-\infty}^{+\infty} \left(\frac{dh}{dt} \right)^2 dt$$

The symbol τ represents the risetime of the output signal. It is easy to see that the quantities a_1 and a_2 are dimensionless numbers. Moreover, for any realistic shaping function, these coefficients are of order unity. This is illustrated in the Table 8.1 showing the value of these coefficients for a few simple shaping functions.

Equation (8.7) is the commonly used expression for the noise of an amplifier. These equations make it clear that any parallel resistor should be as large as possible. Here we see the reason why the impedance of the feedback resistor should be large. The equations also tell us that any series resistor, if present at all, should be as small as possible. The serial resistor and the parallel resistor in this expression are, in fact, the combined effect of a number of different resistors at different places in the circuit. Below we show that, for example, the first transistor of the amplifier has a noise effect as if it was a resistor in series with the detector.

It is instructive to look at a numerical example. Consider a detector with a capacitance of 30 pF and a series resistor of 1 kΩ. Assume, furthermore, a risetime of 100 ns and a feedback resistor of 100 MΩ. In this calculation, we take the coefficients $a_1 = a_2 = 1$. The equivalent noise charge caused by this feedback resistor is 25 electrons and the equivalent noise charge caused by the series resistor is 55 electrons.

Table 8.1   Value of the coefficients a_1 and a_2 for some typical shaping functions

Shaping	a_1	a_2
$\frac{1}{a} \frac{t}{a} e^{-\frac{1}{a}}$	0.93	0.92
$\frac{1}{a} \left(\frac{t}{a} \right)^4 e^{-\frac{1}{a}}$	0.45	1.0
Equilateral triangle	0.33	1.0

For the sake of definiteness let us assume that the amplifier has a bandwidth of 10 MHz. One can consider that we have 10^7 independent noise amplitude samples per second. (This is not quite exact but close enough for the sake of the argument.) If no true signal is present, the samples contain only noise and this noise has a Gaussian distribution with average value zero and an r.m.s. equal to 1 *ENC*. A sample from a Gaussian distribution has 15% chance of exceeding the average value by one standard deviation. If we use a discriminator set at a threshold corresponding to 1 *ENC*, it will be triggered almost 10^6 times per second. A signal of 1 *ENC* will be completely lost in the noise. If we set a threshold corresponding to 6 *ENC* charges, the probability that a noise fluctuation produces such an event is about 10^{-7}. Hence the noise will fake a true signal pulse about once per second. As a rule of thumb, we can say that true events should have a charge of at least about 10 *ENC* to be comfortably visible above the noise.

8.5.2 Noise Due to the First Transistor

We now turn to the calculation of the noise of the amplifier itself. The front-end part of a typical low-noise amplifier is shown in Fig. 8.28. The conducting channel in the first transistor is a resistor, and the value of this resistor is modulated by the input voltage of the amplifier. This resistor will give rise to noise. This first transistor is very often a 'field effect transistor' (FET). The reasons for this and the relative merits of FET transistors compared to bipolar transistors will become clear later. For the time being, let us consider the case of an FET transistor.

The internal structure of an FET transistor is shown in Fig. 8.29(a). The two regions of n-type silicon are called the source and the drain. There is a gate electrode separated from the surface of the p-type silicon by a thin silicon oxide insolating layer. If no voltage is present on the gate, no current will flow between the source and the drain because there is always an n–p junction preventing this current flow. If the gate is brought at a positive voltage of a few volts, the electric field will open up a conducting channel between the source and the drain and a current will flow. The intensity of the current will depend on the gate voltage. The transconductance of the FET is defined as

$$g_m(V_g) = \frac{dI_d}{dV_g}$$

In this equation, V_g and I_d are the gate voltage and the source-to-drain current, respectively. The FET behaves like a resistor in the source-to-drain channel. Notice that the source-to-drain resistor, R_{sd}, in the transistor does not behave like a normal resistor and we will use the expression 'effective noise resistance' for it. There is also a small capacitance C_t between the gate and the source. This capacitance is in parallel with the detector capacitance and therefore should be added to it. This transistor capacitance may seem a minor complication, but it will turn out that this capacitance plays an essential role in determining the noise and we have therefore included it explicitly in our calculations.

The resistor R_{sd} will cause an average square noise current in the source-to-drain channel given by

$$\left\langle I_{noise}^2 \right\rangle = \frac{4kT}{R_{sd}} \int_0^\infty |H(\omega)|^2 \, d\omega \tag{8.8}$$

To find the *ENC*, we need to calculate what hypothetical noise source in parallel with the detector will cause a current in the source-to-drain channel identical to Eq. (8.8). The noise will be due to an apparent noise voltage at the gate given by

$$\langle V_g^2 \rangle = \frac{4kT}{g_m^2 R_{sd}} \int_0^{+\infty} |H(\omega)|^2 \; d\omega$$

Following the same argument as used when calculating the noise due to a series resistor, we conclude that this gate voltage corresponds to an apparent noise current source in parallel with the detector and with a noise current given by

$$\left\langle I_{noise}^2 \right\rangle = (C_d + C_t)^2 \frac{4kT}{g_m^2 R_{sd}} \int_0^\infty \omega^2 \, |H(\omega)|^2 \, d\omega \tag{8.9}$$

The noise caused by the resistance of the source-to-drain channel is therefore equivalent to an apparent noise current source in parallel with the detector with a noise spectrum given by Eq. (8.9). The corresponding equivalent noise charge is therefore

$$ENC^2 = \frac{4kT}{R_{sd}} \frac{(C_d + C_t)^2}{g_m^2} \frac{1}{h_{max}^2} \int_0^\infty \omega^2 |H(\omega)|^2 d\omega$$

$$ENC^2 = \frac{4kT}{R_{sd}} \frac{(C_d + C_t)^2}{g_m^2} \frac{a_2}{\tau}$$

We see that this noise depends on the temperature and the shaping time exactly in the same way as if it were a resistor in series.

The structure of an FET is shown in Fig. 8.28. It is a narrow strip that can be made arbitrarily long. Obviously, the transconductance of the FET is proportional to the length of this strip. It therefore seems that we can make this noise as small as we wish by making the strip sufficiently long and therefore the transconductance sufficiently large. It is indeed possible to make the transconductance very large, but in doing so we will also make the capacitance C_t very large. The FET capacitance C_t is proportional to the transconductance as will be shown below.
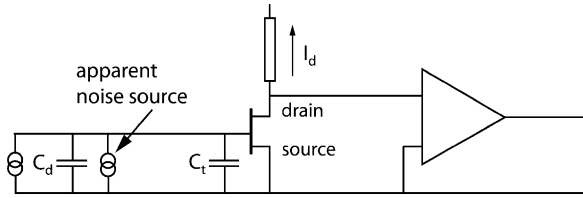
Fig. 8.28 First transistor in a detector for nuclear electronics. The capacitance C_t in this figure represents the capacitance of the transistor itself. The noise of the source-to-drain channel is equivalent to noise source in parallel with the detector

Consider the volume indicated by the box in Fig. 8.29(b).
Applying Maxwell's equation to this volume we can write

$$\int Dds = \int \rho dv = 0$$

The surface integral over the electric field is zero, therefore the total charge inside this volume must be zero. We can only change the number of charges in the conduction channel of the FET by changing the number of charges in the gate and we have

$$dQ_{channel} = dQ_{gate} = C_t dV_g$$

But the source-to-drain current is related to the number of charges in the conduction channel and the transit time of these charges by
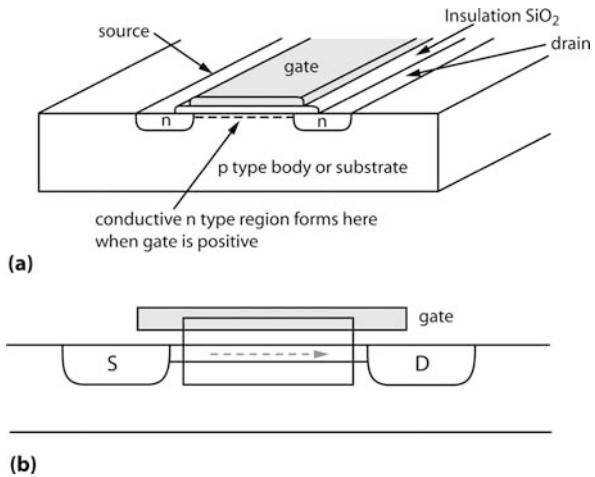
$$I_d = \frac{Q_{channel}}{t_{transit}}$$

Fig. 8.29 (a) Physical layout of an FET transistor, (b) Highly schematic representation of the same

Therefore, any change in the source-to-drain current is caused by a change in the number of charges in the channel and we have

$$dI_d = \frac{dQ_{\text{gate}}}{t_{\text{transit}}} = \frac{C_t\, dV_g}{t_{\text{transit}}}$$

We therefore have the following relation between the transconductance g_m, the capacitance C_t and the transit time of the charges through the FET channel t_{transit}:

$$\frac{dI_d}{dV_g} = g_m = \frac{C_t}{t_{\text{transit}}}$$

Furthermore, it is possible to show that (see Sect. 7.2.4 in [2] in Chap. 5)

$$R_{\text{ds}} = \frac{3}{2} \cdot \frac{1}{g_m}$$

We can now express g_m and R_{sd} as a function of the FET capacitance C_f and insert this into the expression for the *ENC* above

$$ENC^2 = \frac{8}{3} kT \frac{(C_d + C_t)^2}{C_t} t_{\text{transit}} \frac{a_2}{\tau}$$

This noise is minimised if we choose the transconductance of the FET such that $C_t = C_d$. That can be done by taking an FET with an FET strip of the correct length. For such an optimised FET the noise can hence be written as

$$ENC^2 = \frac{8}{3}\, 4kTC_d\, a_2\, \frac{t_{\text{transit}}}{\tau} \tag{8.10}$$

The above equation represents the noise due to the front-end FET transistor in the amplifier. This equation only holds if this transistor is matched to the capacitance of the detector. In this equation τ represents the rise time of the pulse and t_{transit} the transit time of the charges through the FET channel. This transit time is of the order of 0.25 ns in typical modern FET transistors. We see from the above equation that for an optimised amplifier design the noise is proportional to the square root of the detector capacitance. While the noise caused by real resistances in the system can, in principle, be made arbitrarily small, the noise caused by the FET itself is unavoidable. Equation (8.10), therefore, represents a true lowest possible value for the noise.

A numerical example is instructive: assume a detector with a capacitance of 10 pF, a rise time of the signal of 250 ns and a FET with a transit time of 0.25 ns. From the equation above we have $ENC_{\text{FET}} = 126$ electrons. This is a true lower limit on the noise that can be reached. Carefully designed amplifiers can reach a noise level that is close to this theoretical lower limit, but more often the noise will be several times larger. This means that a signal should be at least several thousand electrons in order to be visible.

The noise expression above only holds if for each value of the detector capacitance one uses the input FET that is matched to this capacitance. For a given amplifier, the noise will be independent of the detector capacitance if it is lower than the FET capacitance and increase proportionally with the detector capacitance if it is larger than the transistor capacitance.

8.6 Shot Noise

An electrical current is a flow of discrete electric charges and not a smooth flow. Therefore any current is unavoidably affected by fluctuations associated with the random arrival of these charges. This noise is called shot noise.

For the sake of definiteness let us consider a simple vacuum photodiode. It consists of an evacuated glass tube with a photocathode on one side of the glass and an electrode facing the photocathode on the other side. A potential difference between the photocathode and the electrode ensures collection of the photoelectrons by the electrode (see Fig. 8.30(a)). If a light source illuminates the photocathode, we will measure a photocurrent. Because of the nature of this set-up, it is obvious that these electrons will be emitted randomly with a Poisson frequency distribution.

We furthermore assume that the signal produced by each electron is a square pulse of duration Δt and amplitude $i = Q/\Delta t$, where Q is the charge of the electron. If a weak light source illuminates the photocathode the photocurrent signal will look like the signal in Fig. 8.30(b). If at any instant in time we measure the current, we will observe a current equal to $n_i.(Q/\Delta t)$, where n_i is an integer number. In other words, we always see an integer number of electrons. This integer will be equal to all the electrons that arrived within a given time window of duration Δt. This n_i is a
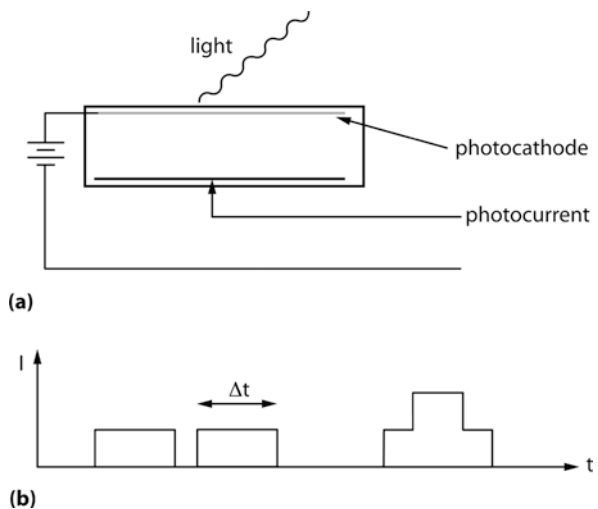

Fig. 8.30 A vacuum photodiode (**a**) and its photocurrent (**b**)

random variable with average value <n_i>. At any instant in time we have

$$I = n_i \frac{Q}{\Delta t}$$

The time averaged values of I and n are therefore related by

$$\langle I \rangle = \langle n_i \rangle \frac{Q}{\Delta t}$$

And the r.m.s. dispersions on the current I and the number of electrons n_i are related by

$$\left\langle (I - \langle I \rangle)^2 \right\rangle = \left\langle (n_i - \langle n_i \rangle)^2 \right\rangle \left(\frac{Q}{\Delta t} \right)^2$$

The random variable n_i has a Poisson distribution, hence the r.m.s. dispersion of this variable equals to the square root of its average value

$$\left\langle (I - \langle I \rangle)^2 \right\rangle = \langle n_i \rangle \left(\frac{Q}{\Delta t} \right)^2 = \frac{\langle I \rangle \, \Delta t}{Q} \left(\frac{Q}{\Delta t} \right)^2 = \frac{\langle I \rangle \, Q}{\Delta t}$$

This can be seen as a noise current I_{noise} with average value zero, superimposed on steady current <I>. The r.m.s. noise current due to shot noise is therefore given by

$$\left\langle I_{noise}^2 \right\rangle = \frac{\langle I \rangle \, Q}{\Delta t}$$

We have derived an expression for the noise current in the special case where the pulse corresponding to one electron is a square pulse. To generalise this result to an arbitrary impulse response $h(t)$ we will use a theorem about random sums of random variables. It can be found in a good textbook on statistics. See also Exercise 3.

Theorem: *Be* x_i *a number of independent random variables, all with the same probability distribution. Consider the random variable* $S = \sum_{i}^{1...n} x_i$, *where the integer n is itself a random variable with a Poisson distribution with average value* λ. *The following relations hold:*

$$\langle S \rangle = \lambda \, \langle x_i \rangle$$
$$\sigma^2 \{S\} = \left\langle (S - \langle S \rangle)^2 \right\rangle = \lambda \left\langle x_i^2 \right\rangle$$

Consider a time interval T long compared to the shaping time of the pulse. Assume that there is one, and only one, charge in this time interval as is illustrated in Fig. 8.31. The impulse response corresponding to the arrival of this charge is $Q \, \hat{h}(t)$ with $\int \hat{h}(t) = 1$.

Fig. 8.31 Response of the
system to one electron
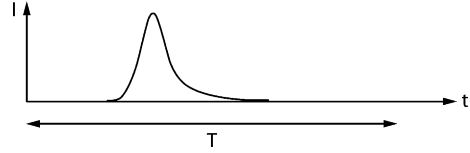
We have for the average current $<I>$ and for the average square current $<I^2>$ the following relations:

$$\langle I \rangle = \frac{Q}{T} \int h(t)dt = \frac{Q}{T}$$

$$\langle I^2 \rangle = \frac{Q^2}{T} \int h^2(t)dt$$

If there are on average λ charges in the time interval T and if the number of charges has a Poisson distribution, we thus have according to the above theorem:

$$\langle I \rangle = \lambda \frac{Q}{T}$$

$$\langle (I - \langle I \rangle)^2 \rangle = \lambda \frac{Q^2}{T} \int h^2(t)dt$$

$$= \langle I \rangle Q \int h^2(t)\, dt$$

Therefore the apparent noise current is given by

$$\left\langle I^2_{\text{noise}} \right\rangle = \left\langle (I - \langle I \rangle)^2 \right\rangle = \langle I \rangle Q \int h^2(t)\, dt$$

As before, this can be written as an integral over the transfer function:

$$\left\langle I^2_{\text{noise}} \right\rangle = 2 \langle I \rangle Q \int_0^\infty |H(\omega)|^2\, d\omega$$

Shot noise is also white noise! One verifies that the above expression reduces to the result derived previously in the particular case of a square shaping function.

Following the same arguments as before, we conclude that the *ENC* for shot noise is given by

$$ENC^2 = 2IQ \frac{1}{h_{\text{max}}} \int_{-\infty}^{+\infty} h^2(t)dt \tag{8.11}$$

$$ENC^2 = 2IQ\tau a_1$$

where a_1 is given by Eq. (8.7).

Equation (8.11) does not apply to the resistive current in a normal resistor. The derivation assumes that the shaping time is large compared to the physical formation time of the pulse and for a normal resistor this assumption is usually not satisfied. Indeed, the formation time of the pulse due to one electron is the time it takes for the electron to travel from one electrode to the other and this time is very long in a normal resistor. The speed of electrons in a resistor is of the order of metres per hour! In addition, the assumption that the number of individual electrons has a Poisson distribution is not valid.

Any detector itself contributes to the noise in two ways: by the shot noise associated with the current through the detector and by the thermal noise of its resistance. The shot noise will increase with the current through the detector; therefore, with increasing voltage over the detector at some point, the shot noise will exceed the thermal noise. This happens when the shot noise associated with the current through a detector (Eq. 8.11) exceeds the thermal noise associated with the resistance of the detector (1st term in Eq. 8.7), therefore when

$$2IQ\,\tau\,a_1 \;>\; \frac{4kT}{R}\,\tau a_1$$
$$IR \;>\; \frac{2kT}{Q} \approx 50\,\text{mV}$$

The quantity I.R is the voltage over the detector. Therefore the shot noise will be larger than the thermal noise if the voltage over the detector is larger than 50 mV.

The voltage over nuclear detectors is always much larger than 50 mV. Therefore, the dominant noise contribution is the shot noise caused by the dark current, rather than the thermal noise due to the resistance of the detector.

In the derivation above we have assumed that the charge quanta were equal to one electron charge. Often the detector has an internal multiplication mechanism and the charges arrive in multiples of the electron charge. In that case the charge Q to be used in the expression for the shot noise is not the electron charge but the electron charge multiplied by the multiplication factor. If all primary events have the same multiplication factor, all one needs to do is to use the correct charge Q in the above formula. However, often the multiplication factor varies from one event to the next. This gives rise to an additional noise called the excess noise, and is represented by a factor that is usually denoted by F. The concept of excess noise factor was introduced in Sect. 6.4. In this case the expression for the *ENC* becomes:

$$ENC^2 = 2IQ\tau F a_1$$

Shot noise will usually be regarded as something undesirable. However, the shot noise can be used to measure physical quantities and this technique is referred to in the literature as the Campbell measuring mode, after the person who developed it. This technique can for example be used when measuring the neutron flux with a proportional tube in the presence of a strong gamma ray background. In this case, the detector signal consists of small pulses of amplitude 'q' caused by gamma rays

and of large pulses of amplitude 'Q' caused by neutrons with $Q \gg q$. If the pulse rate is not too high one can reject the gamma ray pulses using a discriminator at a sufficiently high threshold such as to reject all the gamma ray-induced pulses. If the event rate is very high this is no longer possible. This situation is often encountered when measuring neutron fluxes at nuclear reactors.

The total current I_{total} of the detector is the sum of the current induced by the neutrons I_n and the current induced by the gamma rays I_γ. Therefore

$$I_n = I_{\text{total}} - I_\gamma$$

The measured total current has to be corrected for a poorly known gamma ray-induced current. This gamma ray-induced current is sometimes much larger than the neutron-induced current, making the measurement essentially impossible.

If we measure the current noise rather than the current itself we are much less sensitive to the gamma ray-induced current. Indeed, this current noise is given by

$$\left\langle I_{\text{noise}}^2 \right\rangle = (I_\gamma q + I_n Q) 2 \int\limits_0^{+\infty} |\hat{H}(\omega)|^2 \, d\omega$$

From this we can derive the following expression for the neutron-induced current:

$$I_n = \frac{\langle I_{\text{noise}}^2 \rangle}{2Q \int\limits_0^{+\infty} \left|\hat{H}(\omega)\right|^2 d\omega)} - \left(I_\gamma \frac{q}{Q} \right)$$

We still need to correct for the gamma-induced current I_γ, but this correction is suppressed by a large factor q/Q and we are much less sensitive to any uncertainty in this gamma ray-induced current.

8.7 Summary and Conclusions

In a carefully optimised detector system the two largest noise sources are the shot noise due to the dark current of the detector and the thermal noise associated with the first transistor in the amplifier. Considering only those two terms the *ENC* can be written as

$$ENC^2 \geq a_2 \frac{8}{3} 4kTC_d \frac{t_{\text{transit}}}{\tau} + a_1 \, 2eI_d\tau$$

There is always some additional noise due to other imperfections in the amplifier or other parts of the electronics, therefore this equation is written as an inequality.

The shot noise is due to the dark current of the detector itself. It increases proportionally to the square root of the shaping time. The noise associated with the first
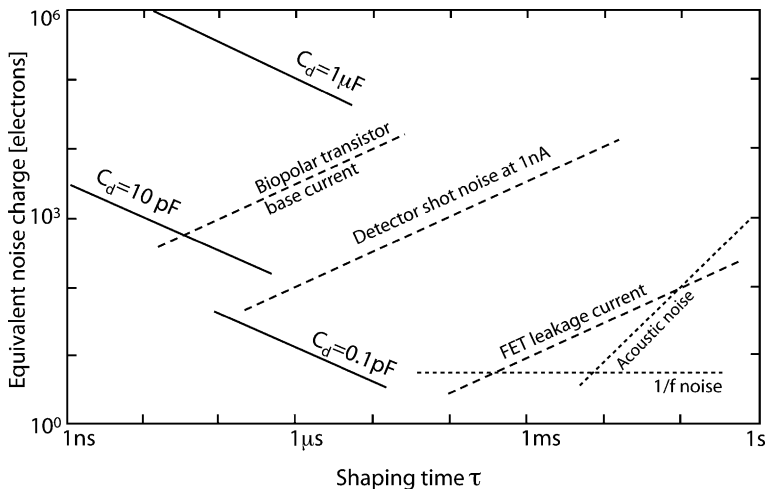
Fig. 8.32 Typical values for the different contributions to the noise in an amplifier for particle detection. The solid lines represent the noise caused by the first transistor for different values of the capacitance of the detector. A number of other contributions to the noise are also shown. Figure from [1]

transistor decreases inversely proportional to the shaping time and increases proportionally to the square root of the detector capacitance. At a high rate one needs a short shaping time and this last term dominates the noise.

Figure 8.32 gives an overview of the different noise contributions in a detector as a function of the shaping time. The solid lines in this figure give the noise contribution from the first transistor in the amplifier for different values of the detector capacitance. The detector capacitance is, of course, related to the size of the detector. For detectors in nuclear and particle physics the capacitance ranges from well below 1 pF in some pixel detectors to several μF for calorimeters used in high-energy physics experiments.

Figure 8.32 also shows the shot noise of a hypothetical detector with a dark current of 1 nA. The shot noise associated with the FET leakage current is also shown. This leakage current has the same shot noise effect as the dark current in the detector. We also show the noise caused by the base current of the bipolar transistor, in case one uses a bipolar transistor instead of a FET. The plot makes it clear why it is usually preferable to use a FET. The shot noise associated with an FET is orders of magnitude smaller than the shot noise of a bipolar transistor. However, if the detector capacitance is very large, or the shaping time very short, other noise contributions become dominant and it is preferable to use a bipolar transistor because this is simpler and has a number of other advantages.

If the shaping time is of the order of 1 ms or longer, other noise sources become important. One important type of noise is the $1/f$ or flicker noise. The term '$1/f$' stands for 'one over the frequency' and this type of noise can have many causes. In an amplifier both capacitances and the resistors contribute to the $1/f$ noise.

A capacitance without dielectric is noiseless, but practical capacitances have a dielectric medium between the plates. The polarisation in this dielectric is not a purely smooth process and the effect is the induction of some $1/f$ noise. Resistors also suffer from fluctuations in resistance, generating an additional noise proportional to the current flowing through them. The magnitude of this noise depends on details of the construction of the resistors, but it always has an $1/f$ frequency spectrum. Typical values are [μV per decade of frequency and per volt over the resistance]

carbon composition	0.01 – 3.0
carbon film	0.05 – 0.3
metal film	0.02 – 0.2
wire wound	0.01 – 0.2

In nuclear and particle physics one is rarely interested in a shaping time that is longer than 1 ms, hence these other noise sources are of little concern.

8.8 Exercises

(1) Derive property 6 of the transfer function in Sect. 8.2.
(2) Calculate the Fourier transform of Eq. (8.4).
(3) Prove the following theorem.
    Be x_i a number of independent random variables, all with the same probability distribution. Consider the random variable

$$R = \sum_i^{1...n} x_i,$$

where the integer n is itself a random variable with a Poisson distribution with average value l. The following relations hold:

$$\langle R \rangle = \lambda \langle x_i \rangle$$
$$\sigma^2\{R\} = \left\langle (R - \langle R \rangle)^2 \right\rangle = \lambda \left\langle x_i^2 \right\rangle$$

(4) Assume that you are measuring the noise voltage of a resistor using a digital oscilloscope with an input impedance of 10 MΩ and a bandwidth of 400 MHz. For what value of the resistor will you measure the largest value for the noise. How much will this maximum noise be in mV?
(5) Consider a silicon strip detector where each strip has a capacitance of 20 pF and a dark current of 20 nA. The rise time of the pulse is 30 ns. Give an under limit for the noise. Take the shape coefficients $a_1 = a_2 = 1$.
(6) Prove that the stationary solutions for a transmission line of length D are given by the following equations

$$
\begin{cases}
I_n(x, t) = I_n \sin\left(\dfrac{n\pi x}{D}\right) \sin\left(\dfrac{n\pi v_0 t}{D} + \varphi_n\right) \\
V_n(x, t) = I_n Z_0 \cos\left(\dfrac{n\pi x}{D}\right) \cos\left(\dfrac{n\pi v_0 t}{D} + \varphi_n\right)
\end{cases} \quad n = 1, .. \infty
$$

References

1. V. Radeka, Low noise techniques in detectors, Ann. Rev. Nucl. Sci. 38, 217–277 (1988).
2. F.S. Goulding and D.A. Landis, IEEE Trans. Nucl. Sci. NS-29, 1125 (1982).

Solutions to Exercises

Chapter 1

1. Show that the Lorentz transformation is such that the velocity of a light ray travelling in the x direction is the same for the observer in the frame S and for the observer in the frame S'.

   Solution: Consider a light ray travelling in the x direction. If the light ray connects two space–time points $\{t_1, x_1\}$ and $\{t_2, x_2\}$, we have

   $$c = \frac{x_2 - x_1}{t_2 - t_1}$$

   The speed of light observed in the frame S' will be

   $$c' = \frac{x'_2 - x'_1}{t'_2 - t'_1} = c\frac{\gamma((x_2 - x_1) - \beta c(t_2 - t_1))}{\gamma(c(t_2 - t_1) - \beta(x_2 - x_1))}$$

   $$= c\frac{\left[\dfrac{x_2 - x_1}{t_2 - t_2} - \beta c\right]}{\left[c - \beta\dfrac{x_2 - x_1}{t_2 - t_2}\right]} = c$$

2. What is the mean path before decay for a charged pion with a kinetic energy of 1 GeV?

   Solution: The pion has a lifetime 2.6×10^{-8} s and a mass of 139.6 MeV. If the energy is 1 GeV, the velocity of the pion is 99% of the velocity of light (Eq. 1.4). The mean path before decay is

   $$= 0.99\, c\, \gamma\, \tau$$
   $$= 0.99\, c\frac{1000 + 139.6}{139.6}\, 2.6\, 10^{-8} = 63\,\text{m}$$

3. Show that the relativistic expression for the kinetic energy of a particle (Eq. 1.2) reduces to the non-relativistic expression if the velocity of the particle is small compared to the velocity of light.

Solution:

$$E = E_{\text{kinetic}} + m_0 c^2 = \frac{m_0 c^2}{\sqrt{1 - (v/c)^2}}$$

$$\approx \frac{m_0 c^2}{(1 - 1/2(v/c)^2)} \approx m_0 c^2 (1 + 1/2(v/c)^2)$$

$$= m_0 c^2 + \frac{1}{2} m_0 v^2$$

4. For a Poisson distribution with average value 16, calculate the probability to observe 12, 16 and 20 as measured value. Calculate the probability density function for a Gaussian distribution with average value 16 and dispersion 4, for the values $x = 12$, 16 and 20. Compare the results.

Solution: For a Poisson distribution $P(12) = 0.0829$, $P(16) = 0.1024$, $P(20) = 0.0418$

For a Gaussian distribution, $f(12) = 0.0605$, $f(16) = 0.0997$, $f(20) = 0.0605$

5. Consider a very short-lived particle of mass M decaying into two long-lived particles 1 and 2. Assume you can measure accurately the energies and momenta of the two long-lived particles. How will you calculate the mass of the short-lived particle from the known energies and momenta of the two long-lived objects?

Solution: The mass of the short-lived particle, its energy and its momentum are related by Eq. (1.1). The energy and momentum of the particle are equal to the sums of the energy and sums of the momenta of the decay products, therefore

$$M^2 c^4 = (E_1 + E_2)^2 - c^2 (\vec{P}_1 + \vec{P}_2)^2$$

6. Calculate the order of magnitude of the energy levels in atoms and in nuclei using the 'particle in a box' approximation, Eq. (1.9). Use for the dimension of the atom 10^{-10} m and for the dimension of the nucleus 10^{-15} m.

Solution: Atomic energy levels: ≈ 40 eV; nuclear energy levels: ≈ 400 MeV.

7. Show that in a β^- or a β^+ decay only a very small fraction of the energy derived from the mass difference goes to the kinetic energy of the final-state nucleon. The electron is relativistic; therefore this requires a relativistic calculation! Hint: the 3-body problem can be reduced to a 2-body problem by considering the electron–neutrino system as one object with a mass of a few MeV.

Solution. Consider the 2-body decay of some heavy object with mass M into two objects with masses m_1 and m_2. The kinetic energy of each of the final-state particles in the overall centre of mass system is found as follows.

Consider two particles with energy and momentum four vectors p_1 and p_2. The symbol p_i stands for the four-vector $\{E_i, c\vec{p}_i\}$. The energy E appearing in this expression is the total energy E, i.e. the rest energy mc^2 plus the kinetic energy. The four-vector product $(p_1.p_2)$ is defined as

$$(p_1.p_2) = \left[E_1 E_2 - c^2 \vec{p}_1 \vec{p}_2 \right]$$

A four-vector product is a Lorentz invariant; this quantity can be evaluated in any reference frame, and the result is the same. Consider now the quantity

$$\frac{(p_1.p_2)}{m_1 c^2}$$

This is a Lorentz invariant. Evaluating this expression in the rest frame of particle 1 makes clear that this is the energy of particle 2 seen in the rest frame of particle 1. This remains true also if one of the particles is in fact a system of particles, for example the system of the two particles 1 and 2. The energy of particle 2, seen in the overall centre of mass frame of the particles 1 and 2 is therefore

$$E_2^* = \frac{(p_1 + p_2).p_2}{(p_1 + p_2)^2}$$

We have the following relations:

$$(p_1 + p_2)^2 = M^2 c^4$$
$$(p_1.p_2) = \frac{1}{2} \left[(p_1 + p_2)^2 - (p_1)^2 - (p_2)^2 \right]$$
$$= M^2 c^4 - m_1^2 c^4 - m_2^2 c^4$$

And therefore finally

$$E_2^* = \frac{M^2 c^4 + m_2^2 c^4 - m_1^2 c^4}{2Mc^2}$$

Let us now apply the above relation to the decay

$$N^* \rightarrow N + e^- + \bar{\nu}_e + Q$$

The symbol Q represents the energy liberated in the reaction. Let us denote by M^* the mass of the parent nucleus, by M the mass of the final-state nucleus and by m the mass of the electron–neutrino system. The kinetic energy of the nucleus in the final state is given by

$$E_{kin} = \frac{M^{*2}c^4 + M^2c^4 - m^2c^4}{2M^*c^2} - Mc^2$$

$$= \frac{M^{*2}c4 + M^2c^4 - m^2c^4 - 2M^*c^2Mc^2}{2M^*c^2}$$

$$= \frac{(M^* - M)^2 c^4 - m^2c^4}{2M^*c^2}$$

$$= \frac{(mc^2 + Q)^2 - m^2c^4}{2M^*c^2}$$

$$= Q\frac{[2mc^2 + Q]}{2M^*c^2}$$

The value of the quantities Q and mc^2 is a few MeV, while the value of M^*c^2 is a few GeV. Therefore the kinetic energy of the final-state nucleon is only of the order of 0.1% of the energy liberated in the reaction Q. The rest of the kinetic energy is given to the electron and the neutrino.

Chapter 2

1. Calculate the approximate mean free path of a high-energy neutron in dry air. Air is 80% ^{14}N and 20% ^{16}O by volume.

   Solution: Obtain the density for N_2 ($A = 14$) and O_2 ($A = 16$) from the gas constant: one mole = 22.4 l at standard temperature and pressure. One finds for the partial densities of nitrogen and oxygen gas: $\rho_N = 1.0 \times 10^{-3}$ g/cm^3; $\rho_O = 0.286 \times 10^{-3}$ g/cm^3.

   Use the expression for the hadronic interaction length in Sect. 2.5, to find the mean free path: $\lambda_N = 843$ m; $\lambda_O = 3083$ m. Use Eq. (2.2) to find the total mean free path in air. The result is 662 m.

2. Derive Eq. (2.10) using energy momentum conservation.

   Solution: A photon has energy $E = \hbar\omega$. Because of Eq. (1.1), it has a momentum $P = \hbar\omega/c$. If the outgoing electron has energy E_{out} and momentum P_{out} and we use φ for the angle between the outgoing electron and the direction of the incoming photon, energy and momentum conservation is written as

$$\begin{cases} mc^2 + \hbar\omega = E_{out} + \hbar\omega' \\ \dfrac{\hbar\omega}{c} = P_{out}\cos\varphi + \dfrac{\hbar\omega'}{c}\cos\theta \\ 0 = P_{out}\sin\varphi + \dfrac{\hbar\omega'}{c}\sin\theta \end{cases}$$

Furthermore Eq. (1.1) gives a relation between energy E_{out} and momentum P_{out}. We hence have four equations and need to eliminate three variables, E_{out},

P_{out} and φ. This is easily done if we rewrite slightly the energy and momentum conservation equations as

$$\begin{cases} (mc^2 + \hbar\omega - \hbar\omega')^2 = E_{\text{out}}^2 \\ (\hbar\omega - \hbar\omega' \cos\theta)^2 = c^2 P_{\text{out}}^2 \cos^2\varphi \\ (\hbar\omega' \cos\theta)^2 = c^2 P_{\text{out}}^2 \sin^2\varphi \end{cases}$$

We thus find

$$m^2 c^4 = (mc^2 + \hbar\omega - \hbar\omega')^2 - (\hbar\omega - \hbar\omega' \cos\theta)^2 - (\hbar\omega' \sin\theta)^2$$
$$m^2 c^4 = m^2 c^4 + 2mc^2(\hbar\omega - \hbar\omega') + 2\hbar\omega\hbar\omega'(\cos\theta - 1)$$

And this gives directly the desired result

$$\hbar\omega' = \frac{\hbar\omega}{\left(1 + \dfrac{\hbar\omega}{m_e c^2}(1 - \cos\theta)\right)}$$

3. Consider two protons with the same kinetic energy and travelling on a head-on collision trajectory. The protons repel each other by the Coulomb force. At the point of closest approach the distance between the protons is 2×10^{-10} m. What is the energy of each of these protons?

Solution: The force between the two protons is given by

$$F = \frac{e^2}{16\pi r^2 \varepsilon_0}.$$

In this equation, r represents the distance of one proton to the centre of mass point, i.e. half of the distance between the two protons. The energy lost by a proton coming from infinity and arriving at a distance R from the centre of mass point is given by

$$-\int_{\infty}^{R} eF \, dr = \frac{e^2}{16\pi \varepsilon_0 R}$$

If this energy loss is equal to the kinetic energy of the proton, this equation gives the distance of closest approach of the particle, and we have

$$\frac{e^2}{16\pi \varepsilon_0 R_{\text{closed}}} = E_{\text{kinetic}}.$$

The numerical values for the constants in this equation can be found in annex 1. If we take the distance of closed approach $= 10^{-15}$ m, we find $E_{\text{kinetic}} = 360$ MeV.

4. Derive the expression for the threshold energy for the Cherenkov effect (Eq. 2.8), starting from $E = \gamma m_0 c^2$.

Solution:

$$E = E_{kin} + m_0 c^2 = \frac{m_0 c^2}{\sqrt{1 - (v/c)^2}}$$

At the Cherenkov threshold $v/c = 1/n$, therefore

$$E_{kin} = \frac{m_0 c^2}{\sqrt{1 - (1/n)^2}} - m_0 c^2$$

$$= m_0 c^2 \left(\sqrt{\frac{n^2}{n^2 - 1}} - 1 \right)$$

5. The diameter of atoms is of the order of the Bohr radius and is given by

$$a = \frac{4\pi \varepsilon_0 \hbar}{m_e e^2}.$$

Use the Heisenberg uncertainty relation to argue that the velocity of the electrons in atoms is of the order $v \approx c\alpha$, where α is the fine structure constant.

$$\alpha = \frac{e^2}{4\pi \, \hbar c \, \varepsilon_0} \approx \frac{1}{137}$$

Solution:

$$\Delta p_x \Delta x \approx \hbar$$
$$v = \frac{\Delta p_x}{m_e} = \frac{\hbar}{m_e a} = c\alpha$$

Chapter 3

1. Assume a linear accelerator as shown in Fig. 3.5 and an alternating voltage source of 10 MHz. Assume we want to use it to accelerate electrons. After a few steps, the electrons will have a velocity close to the velocity of light. How long should each of the tubes be to accelerate each electron further?

Solution:

$$\frac{c}{2f} = 15m$$

2. Assume you have a cyclotron with a magnet of 1.5 Tesla field. The useful diameter of the magnet is 2 m. What is the maximum energy you can reach for protons with this machine?

Solution: The mass of a proton is $mc^2 = 938$ MeV. From Eq. (3.1), the maximum momentum the proton can have is $Pc = 448.5$ MeV. The maximum energy is given by Eq. (1.5): $E = 102$ MeV.

3. Show that in a cyclotron the distance between the successive orbits becomes smaller as the energy of the particles becomes larger.

Solution: With each revolution the kinetic energy increases by the same amount. Differentiating Eq. (1.1), we get the relation between the step in energy and the step in momentum

$$c\Delta P\, cP = \Delta E_{\text{kinetic}}(E_{\text{kinetic}} + mc^2)$$

$$c\Delta P = \Delta E_{\text{kinetic}}\frac{(E_{\text{kinetic}} + mc^2)}{cP}$$

$$c\Delta P = \Delta E_{\text{kinetic}}\frac{\sqrt{m^2c^4 + P^2c^2}}{cP}$$

$$c\Delta P = \Delta E_{\text{kinetic}}\sqrt{1 + \frac{m^2c^4}{(cP)^2}}$$

$$\Delta r = \frac{c\Delta P}{cZeB} = \frac{\Delta E_{\text{kinetic}}}{c\,Ze\,B}\sqrt{1 + \frac{m^2c^4}{(c\,Ze\,Br)^2}}$$

This equation clearly shows that identical steps in energy will correspond to smaller steps in radius as the energy or momentum of the particle increases.

4. Show that the equation for the radius of curvature of the track of a charged particle in a magnetic field: $P = Ze\,B\,r$, can be rewritten as Eq. (3.1).

Solution: $cP = Ze\,B\,r\,c$. If e is expressed in coulomb, B in tesla and r in meter, we obtain cP in joule. To obtain the quantity Pc in eV, divide the expression by the charge of one electron

$$\{cP\}[eV] = Z\,B[\text{tesla}]\,r[m]\,c[\text{m/s}] = 2.9979 \times 10^8 Z\,B[\text{tesla}]\,r[m].$$

5. Assume that to drive a nuclear reactor one needs a beam of protons with an energy of 1 GeV and a beam current of 20 mA. Assume that the accelerator has an efficiency for converting electrical energy to beam energy of 33%. How much electrical power will this accelerator use?

Solution. If 'e' is the charge of a proton, 'I' the beam current and 'E' the beam energy, the number of protons per second in the beam is given by $I[A]/e[\text{coulomb}]$.

The energy of one proton of $E[eV]$ expressed in joule is $= e[\text{coulomb}]\, E[eV]$

The power in the beam in watt is hence $I[A].E[eV]$ or 20 MW.

With an efficiency of 33% the electrical power consumption of the accelerator is 60 MW.

6. What is the speed of a train that has the same kinetic energy as the energy stored in one of the proton beams of the LHC accelerator. A typical train weighs 400 metric tons.

Solution: There are 3×10^{14} protons in one of the beams of LHC. The total energy E_{tot} in the beam therefore is 33.6×10^7 joule.

The speed of the train with kinetic energy E_{tot} is given by

$$v[m/s] = \sqrt{\frac{2E_{tot}}{m[kg]}}$$

This gives 40 m/s or $\approx$150 km/h for the speed of the train.

7. Assume we accelerate protons and make them collide with protons at rest. What should be the energy of the proton beam to produce the same centre of mass energy as is achieved in collisions at the LHC collider.

Solution: The beam energy of the proton producing the same centre of mass energy in a fixed target collision is given by

$$E_{beam} = \frac{E_{CM}^2}{2\,m_p c^2}$$

For a centre of mass energy of 14 TeV this gives 2.6×10^{16} eV

8. In the SPS proton synchrotron, the frequency of the RF cavities at the maximum energy of 450 GeV is 200.2 MHz. How much should the frequency be at the injection energy of 10 GeV?

Solution: The nominal trajectories of the particles should be the same at both energies. We therefore have

$$f = \frac{v}{L} = \frac{c\sqrt{1 - (mc^2/E)^2}}{L}$$

where v and E are the velocity of the protons and L the length of the trajectory. The frequencies at 10 and 450 GeV are hence related by.

$$f_{10} = f_{450}\frac{v_{10}}{v_{450}} = f_{450}\frac{\sqrt{1 - \left(mc^2/10\right)^2}}{\sqrt{1 - \left(mc^2/450\right)^2}}$$

The result is $f_{10} = 199.3$ MHz.

9. Assume a synchrotron for electrons with a beam energy of 1 GeV. What is the power dissipated by synchrotron radiation? Assume that the bending magnets

have a field of 2 tesla, that the number of particles stored is 10^{12}, and that 33% of the circumference is occupied by the bending magnets. The rest of the circumference has quadrupoles and straight sections. Neglect the power dissipated in the quadrupoles.

Solution: Use Eq. (4.2) to find the radius of the beam trajectory in the bending magnets ($r = 1.66$ m). The circumference of the cyclotron is therefore $\approx$30 m and the rotation frequency $\approx 10^7$. The total power dissipated is the power per turn (Eq. 3.3) times the rotation frequency times the number of electrons in the ring. The result is 8×10^4 Watt.

Chapter 4

1. Consider an MWPC with wire spacing Δ. Assume that for perpendicular tracks the biggest signal is always on the nearest wire. Show that the r.m.s. position resolution obtainable with such a detector is given by $\sigma = \dfrac{\Delta}{\sqrt{12}}$

   Solution: The r.m.s. measurement error σ is given by

   $$\sigma^2 = \int (x_{\text{measured}} - x)^2 f(x)\, dx$$

   In this expression x represents the true position of the track, $f(x)$ the probability density function for the true position x and x_{measured} represents the measured value of the position. In this case, the measured position is the position of the wire nearest to the true position. Let us first consider the particular case where the true position is within a distance $\Delta/2$ of a particular wire at position P. In this case the integral becomes

   $$\sigma^2 = \int (P - x)^2 f(x)\, dx$$
   $$\begin{cases} f(x) = \dfrac{1}{\Delta} & P - \Delta/2 < x < P + \Delta/2 \\ f(x) = 0 & \text{all other values of x} \end{cases}$$

   The r.m.s. position error is hence given by

   $$\sigma^2 = \int_{P-\Delta/2}^{P+\Delta/2} (P - x)^2\, f(x)\, dx = \frac{\Delta^2}{12}$$

   Since the true position is always within a distance $\Delta/2$ of one of the wires this result holds for any value of the true position of the track.

2. You suspect that the gas in a cave is heavily contaminated by radon [$^{222}_{86}$Rn] gas. To determine the radon contamination, you measure the current caused by the radon in an ionisation chamber containing one litre of air from the cave. You

measure 0.1 pA. How much radiation expressed in pico Curie (pCi) per litre is there in the air of the cave? How many radon atoms per litre are there in the air of the cave?

Radon has a half-life of 3.8 days and decays into alpha particles of 5.6 MeV nearly 100% of the time. To simplify the calculation, ignore the fact that radon decay products will also be present and will significantly contribute to the current. Also ignore the fact that often the alpha particle will hit the wall of the ionisation chamber and therefore will not use all its energy to ionise the air.

Solution: The average energy needed to produce one electron–ion pair in air is 33.8 eV. (see Table 4.1). The current produced if there is one alpha particle per second therefore is 2.65×10^{-14} A. The number of alpha particles per second corresponding to 0.1 pA therefore is 3.77 This corresponds to an activity of ≈ 100 pCi. The decay time constant for alpha emission by radon is $\tau = T_{1/2}/\ln 2 = 5.48$ days.

The number on radon atoms in one litre is $3.77 \times \tau$ [s] $= 1.88 \times 10^6$.

3. A GEM detector has a conversion gap of 2 mm. The gas filling is 90% Ar and 10% CH4. Cosmic ray muons are falling perpendicularly on this detector. What is the probability that a muon will be go undetected because there is no primary ionisation event in the conversion gap?

Solution: The number of primary ionisation events in this gas mixture is the weighted average of the number of primary ionisations of the two components. In this case the number of primary ionisations per cm is therefore 26.2/cm. The average number of primary ionisation events in two mm is hence 5.2. The number of primary ionisation events has a Poisson statistics. The probability to have nothing is $e^{-5.2} = 0.005$.

4. Calculate the mobility of nitrogen ions in nitrogen gas assuming that the cross section for the collision is 3.7×10^{-15} cm^2.

Solution: The mean free path is $\lambda = 1/N\sigma$. The number of nitrogen molecules/cm^2 is found from the fact that one mole of gas has a volume of 22.4 l. The mean fee path therefore is ≈ 100 nm. The mobility is obtained from Eq. (4.1). Remember to express all quantities in MKSA units. The result is 3.4 cm^{2}V^{-1}s^{-1}.

Chapter 5

1. Calculate the dopant concentration in n-type silicon with a resistivity of 2000 Ω cm.

Solution: In n-type silicon, the number of electrons is much larger than the number of holes and the number electrons per unit volume is equal to the donor concentration N_d. Equation (5.4) can therefore be written as

$$N_d = \frac{1}{e\,\rho\,\mu_e}$$

If in this equation e is expressed in coulomb ($e = 1.6 \times 10^{-19}$ coulomb), ρ in Ωcm and the mobility μ_e in cm^2/Vs ($\mu_e = 1350$ cm^2/Vs , see Table 5.1), we obtain $N_d = 2.3 \times 10^{12}$ cm^{-3}

2. For a silicon strip detector made starting from n-type silicon with a resistivity of 2000 Ωcm, calculate the voltage to fully deplete a silicon microstrip detector of 300 μm thickness.

Solution: From Eq. (5.5), we get the voltage needed to obtain a given depletion layer

$$ V_0 = \frac{d^2\, e\, N_d}{2\varepsilon_0\varepsilon_r} = \frac{d^2}{\rho\mu_e 2\varepsilon_0\, \varepsilon_r} $$

In this equation, ε_0 is the permittivity (also called dielectric constant) in vacuum ϵ_0, see annex 1 'Physical constants' and ε_r is the dielectric constant of silicon ($\varepsilon_r = 12$, see Table 5.1). The dopant concentration is calculated in Exercise 1. Note that if we take ε_0 in MKSA units, the dopant concentration should be expressed as atoms/m^3 and the thickness d should be in meter. The result is $V_0 = 156$ V.

3. Derive an expression for the energy resolution (FWHM and in %) of silicon as an X-ray detector at room temperature. How much will the energy resolution be for X-rays of 50 keV?

Solution: The number of e–h pairs produced by the X-ray is $n_e = E[\text{eV}]/3.62$.

If the number e–h pairs were Poisson distributed, the r.m.s dispersion on the measured number of charges would be $\sigma = \sqrt{n_e}$. Because the total energy needed to produce the e–h pairs is a large fraction of the energy of the X-ray, this is not a Poisson distribution and the dispersion is given by $\sigma = \sqrt{F\, n_e}$ where F is the Fano factor (see Table 5.1). The signal is proportional to n_e and the variance on the signal is proportional to $\sigma = \sqrt{F\, n_e}$. The relative resolution is given by

$$ = \frac{\sqrt{Fn_e}}{n_e} = \sqrt{\frac{F}{n_e}} $$

We need to multiply this by 2.35 to express the resolution as 'Full Width at Half Maximum' and to multiply it by 100 to express it in %.

The result is therefore

$$ \text{resolution FWHM[\%]} = 235\sqrt{\frac{F\, 3.62}{E[\text{eV}]}} $$

For an X-ray of 50 keV this gives 0.63%.

4. You deposit a very thin layer of 241 Am with an activity of 2 MBq on a surface barrier alpha particle detector. This isotope emits alpha particles, see annex 6

for the characteristics of the emission. Calculate the magnitude of the pulses in number of electrons and the current in the detector.

Solution: The charge in each pulse is ≈5,480,000/3.62 = 1,513,000 electrons. The current is the charge of each pulse times the electron charge, times the number of decays (2,000,000/s) divided by 2 since only half of the alpha particle are emitted in the right direction to penetrate into the detector. The result is 240 nA.

5. Calculate the number of charges produced in a silicon strip detector of 500 μm thick by a minimum ionising particle.

Solution: Use the value for the energy loss and the mean energy per electron–ion pair from Table 5.1. The result is 53,453 electrons

6. Calculate the density of states for an electron enclosed in an infinitely deep and cubic potential well. Use the expression below for the energy levels of the electron in such a potential well. The numbers n_1, n_2 and n_3 are positive integers.

$$E = \frac{p^2}{2m} = \frac{\pi^2 (\hbar c)^2}{a^2 \, 2mc^2} (n_1^2 + n_2^2 + n_3^2)$$

Solution: The energy levels of an electron enclosed in an infinitely deep potential well in one dimension can be found by solving the Schrödinger equation. This problem is solved in any good textbook on quantum mechanics. Here we only give a simple and heuristic derivation of the result.

A wave is associated with every electron and the corresponding wavelength is given by

$$\lambda = \frac{h}{p}$$

The energy eigenstates of the electron correspond to the values of the momentum such that the associated wavelength satisfies

$$n\lambda/2 = a$$

where n is a positive integer and a the dimension of the well. We therefore have the following relations

$$p = \frac{\pi \hbar}{a} n$$

$$E = \frac{p^2}{2m} = \frac{\pi^2 (\hbar c)^2}{a^2 \, 2mc^2} n^2$$

In the three-dimensional well this becomes

$$E = \frac{p^2}{2m} = \frac{\pi^2 (\hbar c)^2}{a^2 \, 2mc^2} (n_1^2 + n_2^2 + n_3^2)$$

The energy eigenstates of the electron are given by the positive integer values of the quantum numbers n_1, n_2 and n_3. To derive the number of electron states in the infinitesimal energy interval $\{E, E+dE\}$, we slightly rewrite this as

$$R^2 = (n_1^2 + n_2^2 + n_3^2) \quad \text{with} \quad R^2 = E \frac{a^2 \, 2mc^2}{\pi^2 \, (\hbar c)^2}$$

Each electron state corresponds to a point of a regular grid in a three-dimensional Euclidian space. The probability to have an electron state in the infinitesimal energy interval $\{E, E+dE\}$ is equal to the volume of one octant of the skin of a sphere with radius R and with thickness dR, divided by the number of points per unit volume. Since the number of points per unit volume is equal to one, the probability to have one electron state in the interval is simply given by.

$$\rho(R) \, dR = \frac{1}{8} \, 4\pi \, R^2 dR$$

To obtain the probability density of states per unit volume in the variable E rather than in the variable R, we only need to make a change of variables and divide by a^3.

$$\rho(E) = 4\pi \left(\frac{2m}{h^2}\right)^{3/2} \sqrt{E} \, dE$$

In this last result we have multiplied the density of states by a factor two to take into account the fact that an electron can have two spin states.

7. Calculate the potential difference over a p–n junction if the dopant concentrations in the n-type silicon and the p-type silicon are $N_D = 10^{12}/\text{cm}^3$ and $N_A = 10^{16}$ cm^3, respectively.

Solution: The potential difference over the junction equals the difference in the Fermi levels between the two parts. In the silicon we have

$$n - type \quad E_F - E_i = kT \, \ln \frac{N_D}{n_i}$$

$$p - type \quad E_i - E_F = kT \, \ln \frac{N_A}{n_i}$$

And therefore the potential difference is

$$(E_F)_n - (E_F)_p = kT \; \ln \frac{N_D \, N_A}{n_i^2} = 0.025 \; \ln \frac{10^{12} \times 10^{16}}{1.5 \, 10^{10} \times 1.5 \, 10^{10}} = 0.44 \, V$$

Chapter 6

1. Assume a detector for gamma rays consisting of LSO scintillator and a photo-multiplier tube. The signal is taken from the anode of the PMT with a 50 Ω coax cable and brought to an oscilloscope. The input impedance of the oscilloscope is 50 Ω and the gain of the PMT is 10^5. What will be the signal amplitude when observing gamma rays of 1 MeV? Note: assume a light collection efficiency 50% and a photocathode quantum efficiency of 25%.

   Solution: From Table 6.2, obtain the light yield and decay time of LSO: 28,000 photons/MeV and 40 ns. The total charge in the pulse is $Q_t = 5.6 \times 10^{-11}$ coulomb.
   The current pulse generated by the PMT is given by

   $$i(t) = Q_t \frac{1}{\tau} e^{-t/\tau}$$

   The peak value of the current pulse is $i = Q_t/\tau$ and the peak voltage

   $$V_{\max} = 50 Q_t/\tau = 70\text{mV}.$$

2. Consider a source emitting gamma rays of 511 keV. Calculate the energy where you expect the backscatter peak in the pulse height spectrum.

   Solution: Using Eq. (2.10), we calculate the energy of the scattered photon for $\theta = 180°$. This gives $511/3 = 170$ keV.
3. Consider a CsI:Tl scintillator. What fraction of the energy lost due to the interactions with the electrons in the material is converted to scintillation light?

   Solution: The CsI:Tl emits 65,000 scintillation photons per MeV of energy lost. The wavelength of the scintillation light is $\lambda = 565$ nm. The energy of each photon is

   $$E = \frac{hc}{\lambda} = \frac{1.242[\text{eV}\,\mu\text{m}]}{0.565} = 2.19\,\text{eV}$$

   The total energy in the scintillation light is 143,000 eV or 14.3% of the energy lost in ionising and exciting electrons.
4. Consider a PIN diode with a quantum efficiency of *Eff* = 60%. Assume it is exposed to a light flux one μW at a wavelength of 565 nm. What will be the photocurrent?

   Solution: The energy of a photon of 565 nm is $E_\lambda = 2.19$ eV (see Exercise 3).
   The number of photons per second is

   $$\frac{P[\text{W}]}{E_\lambda[\text{Joule}]} = \frac{P[\text{W}]}{eE_\lambda\,[\text{eV}]}$$

The current is

$$I = \frac{P[\mathrm{W}]}{e\, E_\lambda[\mathrm{eV}]} e\, \mathit{Eff} = 2.710^{-7}\, A$$

5. Consider a PET scanner with a solid angle covering around its centre point of $\Omega = 10\%$. Assume the detection efficiency for a gamma ray of 511 keV and within the solid angle is $\mathit{Eff} = 20\%$. Assume you place a point source in its centre with an activity of 1 mCi. What will be the single count rate and the coincidence count rate?

Solution: An activity of 1 mCi is the same as 37×10^6 Bq.
Single rate $= 37 \times 10^6 \times 2 \times \Omega \times \mathit{Eff} = 1.48\ 10^6$ Hz
Coincidence rate $= 37 \times 10^6 \times \Omega \times \mathit{Eff}^2 = 1.48\ 10^5$ Hz

6. In several photodetectors the charge multiplication gives rise to an exponential pulse height distribution for single primary charges. Show that in this case the excess noise factor equals 2.

Solution: The pulse height distribution is $f(x)$

$$f(x) = a e^{-ax}$$
$$\langle x \rangle = \int x f(x)\, dx = 1/a$$
$$\langle x^2 \rangle = \int x^2 f(x)\, dx = 2/a^2$$
$$\sigma^2\{x\} = \langle x^2 \rangle - \langle x \rangle^2 = 1/a^2$$
$$F = \left[1 + \frac{\sigma^2\{x\}}{\langle x \rangle^2} \right] = 2$$

Chapter 7

1. Consider the neutron absorption reaction below, assume thermal neutrons.
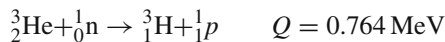
$$^{3}_{2}\mathrm{He} + ^{1}_{0}\mathrm{n} \rightarrow ^{3}_{1}\mathrm{H} + ^{1}_{1}p \qquad Q = 0.764\,\mathrm{MeV}$$

Use energy and momentum conservation to derive the expression giving the kinetic energy for the two final state particles and calculate the value of these for the reaction above.

Solution: In this case the centre-of-mass frame and the laboratory frame are the same. The two final state particles have energy, momentum and mass denoted by E_i, P_i and m_i, $i = 1.2$. Energy and momentum conservation requires

$$\begin{cases} E_1 + E_2 = Q \\ P_1 = P_2 \end{cases}$$

using $P^2 = 2\,mE$, this becomes

$$\begin{cases} E_1 + E_2 = Q \\ 2m_1E_1 = 2m_2E_2 \end{cases}$$

and we readily find

$$E_1 = \frac{m_2 Q}{(m_1 + m_2)}$$

   In the numerical calculation use 938 MeV/c^2 and 2809 MeV/c^2 for the mass of the proton and the tritium. The energy of the proton and tritium are therefore 191 keV and 573 keV, respectively.

2. A commonly used detector for thermal neutrons is a proportional tube filled with ^{3}He gas. Calculate the mean free path of the neutrons in the gas if this gas is at a pressure of 5 atmospheres. If the tube has an inner diameter of 4 cm, what is the probability that a thermal neutron going through its centre will be detected?

   Solution: One mole of gas at standard temperature and pressure has a volume of 22.4 l. The mean free path (m.f.p.) of a neutron is given by (see Chap. 2, Sect. 2.6) mfp $= 1/(N\sigma)$; $N =$ number of scatter centres/cm^3; $\sigma =$ cross section. In the calculation below, I use cm as unit of length. N can be found knowing that 22.4 l of gas at atmospheric pressure contains 6.022×10^{23} atoms of He$_3$. Therefore at 5 atmosphere there are $N = 1.344 \times 10^{20}$ atoms/cm^3. The cross section for the absorption of thermal neutrons in He$_3$ is found on Fig. 8.1: $\sigma = 5000$ barn $= 5 \times 10^{-21}$ cm^2. We therefore have m.f.p. $= 1.49$ cm. The probability to interact in 4 cm is

$$\text{Prob} = \int_0^4 \frac{1}{mfp} e^{-\frac{l}{mfp}} \, dl = 1 - e^{-\frac{4.0}{1.49}} = 0.93$$

3. Calculate the fractional decrease in sensitivity of a self-powered detector with rhodium emitter, after exposure during 6 months to a flux of 3×10^{13} neutrons/cm^2/s.

   Solution: The exact geometry of the rhodium is unimportant as long as the thickness of the rhodium is small compared to the mean free path of the neutrons. We therefore consider a small square of rhodium with surface S and thickness dl. The number of rhodium atoms per unit volume is denoted as N and the total neutron flux as Φ. The probability for a neutron to interact is $P = N\sigma dl$. The number of rhodium atoms lost is therefore

$$\Delta N = N \sigma \, dl \, \Phi$$

and the relative change in sensitivity of the detector is

$$\text{fractional change} \; = \; \frac{\Delta N}{NSdl} = \frac{\sigma \, \Phi}{S} = \sigma \frac{\Phi}{S \, \Delta T} \Delta T$$

The numerical calculation gives for the relative change $= 0.072$. The amount of rhodium lost is of course decreasing exponentially, but because only a small fraction is lost we can use a linear approximation.

4. Derive Eq. (7.1) in Sect. 7.3.2.

Solution: Figure 7.11 shows the geometry of the problem. In this calculation, we can use non-relativistic kinematics; therefore, $2mE = P^2$. We consider the collision in the laboratory frame and use un-primed symbols for the quantities before the collision and primed symbols for the quantities after the collision. Furthermore, we use ϕ to denote the scattering angle of the scattered proton and the Subscript A refers to the target nucleus. Energy and momentum conservation requires

$$\begin{cases} E_n = E'_n + E'_A \\ P_n = P'_A \cos\theta + P'_n \cos\phi \\ 0 = P'_A \sin\theta + P'_n \sin\phi \end{cases}$$

To eliminate the variable ϕ, isolate the term containing the variable ϕ in the 2nd and 3rd equation, square and add.

$$\begin{cases} E_n = E'_n + E'_A \\ P_n^2 - 2P_n P'_A \cos\theta + (P'_A)^2 = (P'_n)^2 \end{cases}$$

Replace P^2 everywhere by $2mE$ and eliminate E'_n using the first equation. After some reshuffling of the terms we get the desired result. E'_A is the recoil energy of the target nucleus.

$$E'_A = E_{\text{recoil}} = E_n \frac{4 m_A m_n}{(m_A + m_n)^2} \cos^2\theta$$

5. Derive Eq. (7.2) in Sect. 7.3.2.

Solution: Figure 7.11 shows the geometry of the problem. To obtain the relation between the scattering angle in the laboratory frame θ and in the centre of mass frame Θ, we notice that in the centre of mass frame (CM frame), the velocity of the nucleon A, both before and after the collision, is equal to the velocity of the centre of mass system in the laboratory frame v_{cm}, The momentum of the recoil nucleus in the CM frame is therefore $m_a v_{\text{cm}}$, Furthermore, the perpendicular component of the momentum of the recoil nucleon is the same in both frames. The symbol v_n denotes the velocity of the initial neutron in the

laboratory frame and P'_A the momentum of the nucleus A in the laboratory frame after the collision.

The value of v_{cm} is found by requiring that the momenta of the neutron and the nucleus A before collision are the same in the CM frame.

$$m_n(v_n - v_{cm}) = m_a v_{cm}$$
$$v_{cm} = \frac{m_n v_n}{(m_A + m_n)}$$

The perpendicular momentum component of the recoil nucleus is found starting from Eq. (7.1) and using $E_{recoil} = (P'_A)^2/2m_A$.

$$P'_A \sin\theta = v_n \frac{2m_A m_n}{(m_A + m_n)} \cos\theta \, \sin\theta$$

The quantity $\sin\Theta$ is given by the ratio of the perpendicular component of the momentum of the target nucleus A divided by the total momentum and we get

$$\sin(\Theta) = 2\cos\theta \, \sin\theta$$

We rewrite this last equation as

$$4\cos^4\theta - 4\cos^2\theta + (1 - \cos^2\Theta) = 0$$

This is a quadratic equation in $\cos^2\theta$. Solving this equation we get

$$\cos^2\theta = \frac{1 \pm \cos\Theta}{2}$$

Simple inspection shows that the solution with the minus sign is the correct one.

6. Consider a proportional tube filled with ^{3}He and used as a slow neutron detector. The gas gain of the tube is 1000 and the capacitance of the anode wire is 100 pF. What will the amplitude of a neutron pulse [in mV] be if the integration time of the pulse is very long?

Solution: The average ionisation energy in He is 41.6 eV, see Table 4.1. The total charge collected on the anode is $Q = 764,000 \times 1000/41.6$.
The voltage is $Q/C = 29$ mV.

Chapter 8

1. Derive property 6 of the transfer function in Sect. 8.2

Solution:

$$V_{in}(t) = \sin(\omega_0 t) = \frac{e^{+j\omega_0 t} - e^{-j\omega_0 t}}{2j}$$

$$V_{\text{out}}(t) = \int\limits_{-\infty}^{+\infty} \sin\left(\omega_0 t'\right) h\left(t - t'\right) dt$$

$$= \int\limits_{-\infty}^{+\infty} \{\frac{e^{+j\omega_0 t'} - e^{-j\omega_0 t'}}{2j} \frac{1}{\sqrt{2\pi}} \int e^{+j\omega(t-t')} H(\omega) d\omega\} dt'$$

There are two terms and we will make the calculation in detail for the first term

$$\text{first term} = \frac{1}{2j\sqrt{2\pi}} \int \{e^{+j\omega_0 t'} \int e^{j\omega(t-t')} H(\omega) d\omega\} dt'$$

$$= \frac{\sqrt{2\pi}}{2j} \int \{\frac{1}{2\pi} \int e^{j(\omega_0-\omega)t'} dt'\} e^{j\omega t} H(\omega) d\omega$$

We recognise the delta function in the expression between {} brackets, and the results is

$$\text{first term} = \frac{\sqrt{2\pi}}{2j} e^{j\omega_0 t} H(\omega_0)$$

After a similar calculation for the second term we obtain

$$V_{\text{out}}(t) = \frac{\sqrt{2\pi}}{2j} \{e^{j\omega_0 t} H(\omega_0) - e^{-j\omega_0 t} H(-\omega_0)\}$$

Writing $H(\omega_0)$ as $H(\omega_0) = |H(\omega_0)| e^{j\varphi(\omega_0)}$

$$V_{\text{out}}(t) = \frac{\sqrt{2\pi}}{2j} |H(\omega_0)| \{e^{j(\omega_0 t+\varphi)} - e^{-j(\omega_0 t+\varphi)}\}$$

$$V_{\text{out}}(t) = \sqrt{2\pi} |H(\omega_0)| \sin\left[\omega_0 t + \varphi(\omega_0)\right]$$

2. Calculate the Fourier transform of Eq. (8.4)

$$H(\omega) = \frac{1}{\sqrt{2\pi}} \frac{1}{1 + j\omega\tau_f} \frac{j\omega\tau}{(1 + j\omega\tau)^2}$$

Solution: In a table with commonly used Fourier transforms, we find

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \frac{1}{1 + j\omega\tau} \quad f(t) = \frac{1}{\tau} e^{-t/\tau} \quad t \geq 0; f(t) = 0 \text{ otherwise}$$

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \frac{1}{(1 + j\omega\tau)^2} \quad f(t) = \frac{1}{\tau^2} e^{-t/\tau} \quad t \geq 0; f(t) = 0 \text{ otherwise}$$

We first decompose the expression given in Eq. (8.4) in a sum of pole terms

$$\sqrt{2\pi} H(\omega) = \frac{1}{1 + j\omega\tau_f} \frac{j\omega\tau}{(1 + j\omega\tau)^2}$$

$$= \frac{A}{1 + j\omega\tau_f} + \frac{B}{1 + j\omega\tau} + \frac{C}{(1 + j\omega\tau)^2}$$

The constants A, B and C are obtained by bringing all terms to a common denominator and requiring the coefficients of the same powers of $j\omega$ to be equal

$$A = -\frac{\tau_f \tau}{(\tau_f - \tau)^2}; \qquad B = \frac{\tau^2}{(\tau_f - \tau)^2}; \qquad C = \frac{\tau_f \tau - \tau^2}{(\tau_f - \tau)^2}$$

Together with the Fourier transforms found in the tables this immediately gives the desired result

$$\begin{cases} \hat{h}(t) = -\dfrac{1}{\tau(\tau_f - \tau)^2} \left[(\tau^2 + t(\tau_f - \tau))e^{-t/\tau} - \tau^2 e^{-t/t_f} \right] \ t \ge 0 \\ \hat{h}(t) = 0 \ t < 0 \end{cases}$$

3. Prove the following theorem.

   Let x_i be a number of independent random variables, all with the same probability distribution. Consider the random variable $R = \sum_i^{1...n} x_i$, where the integer n is itself a random variable with a Poisson distribution with average value λ. The following relations hold:

$$\langle R \rangle = \lambda \langle x \rangle$$
$$\sigma^2\{R\} = \lambda \left\langle x^2 \right\rangle$$

Solution: Consider any random variable x. Its mean value $<x>$, mean square value $<x2>$ and dispersion $\sigma\{x\}$ are related by

$$\sigma^2\{x\} = \left\langle x^2 \right\rangle - <x>^2 .$$

Consider the statistical variable S_N defined as the sum on N independent random variables x_i, all with the same probability distribution. The average value and an average square value of the variable S_N are given by

$$S_N = \sum_i^{1...N} x_i$$

$$\langle S_N \rangle = \sum_i^{1...N} \langle x \rangle = N \langle x \rangle$$

$$\left\langle (S_N)^2 \right\rangle = \left\langle \left(\sum_i^{1...N} x_i \right) \left(\sum_j^{1...N} x_j \right) \right\rangle = \left\langle \sum_i^{1...N} x_i^2 + \sum_{i \ne j}^{1...N} x_i x_j \right\rangle$$

$$= N \left\langle x^2 \right\rangle + N(N - 1) \langle x \rangle^2$$

Consider the statistical variable R defined as the sum of n independent random variables x_i, all with the same probability distribution and where n itself is a random variable and P_k is the probability that $n = k$. The average value and an average square value of the variable R are given by

$$\langle R \rangle = \sum_{n}^{0...\infty} P_n \langle S_n \rangle$$
$$= \langle x \rangle \sum_{n}^{0...\infty} n P_n$$
$$= \langle x \rangle \langle n \rangle$$
$$\langle R^2 \rangle = \sum_{n}^{0...\infty} P_n \left\langle (S_n)^2 \right\rangle$$
$$= \langle x^2 \rangle \sum_{n}^{0...\infty} n P_n + \langle x^2 \rangle^2 \sum_{n}^{0...\infty} n(n-1) P_n$$
$$= \langle x^2 \rangle \langle n \rangle + \langle x \rangle^2 [\langle n^2 \rangle - \langle n \rangle]$$

The dispersion of the variable R is therefore given by

$$\sigma^2\{R\} = \langle R^2 \rangle - \langle R \rangle^2 = \langle x^2 \rangle \langle n \rangle + \langle x \rangle^2 (\langle n^2 \rangle - \langle n \rangle - \langle n \rangle^2)$$

For a Poisson distribution we have

$$\sigma\{n\} = \langle n^2 \rangle - \langle n \rangle^2 = \langle n \rangle$$

The dispersion of R therefore reduces to

$$\sigma^2\{R\} = \langle n \rangle \langle x^2 \rangle$$

4. Assume you are measuring the noise voltage of a resistor using a digital oscilloscope with an input impedance of 10 MΩ and a bandwidth of 400 MHz. For what value of the resistor will you measure the largest value for the noise. How much will this maximum noise be in mV?

Solution: The square of the noise is proportional to

$$\frac{R_{inp}^2 R}{(R + R_{inp})}$$

The value where this expression reaches its maximum is given by the condition that the derivative to R equals zero. Therefore the maximum is reached for $R = R_{inp}$.

The square root of the average noise squared is 5.7 mV

5. Consider a silicon strip detector where each strip has a capacitance of 20 pF and a dark current of 20 nA. The rise-time of the pulse is 30 ns. Give an under limit for the noise. Take the shape coefficients $a_1 = a_2 = 1$.

Solution: The contribution from the front end FET is given by Eq. (8.10). Take 0.25 ns for the charge transit time in the FET. Divide the value of ENC by the charge of one electron to obtain the result in number or electrons. The result is 527 electrons. The contribution of the dark current to the noise is given by Eq. (8.11). The result is ≈ 62 electrons. Combine the two noise contributions quadratically.

The total noise is therefore at least ≈ 530 electrons.

6. Prove that the stationary solutions for a transmission line of length D are given by the following equations

$$\begin{cases} I_n(x,t) = I_n \sin\left(\dfrac{n\pi x}{D}\right) \sin\left(\dfrac{n\pi v_0 t}{D} + \varphi_n\right) \\ V_n(x,t) = I_n Z_0 \cos\left(\dfrac{n\pi x}{D}\right) \cos\left(\dfrac{n\pi v_0 t}{D} + \varphi_n\right) \end{cases} \quad n = 1,..\infty$$

Solution: The equations of the transmission line are:

$$\begin{cases} \dfrac{d^2 V(x,t)}{dx^2} = LC\dfrac{d^2 V(x,t)}{dt^2} \\ \dfrac{d^2 I(x,t)}{dx^2} = LC\dfrac{d^2 I(x,t)}{dt^2} \end{cases} \tag{1}$$

$$\begin{cases} \dfrac{dV(x,t)}{dx} = -L\dfrac{dI(x,t)}{dt} \\ \dfrac{dI(x,t)}{dx} = -C\dfrac{dV(x,t)}{dt} \end{cases} \tag{2}$$

The velocity of the line v_0 and the characteristic impedance Z_0 are given by

$$v_0 = \frac{1}{\sqrt{LC}} \qquad Z_0 = \sqrt{\frac{L}{C}}$$

The stationary solutions are of the form

$$V(x,t) = V'(x).V''(t) \tag{3}$$

$$I(x,t) = I'(x).I''(t) \tag{4}$$

Let us substitute expression 4 in Eq. (1). This gives an equation where we can separate the variables and the two parts must be equal to a constant. Solutions

satisfying the boundary conditions can only be obtained if this constant is negative, and we therefore write this constant as $-k^2$.

$$\frac{1}{I'(x)}\frac{d^2 I'(x)}{dx^2} = LC\frac{1}{I''(t)}\frac{d^2 I''(t)}{dt^2} = -k^2$$

This is equivalent to the two following equations

$$\frac{d^2 I'}{dx^2} = -k^2 I'(x)$$

$$\frac{d^2 I''}{dt^2} = -k^2\frac{1}{LC}I''(t) = -k^2 v_0^2 I''(t)$$

The most general solution to these equations is (using the convention $j^2 = -1$)

$$I''(t) = Ae^{+jkv_0 t} + A'e^{-jkv_0 t}$$
$$I'(x) = Be^{+jkx} + B'e^{-jkx}$$

Imposing the boundary condition $I(x,t) = 0$ for $x = 0$, we have $B = -B'$, and the second equation is written as

$$I'(x) = jB \sin kx$$

The boundary condition $I(x,t) = 0$ for $x = D$ gives

$$k = n\pi/D, \text{ where } n \text{ is an integer}$$

The most general solution satisfying these boundary conditions hence is

$$I(x,t) = jB \sin\left(\frac{n\pi}{D}t\right)\left[(A + A')\cos\left(\frac{n\pi}{D}v_0 t\right) + j(A - A')\sin\left(\frac{n\pi}{D}v_0 t\right)\right]$$

The function $I(x,t)$ must be real, therefore

$$jB(A + A') = K_1$$
$$-B(A - A') = K_2$$

where K_1 and K_2 are arbitrary real constants.

The most general solution satisfying all conditions therefore is

$$I(x,t) = \sin\left(\frac{n\pi x}{D}\right)\left[K_1 \cos\left(\frac{n\pi}{D}v_0 t\right) + K_2 \sin\left(\frac{n\pi}{D}v_0 t\right)\right]$$

It is always possible to write the constants K_1 and K_2 as

$$\begin{cases} K_1 = K \sin\varphi \\ K_2 = K \cos\varphi \end{cases} \Rightarrow \begin{cases} K^2 = K_1^2 + K_2^2 \\ tg\varphi = \dfrac{K_1}{K_2} \end{cases}$$

The solution can hence be written as

$$I(x,t) = K \sin\left(\frac{n\pi}{D}x\right) \left[\sin\varphi \cos\left(\frac{n\pi}{D}v_0 t\right) + \cos\varphi \sin\left(\frac{n\pi}{D}v_0 t\right)\right]$$

$$I(x,t) = K \sin\left(\frac{n\pi}{D}x\right) \sin\left(\frac{n\pi}{D}v_0 t + \varphi\right)$$

The expression for the potential in the transmission line is obtained from

$$\frac{dV}{dx} = -L\frac{dI}{dt}$$

$$\frac{dV}{dx} = -LK \sin\left(\frac{n\pi}{D}x\right)\left(\frac{n\pi}{D}v_0\right) \cos\left(\frac{n\pi}{D}v_0 t + \varphi\right)$$

$$V(x,t) = \int \frac{dV}{dx} dx = -LKv_0 \cos\left(\frac{n\pi}{D}v_0 t + \varphi\right) \int \sin\left(\frac{n\pi}{D}x\right) \frac{n\pi}{D} dx$$

$$= KZ_0 \cos\left(\frac{n\pi}{D}v_0 t + \varphi\right) \cos\left(\frac{n\pi}{D}x\right)$$

We omitted the integration constant because for any system the potential is only meaningful up to an arbitrary constant.

Annex 1 : Physical Constants

Table 1.1. Reviewed 2007 by P.J. Mohr and B.N. Taylor (NIST). Based mainly on the "CODATA Recommended Values of the Fundamental Physical Constants: 2006" by P.J. Mohr, B.N. Taylor, and D.B. Newell (to be published in Rev. Mod. Phys, and J. Phys. Chem. Ref. Data). The last group of constants (beginning with the Fermi coupling constant) comes from the Particle Data Group. The figures in parentheses after the values give the 1-standard-deviation uncertainties in the last digits; the corresponding fractional uncertainties in parts per 10^9 (ppb) are given in the last column. This set of constants (aside from the last group) is recommended for international use by CODATA (the Committee on Data for Science and Technology). The full 2006 CODATA set of constants may be found at `http://physics.nist.gov/constants`.

Quantity	Symbol, equation	Value	Uncertainty (ppb)
speed of light in vacuum	c	299 792 458 m s^{-1}	exact*
Planck constant	h	6.626 068 96(33)$\times10^{-34}$ J s	50
Planck constant, reduced	$\hbar \equiv h/2\pi$	1.054 571 628(53)$\times10^{-34}$ J s	50
		$= 6.582\ 118\ 99(16)\times10^{-22}$ MeV s	25
electron charge magnitude	e	1.602 176 487(40)$\times10^{-19}$ C = 4.803 204 27(12)$\times10^{-10}$ esu	25, 25
conversion constant	$\hbar c$	197.326 9631(49) MeV fm	25
conversion constant	$(\hbar c)^2$	0.389 379 304(19) GeV^2 mbarn	50
electron mass	m_e	0.510 998 910(13) MeV/c^2 = 9.109 382 15(45)$\times10^{-31}$ kg	25, 50
proton mass	m_p	938.272 013(23) MeV/c^2 = 1.672 621 637(83)$\times10^{-27}$ kg	25, 50
		$= 1.007\ 276\ 466\ 77(10)$ u = 1836.152 672 47(80) m_e	0.10, 0.43
deuteron mass	m_d	1875.612 793(47) MeV/c^2	25
unified atomic mass unit (u)	(mass ^{12}C atom)/12 = (1 g)/(N_A mol)	931.494 028(23) MeV/c^2 = 1.660 538 782(83)$\times10^{-27}$ kg	25, 50
permittivity of free space	$\epsilon_0 = 1/\mu_0 c^2$	8.854 187 817 ... $\times10^{-12}$ F m^{-1}	exact
permeability of free space	μ_0	$4\pi \times 10^{-7}$ N A^{-2} = 12.566 370 614 ... $\times10^{-7}$ N A^{-2}	exact
fine-structure constant	$\alpha = e^2/4\pi\epsilon_0\hbar c$	7.297 352 5376(50)$\times10^{-3}$ = 1/137.035 999 679(94)†	0.68, 0.68
classical electron radius	$r_e = e^2/4\pi\epsilon_0 m_e c^2$	2.817 940 2894(58)$\times10^{-15}$ m	2.1
(e^- Compton wavelength)/2π	$\lambdabar_e = \hbar/m_e c = r_e\alpha^{-1}$	3.861 592 6459(53)$\times10^{-13}$ m	1.4
Bohr radius ($m_{\rm nucleus} = \infty$)	$a_\infty = 4\pi\epsilon_0\hbar^2/m_e e^2 = r_e\alpha^{-2}$	0.529 177 208 59(36)$\times10^{-10}$ m	0.68
wavelength of 1 eV/c particle	$hc/(1\ {\rm eV})$	1.239 841 875(31)$\times10^{-6}$ m	25
Rydberg energy	$hcR_\infty = m_e e^4/2(4\pi\epsilon_0)^2\hbar^2 = m_e c^2\alpha^2/2$	13.605 691 93(34) eV	25
Thomson cross section	$\sigma_T = 8\pi r_e^2/3$	0.665 245 8558(27) barn	4.1
Bohr magneton	$\mu_B = e\hbar/2m_e$	5.788 381 7555(79)$\times10^{-11}$ MeV T^{-1}	1.4
nuclear magneton	$\mu_N = e\hbar/2m_p$	3.152 451 2326(45)$\times10^{-14}$ MeV T^{-1}	1.4
electron cyclotron freq./field	$\omega^e_{\rm cycl}/B = e/m_e$	1.758 820 150(44)$\times10^{11}$ rad s^{-1} T^{-1}	25
proton cyclotron freq./field	$\omega^p_{\rm cycl}/B = e/m_p$	9.578 833 92(24)$\times10^{7}$ rad s^{-1} T^{-1}	25
gravitational constant‡	G_N	6.674 28(67)$\times10^{-11}$ m^3 kg^{-1} s^{-2}	1.0×10^5
		$= 6.708\ 81(67)\times10^{-39}\ \hbar c$ (GeV/c^2)$^{-2}$	1.0×10^5
standard gravitational accel.	g_N	9.806 65 m s^{-2}	exact
Avogadro constant	N_A	6.022 141 79(30)$\times10^{23}$ mol^{-1}	50
Boltzmann constant	k	1.380 6504(24)$\times10^{-23}$ J K^{-1}	1700
		$= 8.617\ 343(15)\times10^{-5}$ eV K^{-1}	1700
molar volume, ideal gas at STP	$N_A k(273.15\ {\rm K})/(101\ 325\ {\rm Pa})$	22.413 996(39)$\times10^{-3}$ m^3 mol^{-1}	1700
Wien displacement law constant	$b = \lambda_{\rm max}T$	2.897 7685(51)$\times10^{-3}$ m K	1700
Stefan-Boltzmann constant	$\sigma = \pi^2 k^4/60\hbar^3 c^2$	5.670 400(40)$\times10^{-8}$ W m^{-2} K^{-4}	7000
Fermi coupling constant**	$G_F/(\hbar c)^3$	1.166 37(1)$\times10^{-5}$ GeV^{-2}	9000
weak-mixing angle	$\sin^2 \hat{\theta}(M_Z)$ ($\overline{\rm MS}$)	0.231 19(14)††	6.5×10^5
$W^\pm$ boson mass	m_W	80.398(25) GeV/c^2	3.6×10^5
Z^0 boson mass	m_Z	91.1876(21) GeV/c^2	2.3×10^4
strong coupling constant	$\alpha_s(m_Z)$	0.1176(20)	1.7×10^7

$\pi = 3.141\ 592\ 653\ 589\ 793\ 238$	$e = 2.718\ 281\ 828\ 459\ 045\ 235$	$\gamma = 0.577\ 215\ 664\ 901\ 532\ 861$

1 in $\equiv 0.0254$ m	1 G $\equiv 10^{-4}$ T	1 eV $= 1.602\ 176\ 487(40) \times 10^{-19}$ J	kT at 300 K $= [38.681\ 685(68)]^{-1}$ eV
1 Å $\equiv 0.1$ nm	1 dyne $\equiv 10^{-5}$ N	1 eV/$c^2 = 1.782\ 661\ 758(44) \times 10^{-36}$ kg	0 °C $\equiv 273.15$ K
1 barn $\equiv 10^{-28}$ m^2	1 erg $\equiv 10^{-7}$ J 2.997 924 58 $\times 10^9$ esu = 1 C		1 atmosphere $\equiv 760$ Torr $\equiv 101\ 325$ Pa

* The meter is the length of the path traveled by light in vacuum during a time interval of 1/299 792 458 of a second.
† At $Q^2 = 0$. At $Q^2 \approx m_W^2$ the value is $\sim 1/128$.
‡ Absolute lab measurements of G_N have been made only on scales of about 1 cm to 1 m.
** See the discussion in Sec. 10, "Electroweak model and constraints on new physics."
†† The corresponding $\sin^2 \theta$ for the effective angle is 0.23149(13).

Annex 2 : International System of Units

Physical quantity	Name of unit	Symbol
Base units		
length	meter	m
mass	kilogram	kg
time	second	s
electric current	ampere	A
thermodynamic temperature	kelvin	K
amount of substance	mole	mol
luminous intensity	candela	cd
Derived units with special names		
plane angle	radian	rad
solid angle	steradian	sr
frequency	hertz	Hz
energy	joule	J
force	newton	N
pressure	pascal	Pa
power	watt	W
electric charge	coulomb	C
electric potential	volt	V
electric resistance	ohm	Ω
electric conductance	siemens	S
electric capacitance	farad	F
magnetic flux	weber	Wb
inductance	henry	H
magnetic flux density	tesla	T
luminous flux	lumen	lm
illuminance	lux	lx
celsius temperature	degree celsius	°C
activity (of a radioactive source)*	becquerel	Bq
absorbed dose (of ionizing radiation)*	gray	Gy
dose equivalent*	sievert	Sv

* See Sect. 3.2.

Annex 3 : Atomic and Molecular Properties of Materials

Table 6.1 Abridged from `pdg.lbl.gov/AtomicNuclearProperties` by D. E. Groom (2007). See web pages for more detail about entries in this table including chemical formulae, and for several hundred other entries. Quantities in parentheses are for NTP (20° C and 1 atm), and square brackets indicate quantities evaluated at STP. Boiling points are at 1 atm. Refractive indices n are evaluated at the sodium D line blend (589.2 nm); values $\gg$1 in brackets are for $(n-1) \times 10^6$ (gases).

Material	Z	A	$\langle Z/A \rangle$	Nucl.coll. length λ_T {g cm^{-2}}	Nucl.inter. length λ_I {g cm^{-2}}	Rad.len. X_0 {g cm^{-2}}	$dE/dx\|_{min}$ { MeV g^{-1}cm^2}	Density {g cm^{-3}} ({gℓ^{-1}})	Melting point (K)	Boiling point (K)	Refract. index (@ Na D)
H$_2$	1	1.00794(7)	0.99212	42.8	52.0	63.04	(4.103)	0.071(0.084)	13.81	20.28	1.11[132.]
D$_2$	1	2.01410177803(8)	0.49650	51.3	71.8	125.97	(2.053)	0.169(0.168)	18.7	23.65	1.11[138.]
He	2	4.002602(2)	0.49967	51.8	71.0	94.32	(1.937)	0.125(0.166)		4.220	1.02[35.0]
Li	3	6.941(2)	0.43221	52.2	71.3	82.78	1.639	0.534	453.6	1615.	
Be	4	9.012182(3)	0.44384	55.3	77.8	65.19	1.595	1.848	1560.	2744.	
C diamond	6	12.0107(8)	0.49955	59.2	85.8	42.70	1.725	3.520			2.42
C graphite	6	12.0107(8)	0.49955	59.2	85.8	42.70	1.742	2.210			
N$_2$	7	14.0067(2)	0.49976	61.1	89.7	37.99	(1.825)	0.807(1.165)	63.15	77.29	1.20[298.]
O$_2$	8	15.9994(3)	0.50002	61.3	90.2	34.24	(1.801)	1.141(1.332)	54.36	90.20	1.22[271.]
F$_2$	9	18.9984032(5)	0.47372	65.0	97.4	32.93	(1.676)	1.507(1.580)	53.53	85.03	[195.]
Ne	10	20.1797(6)	0.49555	65.7	99.0	28.93	(1.724)	1.204(0.839)	24.56	27.07	1.09[67.1]
Al	13	26.9815386(8)	0.48181	69.7	107.2	24.01	1.615	2.699	933.5	2792.	
Si	14	28.0855(3)	0.49848	70.2	108.4	21.82	1.664	2.329	1687.	3538.	3.95
Cl$_2$	17	35.453(2)	0.47951	73.8	115.7	19.28	(1.630)	1.574(2.980)	171.6	239.1	[773.]
Ar	18	39.948(1)	0.45059	75.7	119.7	19.55	(1.519)	1.396(1.662)	83.81	87.26	1.23[281.]
Ti	22	47.867(1)	0.45961	78.8	126.2	16.16	1.477	4.540	1941.	3560.	
Fe	26	55.845(2)	0.46557	81.7	132.1	13.84	1.451	7.874	1811.	3134.	
Cu	29	63.546(3)	0.45636	84.2	137.3	12.86	1.403	8.960	1358.	2835.	
Ge	32	72.64(1)	0.44053	86.9	143.0	12.25	1.370	5.323	1211.	3106.	
Sn	50	118.710(7)	0.42119	98.2	166.7	8.82	1.263	7.310	505.1	2875.	
Xe	54	131.293(6)	0.41129	100.8	172.1	8.48	(1.255)	2.953(5.483)	161.4	165.1	1.39[701.]
W	74	183.84(1)	0.40252	110.4	191.9	6.76	1.145	19.300	3695.	5828.	
Pt	78	195.084(9)	0.39983	112.2	195.7	6.54	1.128	21.450	2042.	4098.	
Au	79	196.966569(4)	0.40108	112.5	196.3	6.46	1.134	19.320	1337.	3129.	
Pb	82	207.2(1)	0.39575	114.1	199.6	6.37	1.122	11.350	600.6	2022.	
U	92	[238.02891(3)]	0.38651	118.6	209.0	6.00	1.081	18.950	1408.	4404.	
Air (dry, 1 atm)			0.49919	61.3	90.1	36.62	(1.815)	(1.205)		78.80	
Shielding concrete			0.50274	65.1	97.5	26.57	1.711	2.300			
Borosilicate glass (Pyrex)			0.49707	64.6	96.5	28.17	1.696	2.230			
Lead glass			0.42101	95.9	158.0	7.87	1.255	6.220			
Standard rock			0.50000	66.8	101.3	26.54	1.688	2.650			
Methane (CH$_4$)			0.62334	54.0	73.8	46.47	(2.417)	(0.667)	90.68	111.7	[444.]
Ethane (C$_2$H$_6$)			0.59861	55.0	75.9	45.66	(2.304)	(1.263)	90.36	184.5	
Propane (C$_3$H$_8$)			0.58962	55.3	76.7	45.37	(2.262)	0.493(1.868)	85.52	231.0	
Butane (C$_4$H$_{10}$)			0.59497	55.5	77.1	45.23	(2.278)	(2.489)	134.9	272.6	
Octane (C$_8$H$_{18}$)			0.57778	55.8	77.8	45.00	2.123	0.703	214.4	398.8	
Paraffin (CH$_3$(CH$_2$)$_{n\approx23}$CH$_3$)			0.57275	56.0	78.3	44.85	2.088	0.930			
Nylon (type 6, 6/6)			0.54790	57.5	81.6	41.92	1.973	1.18			
Polycarbonate (Lexan)			0.52697	58.3	83.6	41.50	1.886	1.20			
Polyethylene ([CH$_2$CH$_2$]$_n$)			0.57034	56.1	78.5	44.77	2.079	0.89			
Polyethylene terephthalate (Mylar)			0.52037	58.9	84.9	39.95	1.848	1.40			
Polyimide film (Kapton)			0.51264	59.2	85.5	40.58	1.820	1.42			
Polymethylmethacrylate (acrylic)			0.53937	58.1	82.8	40.55	1.929	1.19			1.49
Polypropylene			0.55998	56.1	78.5	44.77	2.041	0.90			
Polystyrene ([C$_6$H$_5$CHCH$_2$]$_n$)			0.53768	57.5	81.7	43.79	1.936	1.06			1.59
Polytetrafluoroethylene (Teflon)			0.47992	63.5	94.4	34.84	1.671	2.20			
Polyvinyltoluene			0.54141	57.3	81.3	43.90	1.956	1.03			1.58
Aluminum oxide (sapphire)			0.49038	65.5	98.4	27.94	1.647	3.970	2327.	3273.	1.77
Barium flouride (BaF$_2$)			0.42207	90.8	149.0	9.91	1.303	4.893	1641.	2533.	1.47
Bismuth germanate (BGO)			0.42065	96.2	159.1	7.97	1.251	7.130	1317.		2.15
Carbon dioxide gas (CO$_2$)			0.49989	60.7	88.9	36.20	1.819	(1.842)			[449.]
Solid carbon dioxide (dry ice)			0.49989	60.7	88.9	36.20	1.787	1.563	Sublimes at 194.7 K		
Cesium iodide (CsI)			0.41569	100.6	171.5	8.39	1.243	4.510	894.2	1553.	1.79
Lithium fluoride (LiF)			0.46262	61.0	88.7	39.26	1.614	2.635	1121.	1946.	1.39
Lithium hydride (LiH)			0.50321	50.8	68.1	79.62	1.897	0.820	965.		
Lead tungstate (PbWO$_4$)			0.41315	100.6	168.3	7.39	1.229	8.300	1403.		2.20
Silicon dioxide (SiO$_2$, fused quartz)			0.49930	65.2	97.8	27.05	1.699	2.200	1986.	3223.	1.46
Sodium chloride (NaCl)			0.55509	71.2	110.1	21.91	1.847	2.170	1075.	1738.	1.54
Sodium iodide (NaI)			0.42697	93.1	154.6	9.49	1.305	3.667	933.2	1577.	1.77
Water (H$_2$O)			0.55509	58.5	83.3	36.08	1.992	1.000(0.756)	273.1	373.1	1.33
Silica aerogel			0.50093	65.0	97.3	27.25	1.740	0.200	(0.03 H$_2$O, 0.97 SiO$_2$)		

Material	Dielectric constant $(\kappa = \epsilon/\epsilon_0)$ () is $(\kappa-1)\times10^6$ for gas	Young's modulus [10^6 psi]	Coeff. of thermal expansion [10^{-6}cm/cm-°C]	Specific heat [cal/g-°C]	Electrical resistivity [$\mu\Omega$cm(@°C)]	Thermal conductivity [cal/cm-°C-sec]
H_2	(253.9)	—	—	—	—	—
He	(64)	—	—	—	—	—
Li	—	—	56	0.86	8.55(0°)	0.17
Be	—	37	12.4	0.436	5.885(0°)	0.38
C	—	0.7	0.6–4.3	0.165	1375(0°)	0.057
N_2	(548.5)	—	—	—	—	—
O_2	(495)	—	—	—	—	—
Ne	(127)	—	—	—	—	—
Al	—	10	23.9	0.215	2.65(20°)	0.53
Si	11.9	16	2.8–7.3	0.162	—	0.20
Ar	(517)	—	—	—	—	—
Ti	—	16.8	8.5	0.126	50(0°)	—
Fe	—	28.5	11.7	0.11	9.71(20°)	0.18
Cu	—	16	16.5	0.092	1.67(20°)	0.94
Ge	16.0	—	5.75	0.073	—	0.14
Sn	—	6	20	0.052	11.5(20°)	0.16
Xe	—	—	—	—	—	—
W	—	50	4.4	0.032	5.5(20°)	0.48
Pt	—	21	8.9	0.032	9.83(0°)	0.17
Pb	—	2.6	29.3	0.038	20.65(20°)	0.083
U	—	—	36.1	0.028	29(20°)	0.064

Annex 4 : Periodic Table of Elements

Table 4.1. Revised 2008 by C.G. Wohl (LBNL), D.E. Groom (LBNL), and E. Bergren. Atomic weights of stable elements are adapted from the Commission on Isotopic Abundances and Atomic Weights, "Atomic Weights of the Elements 2007," http://www.chem.qmul.ac.uk/iupac/AtWt/. The atomic number (top left) is the number of protons in the nucleus. The atomic mass (bottom) of a stable elements is weighted by isotopic abundances in the Earth's surface. If the element has no stable isotope, the atomic mass (in parentheses) of the most stable isotope currently known is given. In this case the mass is from http://www.nmdc.bnl.gov/amdc/masstables/Ame2003/mass.mas03 and the longest-lived isotope is from www.nmdc.bnl.gov/ensdf/za.form.jsp. The exceptions are Th, Pa, and U, which do have characteristic terrestrial compositions. Atomic masses are relative to the mass of ^{12}C, defined to be exactly 12 unified atomic mass units (u) (approx. g/mole). Relative isotopic abundances often vary considerably, both in natural and commercial samples; this is reflected in the number of significant figures given. As of early 2008 element 112 has not been assigned a name, and there are no confirmed elements with $Z > 112$.

PERIODIC TABLE OF THE ELEMENTS

1 IA	2 IIA	3 IIIB	4 IVB	5 VB	6 VIB	7 VIIB	8 VIII	9 VIII	10 VIII	11 IB	12 IIB	13 IIIA	14 IVA	15 VA	16 VIA	17 VIIA	18 VIIIA
1 H Hydrogen 1.00794																	2 He Helium 4.002602
3 Li Lithium 6.941	4 Be Beryllium 9.012182											5 B Boron 10.811	6 C Carbon 12.0107	7 N Nitrogen 14.0067	8 O Oxygen 15.9994	9 F Fluorine 18.9984032	10 Ne Neon 20.1797
11 Na Sodium 22.98976928	12 Mg Magnesium 24.3050											13 Al Aluminum 26.9815386	14 Si Silicon 28.0855	15 P Phosph. 30.973762	16 S Sulfur 32.065	17 Cl Chlorine 35.453	18 Ar Argon 39.948
19 K Potassium 39.0983	20 Ca Calcium 40.078	21 Sc Scandium 44.955912	22 Ti Titanium 47.867	23 V Vanadium 50.9415	24 Cr Chromium 51.9961	25 Mn Manganese 54.938045	26 Fe Iron 55.845	27 Co Cobalt 58.933195	28 Ni Nickel 58.6934	29 Cu Copper 63.546	30 Zn Zinc 65.38	31 Ga Gallium 69.723	32 Ge German. 72.64	33 As Arsenic 74.92160	34 Se Selenium 78.96	35 Br Bromine 79.904	36 Kr Krypton 83.798
37 Rb Rubidium 85.4678	38 Sr Strontium 87.62	39 Y Yttrium 88.90585	40 Zr Zirconium 91.224	41 Nb Niobium 92.90638	42 Mo Molybd. 95.96	43 Tc Technet. (97.90722)	44 Ru Ruthen. 101.07	45 Rh Rhodium 102.90550	46 Pd Palladium 106.42	47 Ag Silver 107.8682	48 Cd Cadmium 112.411	49 In Indium 114.818	50 Sn Tin 118.710	51 Sb Antimony 121.760	52 Te Tellurium 127.60	53 I Iodine 126.90447	54 Xe Xenon 131.293
55 Cs Cesium 132.9054519	56 Ba Barium 137.327	57–71 Lantha-nides	72 Hf Hafnium 178.49	73 Ta Tantalum 180.94788	74 W Tungsten 183.84	75 Re Rhenium 186.207	76 Os Osmium 190.23	77 Ir Iridium 192.217	78 Pt Platinum 195.084	79 Au Gold 196.966569	80 Hg Mercury 200.59	81 Tl Thallium 204.3833	82 Pb Lead 207.2	83 Bi Bismuth 208.98040	84 Po Polonium (208.98243)	85 At Astatine (209.98715)	86 Rn Radon (222.01758)
87 Fr Francium (223.01974)	88 Ra Radium (226.02541)	89–103 Actinides	104 Rf Rutherford. (267.122)	105 Db Dubnium (268.125)	106 Sg Seaborg. (271.133)	107 Bh Bohrium (270.134)	108 Hs Hassium (269.134)	109 Mt Meitner. (276.151)	110 Ds Darmstadt. (281.162)	111 Rg Roentgen. (280.164)	112 (285.174)						

Lanthanide series	57 La Lanthan. 138.90547	58 Ce Cerium 140.116	59 Pr Praseodym. 140.90765	60 Nd Neodym. 144.242	61 Pm Prometh. (144.91275)	62 Sm Samarium 150.36	63 Eu Europium 151.964	64 Gd Gadolin. 157.25	65 Tb Terbium 158.92535	66 Dy Dyspros. 162.500	67 Ho Holmium 164.93032	68 Er Erbium 167.259	69 Tm Thulium 168.93421	70 Yb Ytterbium 173.054	71 Lu Lutetium 174.9668
Actinide series	89 Ac Actinium (227.02775)	90 Th Thorium 232.03806	91 Pa Protactin. 231.03588	92 U Uranium 238.02891	93 Np Neptunium (237.04817)	94 Pu Plutonium (244.06420)	95 Am Americ. (243.06138)	96 Cm Curium (247.07035)	97 Bk Berkelium (247.07031)	98 Cf Californ. (251.07959)	99 Es Einstein. (252.0830)	100 Fm Fermium (257.09510)	101 Md Mendelev. (258.09843)	102 No Nobelium (259.1010)	103 Lr Lawrenc. (262.110)

Annex 5 : Electromagnetic Relations

Quantity	Gaussian CGS	SI
Conversion factors: Charge: Potential: Magnetic field:	2.99792458×10^9 esu $(1/299.792458)$ statvolt (ergs/esu) 10^4 gauss $= 10^4$ dyne/esu	$= 1$ C $= 1$ A s $= 1$ V $= 1$ J C^{-1} $= 1$ T $= 1$ N A^{-1}m^{-1}
	$\mathbf{F} = q\left(\mathbf{E} + \dfrac{\mathbf{v}}{c} \times \mathbf{B}\right)$	$\mathbf{F} = q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right)$
	$\nabla \cdot \mathbf{D} = 4\pi\rho$ $\nabla \times \mathbf{H} - \dfrac{1}{c}\dfrac{\partial \mathbf{D}}{\partial t} = \dfrac{4\pi}{c}\mathbf{J}$ $\nabla \cdot \mathbf{B} = 0$ $\nabla \times \mathbf{E} + \dfrac{1}{c}\dfrac{\partial \mathbf{B}}{\partial t} = 0$	$\nabla \cdot \mathbf{D} = \rho$ $\nabla \times \mathbf{H} - \dfrac{\partial \mathbf{D}}{\partial t} = \mathbf{J}$ $\nabla \cdot \mathbf{B} = 0$ $\nabla \times \mathbf{E} + \dfrac{\partial \mathbf{B}}{\partial t} = 0$
Constitutive relations:	$\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}, \quad \mathbf{H} = \mathbf{B} - 4\pi\mathbf{M}$	$\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}, \quad \mathbf{H} = \mathbf{B}/\mu_0 - \mathbf{M}$
Linear media:	$\mathbf{D} = \epsilon\mathbf{E}, \quad \mathbf{H} = \mathbf{B}/\mu$ 1 1	$\mathbf{D} = \epsilon\mathbf{E}, \quad \mathbf{H} = \mathbf{B}/\mu$ $\epsilon_0 = 8.854187\ldots \times 10^{-12}$ F m^{-1} $\mu_0 = 4\pi \times 10^{-7}$ N A^{-2}
	$\mathbf{E} = -\nabla V - \dfrac{1}{c}\dfrac{\partial \mathbf{A}}{\partial t}$ $\mathbf{B} = \nabla \times \mathbf{A}$	$\mathbf{E} = -\nabla V - \dfrac{\partial \mathbf{A}}{\partial t}$ $\mathbf{B} = \nabla \times \mathbf{A}$
	$V = \displaystyle\sum_{\text{charges}} \dfrac{q_i}{r_i} = \int \dfrac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d^3x'$ $\mathbf{A} = \dfrac{1}{c}\oint \dfrac{I\, d\boldsymbol{\ell}}{|\mathbf{r}-\mathbf{r}'|} = \dfrac{1}{c}\int \dfrac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d^3x'$	$V = \dfrac{1}{4\pi\epsilon_0}\displaystyle\sum_{\text{charges}} \dfrac{q_i}{r_i} = \dfrac{1}{4\pi\epsilon_0}\int \dfrac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d^3x'$ $\mathbf{A} = \dfrac{\mu_0}{4\pi}\oint \dfrac{I\, d\boldsymbol{\ell}}{|\mathbf{r}-\mathbf{r}'|} = \dfrac{\mu_0}{4\pi}\int \dfrac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d^3x'$
	$\mathbf{E}'_{\parallel} = \mathbf{E}_{\parallel}$ $\mathbf{E}'_{\perp} = \gamma\left(\mathbf{E}_{\perp} + \dfrac{1}{c}\mathbf{v} \times \mathbf{B}\right)$ $\mathbf{B}'_{\parallel} = \mathbf{B}_{\parallel}$ $\mathbf{B}'_{\perp} = \gamma\left(\mathbf{B}_{\perp} - \dfrac{1}{c}\mathbf{v} \times \mathbf{E}\right)$	$\mathbf{E}'_{\parallel} = \mathbf{E}_{\parallel}$ $\mathbf{E}'_{\perp} = \gamma\left(\mathbf{E}_{\perp} + \mathbf{v} \times \mathbf{B}\right)$ $\mathbf{B}'_{\parallel} = \mathbf{B}_{\parallel}$ $\mathbf{B}'_{\perp} = \gamma\left(\mathbf{B}_{\perp} - \dfrac{1}{c^2}\mathbf{v} \times \mathbf{E}\right)$
$\dfrac{1}{4\pi\epsilon_0} = c^2 \times 10^{-7}$ N A^{-2} $= 8.98755\ldots \times 10^9$ m F^{-1} ; $\dfrac{\mu_0}{4\pi} = 10^{-7}$ N A^{-2} ; $c = \dfrac{1}{\sqrt{\mu_0\epsilon_0}} = 2.99792458 \times 10^8$ m s^{-1}		

Impedances (SI units)

ρ = resistivity at room temperature in $10^{-8}\,\Omega$ m:

~ 1.7 for Cu ~ 5.5 for W
~ 2.4 for Au ~ 73 for SS 304
~ 2.8 for Al ~ 100 for Nichrome
(Al alloys may have double the Al value.)

For alternating currents, instantaneous current I, voltage V, angular frequency ω:

$$V = V_0\,e^{j\omega t} = ZI . \tag{1}$$

Impedance of self-inductance L: $Z = j\omega L$.

Impedance of capacitance C: $Z = 1/j\omega C$.

Impedance of free space: $Z = \sqrt{\mu_0/\epsilon_0} = 376.7\ \Omega$.

High-frequency surface impedance of a good conductor:

$$Z = \frac{(1+j)\,\rho}{\delta} , \quad \text{where } \delta = \text{skin depth} ; \tag{2}$$

$$\delta = \sqrt{\frac{\rho}{\pi\nu\mu}} \approx \frac{6.6\ \text{cm}}{\sqrt{\nu\,(\text{Hz})}} \quad \text{for Cu} . \tag{3}$$

Capacitors, inductors, and transmission Lines

The capacitance between two parallel plates of area A spaced by the distance d and enclosing a medium with the dielectric constant ε is

$$C = K\varepsilon A/d, \tag{4}$$

where the correction factor K depends on the extent of the fringing field. If the dielectric fills the capacitor volume without extending beyond the electrodes. the correction factor $K \approx 0.8$ for capacitors of typical geometry.

The inductance at high frequencies of a straight wire whose length ℓ is much greater than the wire diameter d is

$$L \approx 2.0 \left[\frac{\text{nH}}{\text{cm}}\right] \cdot \ell \left(\ln\left(\frac{4\ell}{d}\right) - 1 \right) . \tag{5}$$

For very short wires, representative of vias in a printed circuit board, the inductance is

$$L(\text{in nH}) \approx \ell/d . \tag{6}$$

A transmission line is a pair of conductors with inductance L and capacitance C. The characteristic impedance $Z = \sqrt{L/C}$ and the phase velocity $v_p = 1/\sqrt{LC} = 1/\sqrt{\mu\varepsilon}$, which decreases with the inverse square root of the dielectric constant of the medium. Typical coaxial and ribbon cables have a propagation delay of about 5 ns/cm. The impedance of a coaxial cable with outer diameter D and inner diameter d is

$$Z = 60\,\Omega \cdot \frac{1}{\sqrt{\varepsilon_r}} \ln\frac{D}{d} , \tag{7}$$

where the relative dielectric constant $\varepsilon_r = \varepsilon/\varepsilon_0$. A pair of parallel wires of diameter d and spacing $a > 2.5\,d$ has the impedance

$$Z = 120\,\Omega \cdot \frac{1}{\sqrt{\varepsilon_r}} \ln\frac{2a}{d} . \tag{8}$$

This yields the impedance of a wire at a spacing h above a ground plane,

$$Z = 60\,\Omega \cdot \frac{1}{\sqrt{\varepsilon_r}} \ln\frac{4h}{d} . \tag{9}$$

A common configuration utilizes a thin rectangular conductor above a ground plane with an intermediate dielectric (microstrip). Detailed calculations for this and other transmission line configurations are given by Gunston.

---

* M.A.R. Gunston. Microwave Transmission Line Data, Noble Publishing Corp., Atlanta (1997) ISBN 1-884932-57-6, TK6565.T73G85.

Synchrotron radiation (CGS units)

For a particle of charge e, velocity $v = \beta c$, and energy $E = \gamma mc^2$, traveling in a circular orbit of radius R, the classical energy loss per revolution δE is

$$\delta E = \frac{4\pi}{3}\frac{e^2}{R}\,\beta^3\,\gamma^4 . \tag{10}$$

For high-energy electrons or positrons ($\beta \approx 1$), this becomes

$$\delta E\ (\text{in MeV}) \approx 0.0885\ [E(\text{in GeV})]^4/R(\text{in m}) . \tag{11}$$

For $\gamma \gg 1$, the energy radiated per revolution into the photon energy interval $d(\hbar\omega)$ is

$$dI = \frac{8\pi}{9}\alpha\,\gamma\,F(\omega/\omega_c)\,d(\hbar\omega) , \tag{12}$$

where $\alpha = e^2/\hbar c$ is the fine-structure constant and

$$\omega_c = \frac{3\gamma^3 c}{2R} \tag{13}$$

is the critical frequency. The normalized function $F(y)$ is

$$F(y) = \frac{9}{8\pi}\sqrt{3}\,y \int_y^\infty K_{5/3}(x)\ dx , \tag{14}$$

where $K_{5/3}(x)$ is a modified Bessel function of the third kind. For electrons or positrons,

$$\hbar\omega_c\ (\text{in keV}) \approx 2.22\ [E(\text{in GeV})]^3/R(\text{in m}) . \tag{15}$$

Fig. 1 shows $F(y)$ over the important range of y.
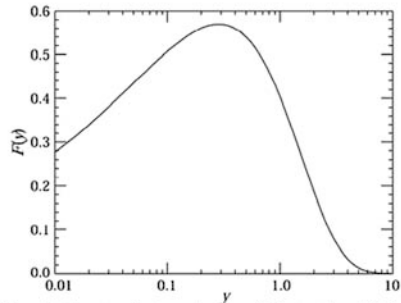
Figure 1: The normalized synchrotron radiation spectrum $F(y)$.

For $\gamma \gg 1$ and $\omega \ll \omega_c$,

$$\frac{dI}{d(\hbar\omega)} \approx 3.3\alpha\,(\omega R/c)^{1/3} , \tag{16}$$

whereas for

$$\gamma \gg 1 \text{ and } \omega \gtrsim 3\omega_c ,$$

$$\frac{dI}{d(\hbar\omega)} \approx \sqrt{\frac{3\pi}{2}}\,\alpha\,\gamma \left(\frac{\omega}{\omega_c}\right)^{1/2} e^{-\omega/\omega_c}\left[1 + \frac{55}{72}\frac{\omega_c}{\omega} + \cdots\right] . \tag{17}$$

The radiation is confined to angles $\lesssim 1/\gamma$ relative to the instantaneous direction of motion. For $\gamma \gg 1$, where Eq. (12) applies, the mean number of photons emitted per revolution is

$$N_\gamma = \frac{5\pi}{\sqrt{3}}\alpha\gamma , \tag{18}$$

and the mean energy per photon is

$$\langle\hbar\omega\rangle = \frac{8}{15\sqrt{3}}\hbar\omega_c . \tag{19}$$

When $\langle\hbar\omega\rangle \gtrsim O(E)$, quantum corrections are important.

Annex 6 : Commonly Used Radioactive Sources

Table 30.1. Revised November 1993 by E. Browne (LBNL).

Nuclide	Half-life	Type of decay	Particle Energy (MeV)	Particle Emission prob.	Photon Energy (MeV)	Photon Emission prob.
$^{22}_{11}$Na	2.603 y	β^+, EC	0.545	90%	0.511	Annih.
					1.275	100%
$^{54}_{25}$Mn	0.855 y	EC			0.835	100%
					Cr K x rays 26%	
$^{55}_{26}$Fe	2.73 y	EC			Mn K x rays:	
					0.00590	24.4%
					0.00649	2.86%
$^{57}_{27}$Co	0.744 y	EC			0.014	9%
					0.122	86%
					0.136	11%
					Fe K x rays 58%	
$^{60}_{27}$Co	5.271 y	β^-	0.316	100%	1.173	100%
					1.333	100%
$^{68}_{32}$Ge	0.742 y	EC			Ga K x rays 44%	
$\rightarrow ^{68}_{31}$Ga		β^+, EC	1.899	90%	0.511	Annih.
					1.077	3%
$^{90}_{38}$Sr	28.5 y	β^-	0.546	100%		
$\rightarrow ^{90}_{39}$Y		β^-	2.283	100%		
$^{106}_{44}$Ru	1.020 y	β^-	0.039	100%		
$\rightarrow ^{106}_{45}$Rh		β^-	3.541	79%	0.512	21%
					0.622	10%
$^{109}_{48}$Cd	1.267 y	EC	0.063 e^-	41%	0.088	3.6%
			0.084 e^-	45%	Ag K x rays 100%	
			0.087 e^-	9%		
$^{113}_{50}$Sn	0.315 y	EC	0.364 e^-	29%	0.392	65%
			0.388 e^-	6%	In K x rays 97%	
$^{137}_{55}$Cs	30.2 y	β^-	0.514	94%	0.662	85%
			1.176	6%		
$^{133}_{56}$Ba	10.54 y	EC	0.045 e^-	50%	0.081	34%
			0.075 e^-	6%	0.356	62%
					Cs K x rays 121%	
$^{207}_{83}$Bi	31.8 y	EC	0.481 e^-	2%	0.569	98%
			0.975 e^-	7%	1.063	75%
			1.047 e^-	2%	1.770	7%
					Pb K x rays 78%	
$^{228}_{90}$Th	1.912 y	6α:	5.341 to 8.785		0.239	44%
		$3\beta^-$:	0.334 to 2.246		0.583	31%
					2.614	36%
($\rightarrow ^{224}_{88}$Ra	$\rightarrow ^{220}_{86}$Rn	$\rightarrow ^{216}_{84}$Po	$\rightarrow ^{212}_{82}$Pb	$\rightarrow ^{212}_{83}$Bi	$\rightarrow ^{212}_{84}$Po)	
$^{241}_{95}$Am	432.7 y	α	5.443	13%	0.060	36%
			5.486	85%	Np L x rays 38%	
$^{241}_{95}$Am/Be	432.2 y	6×10^{-5} neutrons (4–8 MeV) and				
		$4 \times 10^{-5}\gamma$'s (4.43 MeV) per Am decay				
$^{244}_{96}$Cm	18.11 y	α	5.763	24%	Pu L x rays $\sim$ 9%	
			5.805	76%		
$^{252}_{98}$Cf	2.645 y	α (97%)	6.076	15%		
			6.118	82%		
		Fission (3.1%)				
		$\approx$ 20 γ's/fission; 80% < 1 MeV				
		$\approx$ 4 neutrons/fission; $\langle E_n \rangle$ = 2.14 MeV				

"Emission probability" is the probability per decay of a given emission; because of cascades these may total more than 100%. Only principal emissions are listed. EC means electron capture, and e^- means monoenergetic internal conversion (Auger) electron. The intensity of 0.511 MeV e^+e^- annihilation photons depends upon the number of stopped positrons. Endpoint $\beta^\pm$ energies are listed. In some cases when energies are closely spaced, the γ-ray values are approximate weighted averages. Radiation from short-lived daughter isotopes is included where relevant.

Half-lives, energies, and intensities are from E. Browne and R.B. Firestone, *Table of Radioactive Isotopes* (John Wiley & Sons, New York, 1986), recent *Nuclear Data Sheets*, and *X-ray and Gamma-ray Standards for Detector Calibration*, IAEA-TECDOC-619 (1991).

Neutron data are from *Neutron Sources for Basic Physics and Applications* (Pergamon Press, 1983).

Index