Lucio Cerrito

Radiation and Detectors

Introduction to the Physics of Radiation and Detection Devices

Springer

Lucio Cerrito
School of Physics and Astronomy
Queen Mary University of London
London
UK

ISSN 1868-4513 ISSN 1868-4521 (electronic)
Graduate Texts in Physics
ISBN 978-3-319-53179-3 ISBN 978-3-319-53181-6 (eBook)
DOI 10.1007/978-3-319-53181-6

Library of Congress Control Number: 2017938530

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This textbook is based on my lectures for the undergraduate course of radiation detectors, in the School of Physics and Astronomy at Queen Mary University of London, given in the years 2012 and 2013.

The book provides an introduction to the field of radiation, to the mechanisms of interaction between radiation and matter, and to the exploitation of those mechanisms for the purpose of designing radiation detectors. Significant attention is paid to all of those aspects in equal measure, defining and explaining the language and the first principles with little assumption. The mathematical formalism is kept to a minimum, but simple derivations are presented in order to guide the reasoning and understanding of the physics principles. Detectors are introduced by both their general working design and by modern examples of devices currently operating around the world. There are over 140 original experimental figures, detector schematics, and photographs in this book, which help relate the material to the broader research context and can be used to find out more through the selected referenced documents.

Since interaction processes become more or less relevant depending on the energy, on the materials and the detection objectives, in this book the interaction mechanisms and the detectors are presented interleaved, and the description of a technology always follows the particular interaction mechanism which it exploits. For example, in describing the interaction of photons, the photoelectric effect is introduced first in Chap. 9 by only the details needed in the context of photomultipliers, then expanded in Chap. 10 together with Compton scattering and pair production which anticipate the discussion of electromagnetic showers and the technology of calorimeters. In other words, in organising the material in this book only the physics principles most relevant are presented first, clearly and extensively, just before describing a class of detectors.

The first five chapters are dedicated to *radiation*, as we understand it today based on subatomic physics, including the language in use, its metrics, and the most common natural and man-made forms of radiation. Realistic worked examples of the various types of radiation and its energy accompany the presentation. Dosimetry (Chap. 3) is presented from a modern, user-led point of view, and

relativistic kinematics (Chap. 4) is introduced to give the basic knowledge needed to handle the more formal aspects of radiation dynamics and interaction. Part II is dedicated to the *interaction* between radiation and matter, and to *detectors*. The energy loss by ionisation (Chap. 6) is described in some detail, anticipating the principles of ionisation detectors (Chap. 7), semiconductor detectors (Chap. 8), and scintillation detectors (Chap. 9). The topics in Chap. 10 span several interaction mechanisms that underpin the phenomenology of showers and the design of calorimeters. Chapter 11 covers a number of additional phenomena including Cherenkov and transition radiation and the detection of neutrinos. Finally, a summary of statistics and probability distributions is presented in the Appendix.

Through this book, the reader is expected to acquire an awareness of how radiation and its exploitation are rapidly expanding to many diverse contexts in the modern world, including in medical physics. The reader will also acquire the preliminary broad knowledge needed if wishing to undertake advanced studies in one of the areas presented. The introductory level of the material makes the book of particular interest to undergraduate students, graduate students (for an introduction to radiation detectors or selected aspects of it), and lecturers as support for a one-semester undergraduate course on radiation detectors.

I am indebted to Anna Di Ciaccio at the Department of Physics of Rome Tor Vergata, and Jonathan Hays, Teppei Katori, and Jeanne Wilson at the School of Physics of Queen Mary University of London, for providing valuable feedback on earlier drafts of some of the chapters. For their support and feedback, I am grateful to the editorial team at Springer, particularly to Tom Spicer.

Rome, Italy — Lucio Cerrito

Contents

Acronyms

AC	Alternating Current
ACE	Advanced Composition Explorer
ADC	Analog-to-Digital Converter
ALICE	A Large Ion Collider Experiment
AMANDA	Antarctic Muon and Neutrino Detector Array
ANTARES	Astronomy with a Neutrino Telescope and Abyss Environmental Research
APEX	Atacama Pathfinder Experiment
ATLAS	A Toroidal LHC Apparatus
CDF	Collider Detector at Fermilab
CERN	European Organisation for Nuclear Research
CMB	Cosmic Microwave Background
CMS	Compact Muon Solenoid
CNAO	Italian National Centre of Oncological Hadron-therapy
COBE	Cosmic Background Explorer
CT	Computed Tomography
DC	Direct Current
DESY	Deutsches Elektronen-Synchrotron
D0	D0 Experiment
EM	Electromagnetic (Calorimeter)
ESA	European Space Agency
FIRAS	Far Infrared Absolute Spectrophotometer
FNAL	Fermi National Accelerator Laboratory
GEM	Gas Electron Multiplier
ICRP	International Commission on Radiological Protection
ICRU	International Commission on Radiation Units and Measurements
LABOCA	Large APEX Bolometer Camera
LAr	ATLAS Liquid Argon Calorimeter
LHC	Large Hadron Collider

LHCb	Large Hadron Collider beauty
LINAC	Linear Accelerator
LNT	Linear-non-Threshold
LO	Leading Order
LOFT	Large Observatory for X-ray Timing
MICROMEGAS	Micromesh Gaseous Structure
MSGC	Multistrip Gas Chamber
MWPC	Multiwire Proportional Chamber
NASA	National and Aeronautics Space Administration
NLO	Next-to-Leading Order
PDF	Probability Distribution Function
PID	Particle Identification
PMT	Photomultiplier
PVT	Polyvinyl Toluene
QCD	Quantum Chromodynamics
RHIC	Relativistic Heavy Ion Collider
RICH	Ring Imaging Cherenkov
RF	Radio Frequency
RMS	Root Mean Square
RPC	Resistive Plate Chamber
SI	International System of Units
SM	Standard Model
SNO	Sudbury Neutrino Observatory
SNU	Solar Neutrino Unit
SWOOPS	Solar Wind Observations Over the Poles of the Sun
SWICS	Solar Wind Ion Composition Spectrometer
TPC	Time Projection Chamber
TRD	Transition Radiation Detector (ALICE)
TRT	Transition Radiation Tracker (ATLAS)
UV	Ultraviolet
WMAP	Wilkinson Microwave Anisotropy Probe

Part I
Radiation

Part I of this book is dedicated to *radiation*, its features and the terminology in use. The objective is to present what is meant by radiation, the different types of radiation, how it can be measured, and what are some common natural and artificial sources. Special topics include dosimetry, the physics and history of accelerator machines, and elements of relativistic kinematics and collisions.

Chapter 1
Particles and Forces

Abstract Radiation consists, ultimately, of subatomic matter and non-matter particles (e.g. photons) and nuclei. The mechanisms by which radiation impacts the surrounding world is determined by the laws that govern subatomic particle interactions. This chapter presents the elementary and composite particles discovered so far and introduces the forces by which these particles interact. The emphasis is placed on an overview of the current knowledge in this field, rather than an historical or theoretical account of the topic.

1.1 Units of Energy and Mass

The unit traditionally in use for the measurement of particle and radiation energy is the *electronvolt* (eV). It is defined as the kinetic energy gained by an electron after accelerating through a potential difference of 1 volt. Amounts of energy expressed in electronvolts are often found with prefixes kilo- (keV), mega- (MeV), giga- (GeV), tera- (TeV):

$$1\ \text{keV} = 10^3\ \text{eV}; \tag{1.1}$$
$$1\ \text{MeV} = 10^6\ \text{eV}; \tag{1.2}$$
$$1\ \text{GeV} = 10^9\ \text{eV}; \tag{1.3}$$
$$1\ \text{TeV} = 10^{12}\ \text{eV}. \tag{1.4}$$

In SI units,[1] the electronvolt corresponds to an extremely small amount of energy, tribute of the fact that the unit is practical to describe energies proper of the atomic and subatomic world. In fact, by its definition the electronvolt corresponds to 1 volt (1 joule per coulomb) multiplied by the electron charge ($1.602\,176\,565(35) \times 10^{-19}$ C) and is therefore:

$$1\,\text{eV} = 1.602\,176\,565(35) \times 10^{-19}\,\text{J}. \tag{1.5}$$

[1] SI is the International System of Units, in which the energy is measured in joule (J).

L. Cerrito, *Radiation and Detectors*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-53181-6_1

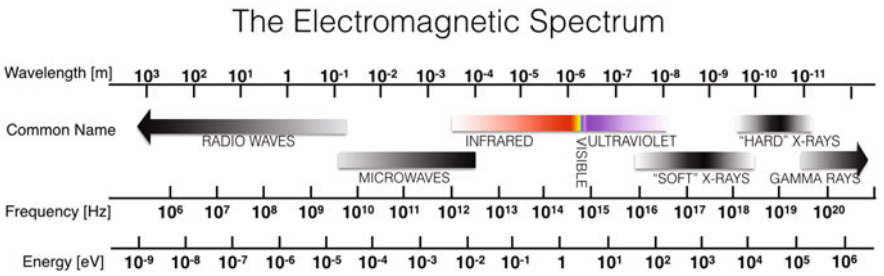


Fig. 1.1 The electromagnetic spectrum with the name of common types of radiation

Two fundamental constants are important in treating elementary particles. The Planck constant (h) relates in quantum mechanics the energy of a photon with its frequency by the relationship:

$$E = h\nu = hc/\lambda \text{ (or also } E = \hbar\omega), \tag{1.6}$$

where ν, ω and λ are the frequency, angular frequency, and wavelength of the photon, respectively, and c is the speed of light. Figure 1.1 shows the electromagnetic spectrum indicating the wavelength, frequency and energy (in eV) of common types of electromagnetic radiation. The human eye is sensitive to electromagnetic radiation with wavelengths between approximately 390 and 700 nm.

The symbol $\hbar$ is the *reduced* Planck constant, given by $\hbar \equiv h/2\pi$. The Planck constant has the dimension of action (or angular momentum), that is energy multiplied by time or momentum multiplied by distance and is an extremely small number when expressed in units of both joule or electronvolt:

$$h = 6.626\,069\,57(29) \times 10^{-34}\,\text{J s} = 4.135\,667\,516(91) \times 10^{-15}\,\text{eV s} \tag{1.7}$$

and therefore $\hbar \simeq 1 \times 10^{-34}\text{J s} = 6.6 \times 10^{-16}\,\text{eV s}$. Another fundamental constant is the speed of light, which in vacuum is:

$$c = 299\,792\,458\ \text{m s}^{-1} \simeq 3 \times 10^{8}\,\text{m s}^{-1}. \tag{1.8}$$

The constants c and h are so common in equations of particle dynamics that it has become customary to define conversion factors such that the speed of light and the reduced Plank constant are equal to 1 ($c = \hbar = 1$, *natural units*). To do this, the length is defined as the distance travelled by light in vacuum in one unit of time, so $c = 1$ and length and time are measured in the same units.

Furthermore, since in relativistic kinematics (see Chap. 4) the relativistic mass of a particle is related to its energy by:

$$E = m\,c^2, \tag{1.9}$$

this implies that masses are measured in units of eV/c^2. Similarly, the *momentum* $\mathbf{p}$ of a particle is related to its relativistic mass and velocity $\mathbf{v}$ by:

$$\mathbf{p} = m\,\mathbf{v}, \tag{1.10}$$

and therefore momentum is measured in units of eV/c. In natural units because $c = 1$ then masses, momenta, and energies (see (1.9) and (1.10)) are all expressed in the same units (eV). With $\hbar = 1$ and $c = 1$, the (1.6) shows that inverse lengths (L) and inverse times (T) are also measured in units of energies (E), since:

$$\left[\frac{1}{L}\right] = \left[\frac{E}{hc}\right] \quad \text{and} \quad \left[\frac{1}{T}\right] = \left[\frac{E}{h}\right] \tag{1.11}$$

Natural units are used throughout this textbook, however numerical results can be restored with the use of the numerical values of $\hbar$, c, and of the conversion factor:

$$\hbar\, c = 1.97 \times 10^{-16}\ \mathrm{GeV\,m}. \tag{1.12}$$

1.2 Elementary Particles and Antiparticles

The quest to understand and categorise the constituents of everyday matter, including the building blocks of our body, the objects surrounding us, the planets and stars in the Universe, has driven the field of experimental physics towards the examination of progressively smaller distances. For instance, structures called molecules are visible in substances at a scale of 10^{-8} to 10^{-10} m, and individual atoms appear on a scale of 10^{-10} to 10^{-11} m (the distance 10^{-10} m is called one ångström, 1 Å). Atoms are formed, in a simplified picture, by a nucleus of protons and neutrons combined with a shell of electrons (e). The nucleus appears to have a size of 10^{-14}–10^{-15} m (the distance 10^{-15} m is called one fermi, 1 fm) and its constituents, the proton and the neutron (see Fig. 1.2) are detectable at a distance of 10^{-15} m. Notice that there is a difference of about three to four orders of magnitude between the size of the atom and the size of its nucleus, so that the majority of the space within an atom is actually empty space. The energy corresponding to the characteristic size of an atom can be calculated from (1.11):

$$\mathscr{E}_{atom} = \frac{\hbar\, c}{1\text{Å}} = 1.97 \times 10^{-16}\ \mathrm{GeV\,m} \times 10^{10}\ \mathrm{m^{-1}} \approx \mathrm{keV} \tag{1.13}$$

and for the size of the nucleus, 10^{-14} m, that energy is $\mathscr{E}_{nucleus} \approx 10$ MeV.

When probing distances smaller than 10^{-18} m, the proton and the neutron reveal a substructure of quarks. Two quark types, called the *up* quark (u) and the *down* quark (d), are the dominant constituents of the proton and neutron. The energy corresponding to a scale of 10^{-18} m (the dimensional scale of quarks) is $\mathscr{E}_{quark} \approx$ 100 GeV.

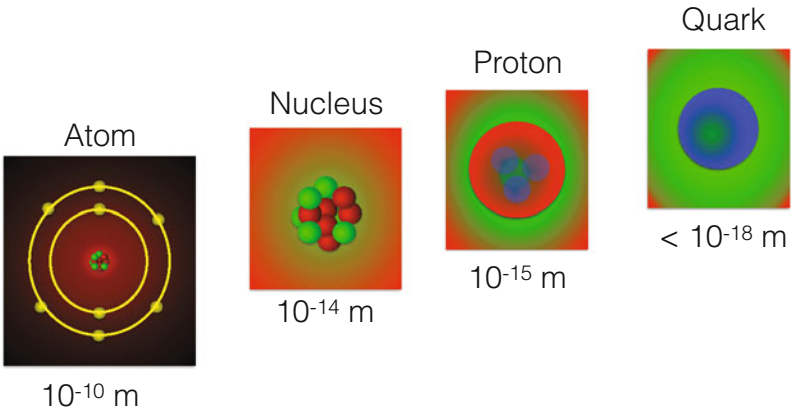


Fig. 1.2 The building blocks of the Universe and all everyday objects seem to consist of only quarks and electrons. On a scale of 10^{-10} m they combine to form atoms

Modern particle accelerators, such as the CERN Large Hadron Collider, are capable of probing distances of the order of up to about 10^{-21} m and at this distance there is no visible substructure to either the electron or the quarks. Quarks and electrons are therefore, as of today, the fundamental building blocks (meaning the matter particles) of the atoms.

Besides the electron and the up and down quarks mentioned above, there are additional fundamental particles that have been observed. They appear to be produced readily as a result of particle collisions as well as in particle decays. The elementary particles discovered so far and forming the matter of the Universe are summarised in Table 1.1.

The fundamental particles are classed into three families of two leptons and two quarks, and each higher-ranked family contains particles identical to the lower family except for a larger mass. The first family of leptons consists of the electron and the electron neutrino. The *muon* (μ) and the *tau* (τ) leptons share the same charge and spin of the electron but are heavier; the heavier mass changes the phenomenology of these particles, e.g. it influences their decays and lifetime compared to the electron which is stable. The muon has a mean lifetime of 2.2×10^{-6} s before decaying to an electron and two neutrinos (one of electron type and one of muon type), whereas the τ has much shorter mean lifetime of 2.9×10^{-13} s and decays primarily to hadrons (see Sect. 1.5). The different lifetime is explained by the large difference of mass between the μ and the τ.

Similarly to the leptons, the first family of quarks (the one containing the up and down quarks) is replicated in two additional families of unstable heavier quarks: *charm (c), strange (s), top (t)* and *bottom (b)*. They all have fractional electric charge of $+2/3$ or $-1/3$ in units of the electron charge. The top quark, in the third family,

Table 1.1 Elementary particles (fermions)

Particle	Family	Class	Symbol	Spin	Electric charge[a]	Mass[b]	Mean lifetime[c] (s)
Electron neutrino	I	Lepton	ν_e	1/2	0	<2 eV	Stable
Electron	I	Lepton	e^-	1/2	-1	0.5 MeV	Stable
Up	I	Quark	u	1/2	+2/3	2 MeV	Hadronise
Down	I	Quark	d	1/2	$-1/3$	5 MeV	Hadronise
Muon neutrino	II	Lepton	ν_μ	1/2	0	<0.19 MeV	Stable
Muon	II	Lepton	μ^-	1/2	-1	106 MeV	2.2×10^{-6}
Charm	II	Quark	c	1/2	+2/3	1.3 GeV	Hadronise
Strange	II	Quark	s	1/2	$-1/3$	100 MeV	Hadronise
Tau neutrino	III	Lepton	ν_τ	1/2	0	<18.2 MeV	Stable
Tau	III	Lepton	τ^-	1/2	-1	1.8 GeV	2.9×10^{-13}
Top	III	Quark	t	1/2	+2/3	173 GeV	5×10^{-25}
Bottom	III	Quark	b	1/2	$-1/3$	5 GeV	Hadronise

[a]Electric charge is expressed in units of the electron charge

[b]Mass values shown are only approximate and measurements can be found in [1]. Except for the top quark, the mass of the quarks are estimates based on the properties of quark bound states (see Sect. 1.5)

[c]Mean lifetime values are only approximate and measurements can be found in [1]. Quarks that form bound states of hadrons might decay as part of an hadronic state (see Sect. 1.5)

is the heaviest fundamental particle ever discovered and its mass[2] of 173 GeV is similar to the mass of the entire nucleus in an atom of gold. As a reference, the mass of the proton is only about 1 GeV (approximately 938 MeV). The leptons and quarks have spin 1/2,[3] and in quantum mechanics particles with spin 1/2 obey Fermi's spin-statistics rule, i.e. they cannot occupy simultaneously the same quantum state; they are called *fermions*. For each fermion there is a corresponding antiparticle of the same mass but opposite electric charge, and collectively antiparticles are referred to as *antimatter*.

Antimatter was introduced in 1928 by Dirac [2] as a theoretical consequence of combining quantum mechanics and special relativity, and it was discovered in 1932 [3] through the detection of a positively charged particle with the same mass as the electron. The antielectron (e^+) is more commonly called a positron and is routinely produced through the decay of heavier particles. For every fermion in Table 1.1 the corresponding antiparticle is indicated by placing a dash above the particle symbol, so that for example $\bar{\nu}_e$ indicates an antineutrino and $\bar{t}$ an antitop, except for the

[2]The mass of particles is expressed in units of energy when using *natural* units, a full explanation is given in Sect. 1.1.

[3]Spin is a form of angular momentum possessed by elementary particles.

charged leptons for which the antiparticle is indicated by the positive charge sign as: e^+, μ^+ and τ^+.

Particle decays are indicated by the notation:

$$initial\ particle \rightarrow final\ particles, \tag{1.14}$$

and the same particle might decay into a number of different modes, called decay channels, each with its own probability. Considering all the correct symbols for particles and antiparticles, the decay of a muon to an electron and two neutrinos is therefore written as:

$$\mu^- \rightarrow e^-\ \bar{\nu}_e\ \nu_\mu. \tag{1.15}$$

It is worth mentioning that there is so far no experimental evidence to confirm that the antineutrino is indeed a different particle than the neutrino. For the positron and the electron for instance, it is sufficient to detect two particles of the same mass but with opposite sign electric charge. For a neutral, nearly massless, and weakly interacting particle the proof is vastly more difficult.

1.3 Fundamental Forces and the Higgs Boson

The elementary particles introduced in Sect. 1.2 interact with each other via one or more fundamental forces. These are the four known fundamental forces of nature: electromagnetic, weak nuclear, strong nuclear, and gravitational. The gravitational force, whilst present, is negligible at the level of subatomic particles when compared to the other three forces, and is therefore neglected in all considerations of particle dynamics. The electromagnetic and the weak nuclear force are conceptually unified in a mathematical description called electroweak interactions. This includes the electric and magnetic forces seen in the macroscopic world, as well as the weak nuclear force which was first observed in neutron decays. All quarks and leptons are subject to the electroweak interactions. Neutrinos are electrically neutral and therefore interact only via the weak force. The strong nuclear force is only felt by quarks.

The relative strength of these forces depends significantly on the distance at which they are probed. At distances of 10^{-18} m, i.e. the quark scale, if the strong interaction is taken as reference (with a strength of 1), the electromagnetic and weak forces have a similar strength of 10^{-2} and the gravitational force is immensely weaker, with a strength of 10^{-43}. The weak force however becomes much more feeble at a marginally larger distance, at 10^{-17} m it becomes about 10,000 times weaker than at 10^{-18} m.

Similarly to the electric charge, the strong nuclear force has an associated *colour* charge. This charge has of course nothing to do with the colours of visible light, but the word is aptly used to label the three types of charges present: red, green and blue. Furthermore, each of them can have two states, referred to as antired, antigreen and antiblue.

Table 1.2 Elementary particles (bosons)

Particle	Symbol	Interaction	Spin	Electric charge	Mass[a] (GeV)	Width (GeV)	Dominant decay mode
Photon	γ	Electromagnetic	1	0	0	Stable	–
W boson	$W^{\pm}$	Weak	1	± 1	80	2.1	$q\bar{q}'$[b], $\ell\nu_\ell$[c]
Z boson	Z^0	Weak	1	0	91	2.5	$q\bar{q}$, $\ell^+\ell^-$, $\nu_\ell\bar{\nu}_\ell$
Gluon	g	Strong	1	0	0	Stable	–
Higgs	H^0	–	1	0	126	$\sim 4 \times 10^{-3}$[d]	$b\bar{b}$, $W^{\pm}W^{\mp *}$

[a] Values for the $W^{\pm}$, Z^0, H^0 bosons are only approximate and measurements can be found in [1]
[b] q indicates any quark except the top
[c] ℓ indicates each type of lepton, e, μ and τ
[d] Prediction based on the Standard Model

The quantum-mechanical relativistic theory of the electroweak and strong interactions between elementary particles is called the *Standard Model*.

The Standard Model describes the particle interactions through the exchange of spin 1 particle carriers of the electromagnetic, weak and strong force. The photon (γ) is the carrier of the electromagnetic force, while the weak force is carried by the $W^{\pm}$ and Z^0 bosons. Finally, the strong force is carried by eight gluons (g). Table 1.2 summarises the fundamental forces and their particle carriers. The Z and W bosons are readily produced in high-energy particle collisions. The W boson is also produced in the decay of the top quark:

$$t \rightarrow \; W^+ \, b, \tag{1.16}$$

and the W itself then proceeds to decay into lighter particles, either to one lepton (e, μ or τ) with the corresponding neutrino, or to a pair of quark and anti-quark (all types except for the top quark, whose mass is too large).

The picture of the building blocks of the Universe is completed by an additional particle: the *Higgs* boson (H^0) [5–7]. The Higgs boson was discovered in 2012 by the ATLAS and CMS experiments [8, 9] at CERN's the Large Hadron Collider. The new particle was seen to have a mass of 126 GeV, and is key to explaining how leptons, quarks and the electroweak bosons acquire their own mass. In the Standard Model, the Higgs boson is the manifestation of the Higgs field and the quarks and leptons acquire their mass through their interaction with such a field. Different masses of leptons and quarks are explained by the different strengths of their coupling with the Higgs field.

Once produced, the Higgs boson decays to ordinary particles in a variety of decay channels. The event display in Fig. 1.3 shows the decay of an Higgs boson candidate as seen by the ATLAS detector to a pair of Z and finally to four electrons

$$H \rightarrow \; Z^0 Z^{0*} \rightarrow e^+ \, e^- \, e^+ \, e^-. \tag{1.17}$$

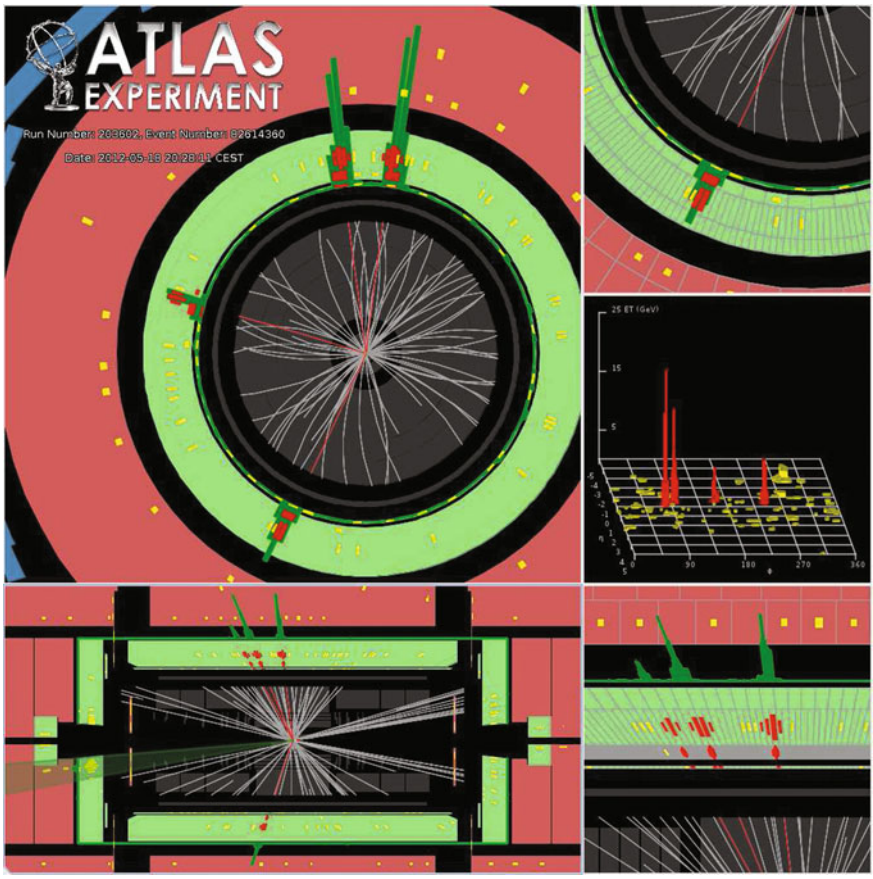


Fig. 1.3 The decay of a Higgs boson to four electrons as detected by the ATLAS experiment. The electrons are visible as *straight* lines in the inner tracking system and large *energy deposits* in the electromagnetic calorimeter. Reprinted with permission from [4]

The quadrant in the upper left shows a cross sectional view of the detector. The beams of protons enter the detector perpendicularly to the plane of the page and collide at the center of the concentric detector layers. Charged particles emerging from the collision are bent by a magnetic field and perform a helicoidal trajectory with radius inversely proportional to the component of their momentum perpendicular to the magnetic field. The four electrons are visible as almost-straight lines in the inner tracking system, and also as large *energy deposits* in the electromagnetic calorimeter. Although the dominant decay modes of the Higgs boson are $H^0 \to b\bar{b}$ and $H^0 \to W^{\pm}W^{\mp *}$, the Higgs boson is more easily distinguished from backgrounds in its rarer $H \to Z^0 Z^{0*}$ and $H^0 \to \gamma\gamma$ final states.

1.4 Feynman Diagrams

The interaction between particles is described pictorially with the aid of *Feynman diagrams*. The sketch in Fig. 1.4 depicts the electromagnetic scattering between two electrons: $e^- + e^- \rightarrow e^- + e^-$ through the exchange of a single photon or a Z^0 boson. A line entering a diagram from the left-hand side is interpreted as a particle in the initial state, and one leaving a diagram on the right-hand side as one present in the final state. The electrons are represented by a solid line, while the photon, W and Z^0 by a wiggly line and the gluon by a curly line (see Fig. 1.5). By convention time flows from left to right, while the arrows point in the direction of motion for particles and in the opposite direction for antiparticles, so that charge conservation is implied at each vertex. In Fig. 1.4, both electrons can emit the exchanged photon.

The diagram in Fig. 1.6 represents one contribution to the so-called *Compton scattering*. The electron line at the center indicates both the initial electron emitting the final photon followed by absorption of the initial photon, or creation of an electron-positron pair by the initial photon followed by electron-positron annihilation of the initial electron.

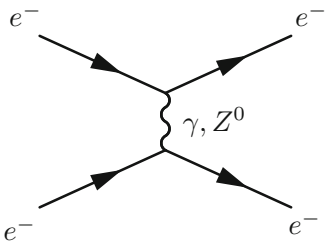


Fig. 1.4 Feynman diagram of an electron–electron scattering process characterised by the exchange of a photon or a Z^0 boson

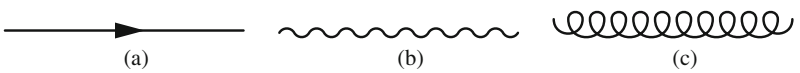
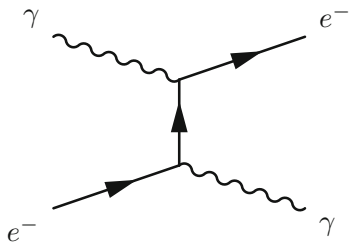


Fig. 1.5 Feynman diagrams of a fermion (**a**), photon or electroweak boson (**b**), and gluon (**c**)

Fig. 1.6 Feynman diagram of Compton scattering, $\gamma\ e \rightarrow \gamma\ e$



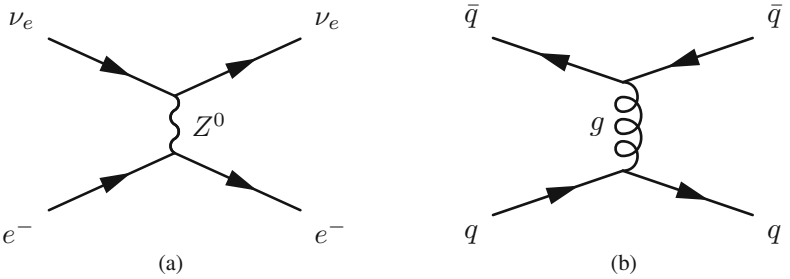
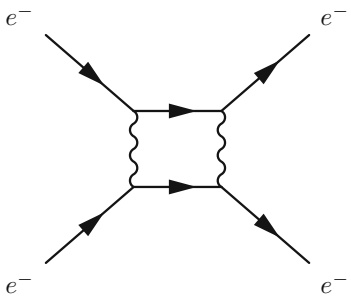


Fig. 1.7 Feynman diagrams showing the exchange of bosons in the interaction of leptons and quarks. **a** neutrino–electron scattering mediated by the exchange of a Z^0 boson. **b** quark–quark scattering mediated by a gluon exchange

Fig. 1.8 Contribution to electron–electron scattering from two-photon exchange



The illustrations in Fig. 1.7 show examples of the $\nu_e + e \rightarrow \nu_e + e$ scattering, mediated by a Z^0, and the gluon exchange process $q + \bar{q} \rightarrow q + \bar{q}$.

The interaction processes can also occur with the exchange of multiple particles, as shown by the two-photon exchange depicted in Fig. 1.8. However, since each vertex in the diagram represents a basic interaction, whose probability is dependent on the strength of the coupling between the particles involved, multiple vertices typically correspond to progressively smaller contributions to the same initial and final-state process. The number of vertices in a given Feynman diagram is called its *order* and the diagram with the least number of vertices for a given process is called a *Leading Order (LO)* diagram.

1.5 Hadrons

The quarks introduced in Sect. 1.2 form bound states (called *hadrons*) composed of either three quarks (qqq or $\bar{q}\bar{q}\bar{q}$, named *baryons*) or a pair of quark–antiquark ($q\bar{q}$, indicated as *mesons*). The quark bound states are the result of the strong nuclear force acting between the quarks. These configurations are the only ones that allow colourless bound states, either by pairing a red–green–blue combination of three

quarks, or a colour–anticolour state of two quarks. All searches for free quarks have so far been unsuccessful, quarks have not been found free from hadrons, and our modeling of the strong nuclear force (a theory called *quantum chromodynamics*, QCD) incorporates the fact that they cannot do so. By pulling apart quarks, as it is for instance common in hadron colliders following the breakup of protons, the quarks will form bound states in a number that increases with the energy of the quark. Escaping quarks will therefore appear as a shower of hadrons moving in the direction of the original quark.

The event display in Fig. 1.9 shows a signature compatible with three high-energy quarks or gluons being emitted from the collision point of a proton–antiproton collision. The active detector is shown in cross sectional view, the colliding beams approaching from opposite directions perpendicular to the plane of the page, with the inner circle representing the tracking volume of the detector and the outer ring indicating the calorimeters. Charged particles appear as tracks in the inner detector, bent by a magnetic field, and coloured blocks in the calorimeters indicate an energy release. Because the quarks hadronise as they move away from the collision point at the center of the picture, they would appear as three jets of collimated particles.

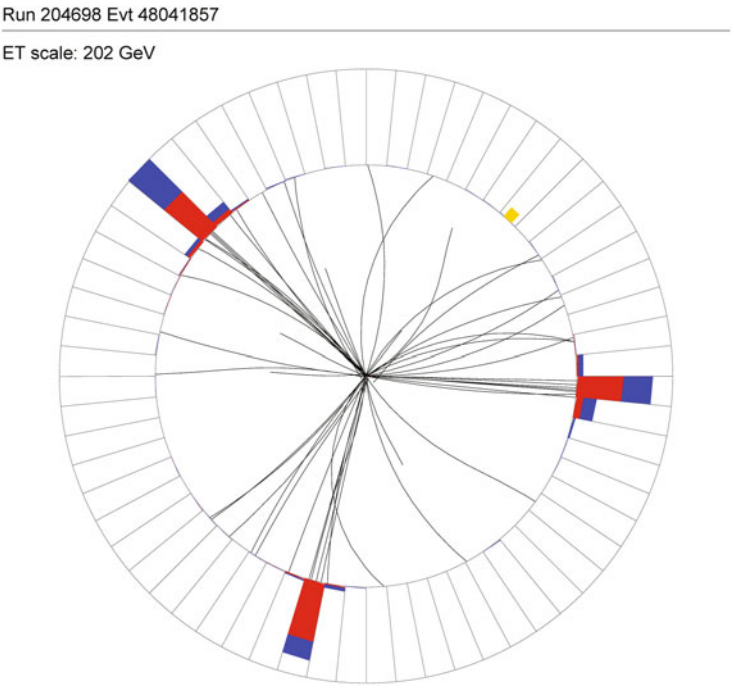


Fig. 1.9 Signature compatible with the hadronisation of three quarks or gluons emitted at high energy from the collision point (center of display) in this proton–antiproton collision, recorded by the D0 detector at the Tevatron collider. Printed with permission from [10]. Credit: Fermilab/DZero Collaboration

Table 1.3 The quark content of a selection of baryons (lowest levels)

Particle	Symbol	Quark content	Electric charge[a]	Mass[b] (GeV)	Mean lifetime[b] (s)	Dominant decay mode
Proton	p	uud	$+1$	1	Stable[c]	–
Neutron	n	udd	0	1	880[d]	$pe^-\bar{\nu}_e$
Lambda	Λ	uds	0	1.1	2×10^{-10}	$p\pi^-$
Xi	Ξ^0	uss	0	1.3	3×10^{-10}	$\Lambda\pi^0$
Omega	Ω^-	sss	-1	1.7	8×10^{-11}	ΛK^0
Lambda (c)	Λ_c^+	udc	$+1$	2.3	2×10^{-13}	p anything
Lambda (b)	Λ_b^0	udb	0	5.6	1×10^{-12}	Λ_c^+ anything

[a]Electric charge is expressed in units of the proton charge
[b]Mass and lifetime values shown are only approximate and measurements can be found in [1]
[c]Current measurements of the proton lifetime yield a limit of $>10^{29}$ years
[d]This is the mean lifetime of a free neutron, i.e. not bound in a nucleus

Table 1.4 The quark content of a selection of mesons (lowest levels)

Particle	Symbol	Quark content	Electric charge[a]	Mass[b] (GeV)	Mean Lifetime[c] (s)	Dominant decay mode
Pion ($\pm$)	π^+, π^-	$u\bar{d}, d\bar{u}$	$+1, -1$	0.14	3×10^{-8}	$\mu\bar{\nu}_\mu, \mu^+\nu_\mu$
Pion (0)	π^0	$u\bar{u}, d\bar{d}$	0	0.14	8×10^{-17}	$\gamma\gamma$
Kaon ($\pm$)	K^+, K^-	$u\bar{s}, s\bar{u}$	$+1, -1$	0.49	1×10^{-8}	$\mu\bar{\nu}_\mu, \mu^+\nu_\mu$, $\pi^\pm\pi^0$
D ($\pm$)	D^+, D^-	$c\bar{d}, d\bar{c}$	$+1, -1$	1.9	1×10^{-12}	$K^0\bar{K}^0$ anything
B (0)	$B^0, \bar{B}^0$	$d\bar{b}, b\bar{d}$	0	5.3	1×10^{-12}	$\bar{D}^0X, D^-X$
J/Ψ(1S)	$J/\Psi(1S)$	$c\bar{c}$	0	3.1	93 keV	hadrons, e^+e^-, $\mu^+\mu^-$
$\Upsilon(1S)$	$\Upsilon(1S)$	$b\bar{b}$	0	9.5	54 keV	$\ell^+\ell^-$ [b]

[a]Electric charge is expressed in units of the proton charge
[b]Mass and lifetime values shown are only approximate and measurements can be found in [1]
[c]ℓ indicates each type of lepton e, μ and τ

The event was recorded by the D0 detector at the US Fermilab's Tevatron Collider in 2005.

Conversely, the top is the only quark that does not form bound states, but only because it decays (almost always to a W boson and b quark) with a mean lifetime of $\sim 5 \times 10^{-25}$ s, which is about ten times faster than the time required to *hadronise*, i.e. to form the bound states of quarks.

The proton is composed primarily of uud quarks, and the neutron is a bound state of udd quarks. Historically, hadrons were discovered in increasing numbers and by 1955, the proton, the neutron and three pions (π^0, π^+ and π^-) were known, along with the antiproton and antineutron.

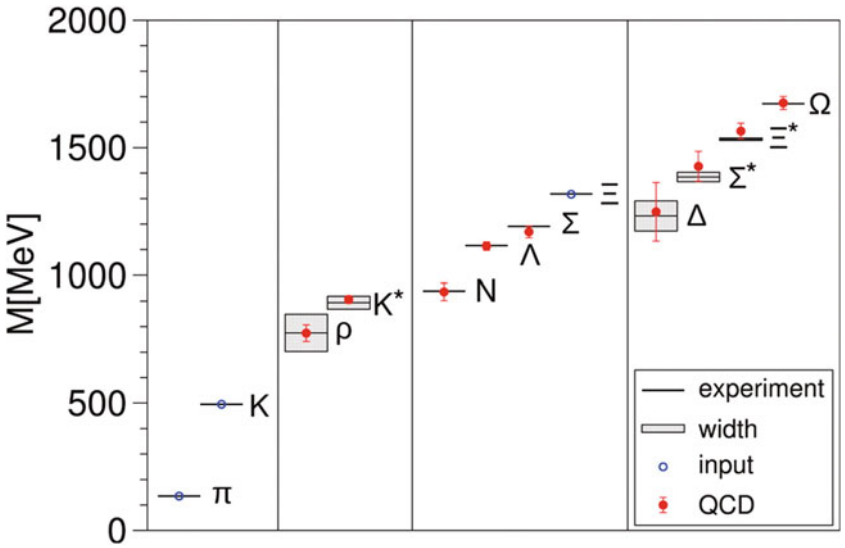


Fig. 1.10 The light hadron spectrum of QCD, from [14]. Reprinted with permission from AAAS

Quarks were subsequently postulated by Gell-Mann [11] and Zweig [12, 13] in 1964 to explain patterns of mass, charge and decay properties that characterised the observed hadrons. The quark content of several of the lowest level hadrons is summarised in Tables 1.3 and 1.4. All hadrons, except for the proton, are unstable and the neutron decays into a proton with a mean lifetime of 880 s [1].

The plot in Fig. 1.10 summarises the mass spectrum of light hadrons from the pion to the Ω. The lightest, pions and kaons, are the most commonly produced.

1.6 Lepton and Quark Numbers

Particle decays have been observed to occur under strict conservation rules. The conservation of energy and momentum are concepts of classical physics that are also valid in relativistic quantum mechanics. The conservation of the electric charge is also respected by the electroweak and strong interactions. Additional conserved quantities which are observed include for instance the *lepton number*, defined as:

$$L_\ell = N(\ell^-) + N(\nu_\ell) - N(\ell^+) - N(\bar{\nu}_\ell), \tag{1.18}$$

where ℓ is separately either e, μ or τ. Lepton number is conserved in decays mediated by electroweak interactions, so for instance the initial state in the muon decay process $\mu \to e\ \bar{\nu}_e\ \nu_\mu$ has muon number 1 and electron and tau numbers zero, and therefore

the final state must also have muon number 1 (given by the ν_μ) and electron and tau numbers zero (the electron number is cancelled by the e and the $\bar{\nu}_e$).

The purely leptonic τ decay:

$$\tau^- \to \mu^- \; \bar{\nu}_\mu \; \nu_\tau \tag{1.19}$$

can also be readily seen to conserve the lepton numbers.

Examples of decays that would violate lepton number conservation include:

$$\mu^- \to e^- \; \gamma, \tag{1.20}$$
$$\tau^- \to e^- \; \mu^+ \; \mu^-, \tag{1.21}$$
$$\tau^- \to \mu^- \; \gamma, \tag{1.22}$$

and these are indeed not seen experimentally.

Hadronic states are also labelled by quark numbers: strangeness S, charm C, bottom B and top T. These quantities relate to the quark content of the hadron, so for instance a strange quark gives a contribution of $S = -1$ to an hadron which contains it. The hadron electric charge Q and the baryon number B, defined as:

$$\tilde{B} \equiv \frac{1}{3}[N(q) - N(\bar{q})] \tag{1.23}$$

complete the list of internal quantum numbers. These are important because strong and electromagnetic interactions conserve all quark numbers, since in these interactions quark are only produced in quark–antiquark pairs. However, weak interactions violate quark numbers; an example is given by the neutron decay

$$n \to \; p \, e^- \bar{\nu}_e, \tag{1.24}$$

which written in terms of quarks is:

$$(udd) \to \; (uud) \, e^- \bar{\nu}_e. \tag{1.25}$$

It is immediately evident in this transition that a d quark is replaced by a u quark in the final state. The corresponding Feynman diagram is shown in Fig. 1.11. The hadron is represented by the parallel lines of its constituent quarks.

The quark and baryon numbers play a fundamental role in understanding the decay of hadrons. The vast majority of hadrons decay via the strong interactions to lighter particles with the same quark and baryon numbers. When there are no lighter hadrons with the same internal quantum numbers the decay may proceed through the weak interaction, in which only the baryon number and electric charge are conserved, but the lifetime of these processes is much longer, by about 10 orders of magnitude.

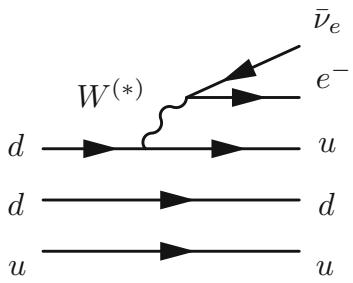


Fig. 1.11 Quark representation of the neutron decay

Glossary

Electronvolt (eV) Kinetic energy gained by an electron after accelerating through a potential difference of 1 volt

$$1\,\text{eV} = 1.602\,176\,565(35) \times 10^{-19}\,\text{J}.$$

SI International System of Units

Planck constant Fundamental constant that relates the energy of a photon with its frequency

$$h = E/\nu = 6.626\,069\,57(29) \times 10^{-34}\,\text{J s} = 4.135\,667\,516(91) \times 10^{-15}\,\text{eV s}$$

Ångström (Å) Distance corresponding to 10^{-10} m which is approximately the scale of an atom

Antimatter Collective name for the antiparticles

Baryons Composite particles which are bound states of three quarks or three anti-quarks

Boson Particle with spin 1, such as the photon and the Higgs particle

Feynman diagram Pictorial mathematical representation of particle interactions

Fermi (f) Distance corresponding to 10^{-15} m which is approximately the scale of the atomic nucleus

Fundamental forces The four known fundamental forces of nature: electromagnetic, weak nuclear, strong nuclear, and gravitational

Hadrons Composite particles which are bound states of quarks

Higgs boson Particle manifestation of the Higgs field, whose existence is key to explaining how leptons, quarks and the electroweak bosons acquire their own mass

Lepton Elementary matter particle with spin 1/2, such as the electron, the muon and the neutrino

Mesons Composite particles which are bound states of one quark and one antiquark

Muon Elementary unstable particle belonging to the second family of the leptons. Its mass is 106 MeV and mean lifetime of 2.2 μs

Quark Elementary matter particle, present for instance in the protons and neutrons of an atomic nucleus

Standard Model Quantum-mechanical relativistic theory of the electroweak and strong interactions between elementary particles

Speed of light Fundamental constant which in vacuum corresponds to

$$c = 299\,792\,458\,\mathrm{m\,s^{-1}} \simeq 3 \times 10^8\,\mathrm{m\,s^{-1}}$$

Top quark The heaviest of the quarks, with a mass of 173 GeV

Ultraviolet (UV) Part of the electromagnetic spectrum corresponding to wavelengths between 100 nm and 400 nm

X-ray (soft) Electromagnetic radiation corresponding to wavelengths between 0.01 nm and 10 nm

X-ray (hard) Electromagnetic radiation corresponding to energies of the order 10–100 keV

References

1. J. Beringer et al., (Particle Data Group). Phys. Rev. D **86**, 010001 (2012). 2013 partial update for the 2014 edition
2. P.A.M. Dirac, Proc. R. Soc. Lond. A **117** (1928). doi:10.1098/rspa.1928.0023
3. C. Anderson, Phys. Rev. **43**, 491 (1933)
4. ATLAS Collaboration, ATLAS-CONF-2012-169 (2012), https://cds.cern.ch/record/1499628. Accessed 10 June 2014
5. F. Englert, R. Brout, Broken symmetry and the mass of the gauge vector mesons. Phys. Rev. Lett. **13**, 321 (1964)
6. P.W. Higgs, Broken symmetries, massless particles and gauge fields. Phys. Lett. **12**, 132 (1964)
7. P.W. Higgs, Broken symmetries and the mass of the gauge bosons. Phys. Rev. Lett. **13**, 508 (1964)
8. ATLAS Collaboration, Phys. Lett. B **716**, 129 (2012)
9. C.M.S. Collaboration, Phys. Lett. B **716**, 3061 (2012)
10. Fermilab Creative Services, image number 11-0290-03D (2011). Credit: Fermilab/DZero Collaboration
11. M. Gell-Mann, Phys. Lett. **8**, 3 214–215 (1964). doi:10.1016/S0031-9163(64)92001-3
12. G. Zweig, CERN Report No.8182/TH.401, https://cds.cern.ch/record/352337/files/CERN-TH-401.pdf. Accessed 30 July 2014
13. G. Zweig, CERN Report No.8419/TH.412, http://library.lanl.gov/cgi-bin/getfile?00323548.pdf. Accessed 30 July 2014
14. S. Dürr et al., Science, **322**, 5905 1224–1227 (2008), http://dx.doi.org/10.1126/science.1163233. Accessed 23 Nov 2016

Chapter 2
Natural Sources of Radiation

Abstract The term *radiation* is commonly used to describe charged or neutral heavy particles (e.g. electrons, protons, neutrons, nuclei and nuclear fragments) and massless neutral particles (X-rays, γ-rays). However, strictly speaking and throughout this text all elementary particles, including neutrinos and any electromagnetic wave, can be thought of as a form of radiation. The sources of radiation in the Universe are many, but can be grouped for simplicity in natural and man-made. Natural radiation is radiation whose source can be traced to be of either terrestrial, solar or astronomical nature and as such is typically found in our everyday life. This chapter presents an account of only a handful of the most common natural radiation sources: the cosmic background radiation, the cosmic and solar radiation and the natural radioactivity. Despite being mostly invisible to our eye, radiation permeates the natural environment in which we live.

2.1 Cosmic Microwave Background

The *cosmic microwave background* (CMB) refers to the low energy photons permeating all the Universe and dominating the energy content of the radiation coming from beyond our Galaxy. Such thermal radiation, like a glow permeating the Universe at wavelengths in the millimeter region, is interpreted today as the remnant of the early Universe, some 380,000 years after the Big Bang. According to the Standard Model of Cosmology, the early Universe was primarily a plasma of protons and electrons with thermal electromagnetic radiation. As the Universe expanded and cooled, electrons and protons formed neutral hydrogen atoms and the radiation could no longer be absorbed, travelling freely and reaching the Earth today from all directions.

The CMB was discovered by Penzias and Wilson using a horn-reflector antenna at the Crawford Hill Laboratory in 1964 [2, 3], which is shown in Fig. 2.1. While attempting to estimate all contributions to background noise at a wavelength of 7.3 cm, it became evident that there was a residual of 3.5 ± 1.0 K of unknown origin. The spectrum of the CMB was measured for the first time by the Cosmic Background Explorer (COBE) satellite (1989–1993) and was found to be consistent with the emission of a *blackbody* at a temperature of $T = 2.726 \pm 0.010$ K [4, 5]. A blackbody is an ideal object that absorbs all incident electromagnetic radiation and

L. Cerrito, *Radiation and Detectors*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-53181-6_2

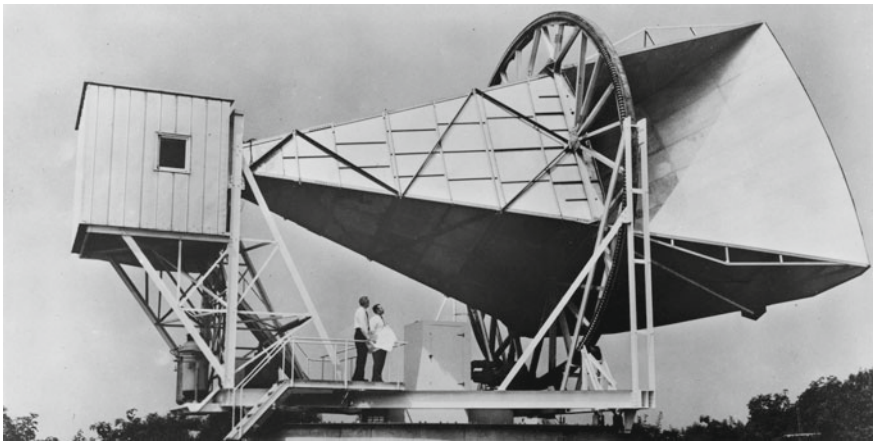


Fig. 2.1 The horn-reflector antenna at Bell Telephone Laboratories in Holmdel, New Jersey, built in 1959 and led to the discovery of the Cosmic Microwave Background radiation. Printed from [1] under CC-BY-2.0 licence

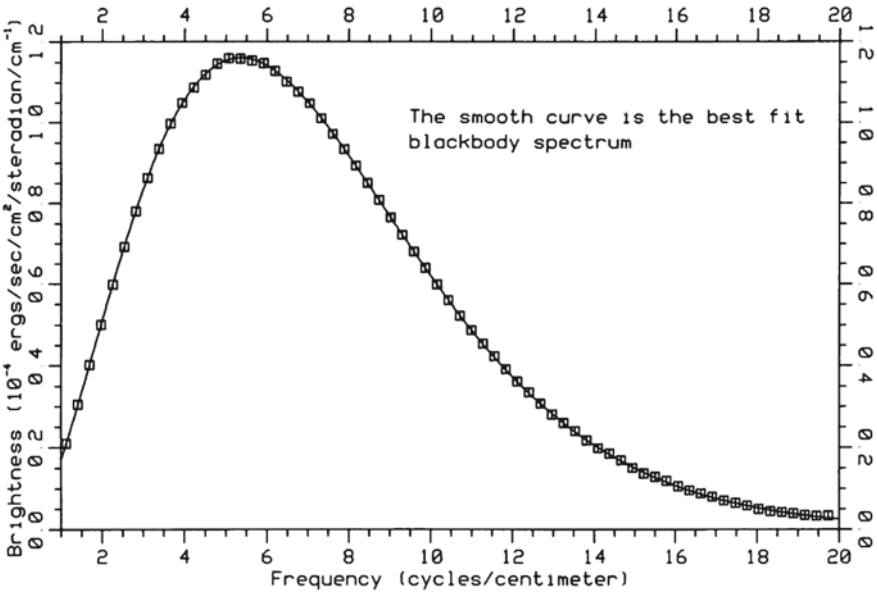


Fig. 2.2 The CMB spectrum measured by the FIRAS spectrometer on COBE compared to a blackbody spectrum. © AAS. Reproduced with permission from [4]

emits, when in thermal equilibrium, electromagnetic radiation called blackbody radiation, whose spectrum depends only on its temperature. The CMB spectrum peaks at approximately 160 GHz, or slightly less than 2 mm of wavelength, and is shown in Fig. 2.2 as measured by the Far InfraRed Absolute Spectrophotometer (FIRAS)

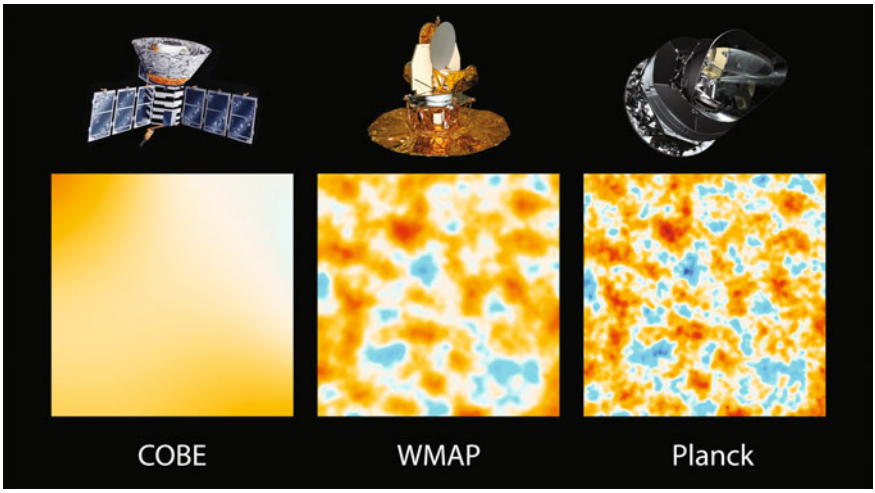


Fig. 2.3 The evolution of satellites dedicated to CMB measurements, the COBE, WMAP and Plank missions, showing the progressively increasing resolution. Printed with permission. Courtesy NASA/JPL-Caltech/ESA [8]

onboard COBE. The FIRAS [6, 7] was a rapid-scan polarising Michelson interferometer, covering the wavelength range 0.1–10 mm and designed to measure small deviations (0.1%) from an internal reference blackbody spectrum. The instrument was cooled at 1.5 K with liquid helium to reduce its termal emissions and enable the use of sensitive *bolometric* detectors (see Sect. 9.5).

The COBE mission did not only measure the CMB spectrum, it also carried an instrument (Differential Microwave Radiometer, DMR) to measure the anisotropy, that is the fluctuations in the brightness of the CMB versus angle. Such fluctuations are crucial to understand the evolution of the universe at the time of formation of neutral atoms. Today the CMB temperature is known with a precision of 1 part in 10,000, $T = 2.7255 \pm 0.0006$ K and shows ripples varying in angle by roughly 1 part in 100,000. The CMB has been measured with increasing precision by the National and Aeronautics Space Administration (NASA) Wilkinson Microwave Anisotropy Probe (2012) and the European Space Agency's (ESA) Planck space observatory (2013), see Fig. 2.3 for a comparison of the resolution.

2.2 Cosmic Radiation

The Earth is constantly reached by a flux of particles whose origin can be traced to sources located within Galaxies in the Universe. This flux is commonly referred to as *cosmic radiation*. The particles incident a the top of the Earth's atmosphere include all stable particles and nuclei with lifetimes of the order of 1 or more million

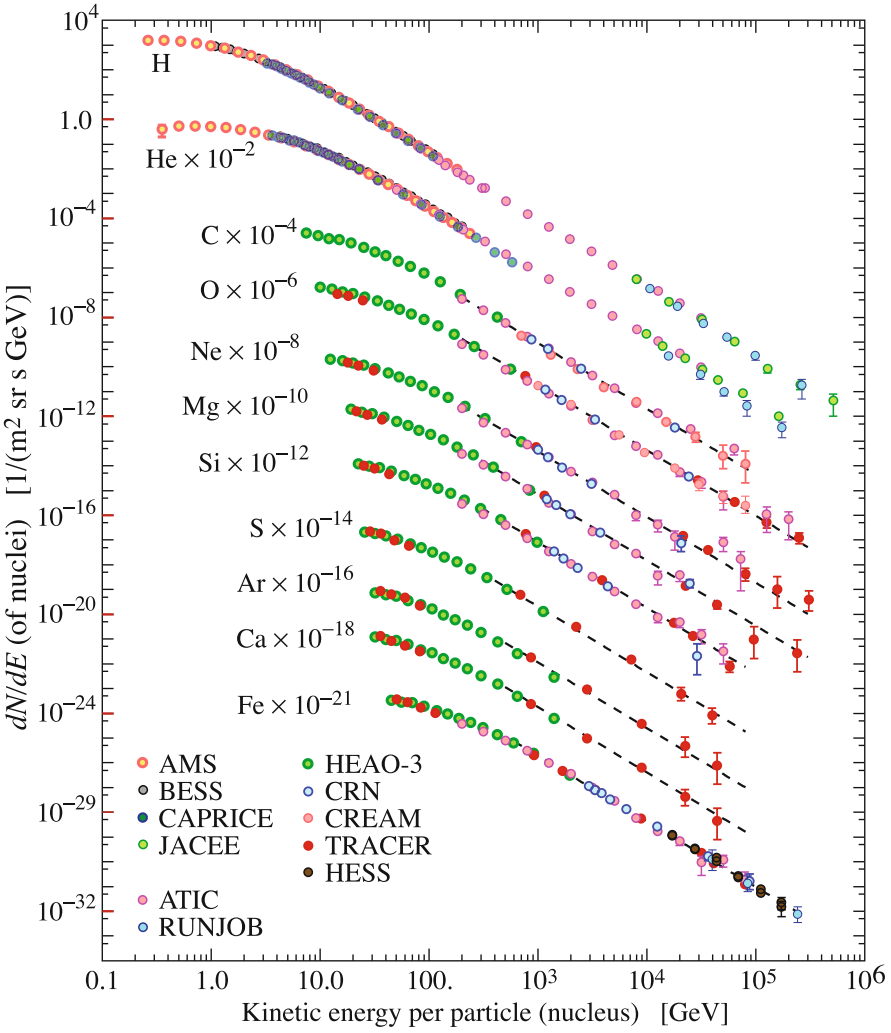


Fig. 2.4 The flux of nuclei of the primary cosmic radiation versus energy-per-nucleus. Reprinted with permission from [9]

years. The cosmic radiation is classed as "primary" if the particles are accelerated at astrophysical sources, while particles produced from subsequent interactions of the primary with interstellar gas are referred to as the "secondary" cosmic radiation. The cosmic flux includes protons, electrons and various nuclei, particularly light nuclei (primarily Helium). Antiprotons and positrons are also present and mainly part of the secondary cosmic radiation.

About 80% of the primary nucleons are free protons, and the flux is nearly isotropic at most energies. Figure 2.4 shows the constituents of the cosmic flux on top of

the atmosphere for energies greater than 2 GeV/nucleon. The energy spectrum falls rapidly from a few GeV to the TeV range with a power law. The flux in the interesting range of thousands of TeV is rather small, approximately one event per square kilometer per year. The Fig. 2.4 shows also how the relative abundance of the primary nuclei is nearly constant. The ratio of positrons to electrons is about 5% at 10 GeV and its evolution with energy is still the subject of much interest. The ratio of antiprotons to protons is of order of 1 in 10,000 at 10–20 GeV.

Once particles from the cosmic flux reach the top of the Earth's atmosphere, the interaction will produce a cascade of particles and the development of a cosmic *shower*. The plot in Fig. 2.5 shows the estimated vertical flux of electrons, nucleons, pions, muons and neutrinos in the atmosphere with energy greater than 1 GeV. The plot also shows the measured negative muon flux [10–15]. Notice that at high altitudes, for instance around 10,000 m which is the cruising altitude of commercial jets, the leading particles of the cosmic shower are protons and neutrons. It is this the reason for commercial airlines to have protocols in place to limit the exposure to cosmic radiation of their employees. The impact of radiation on biological tissues will be discussed in Chap. 3.

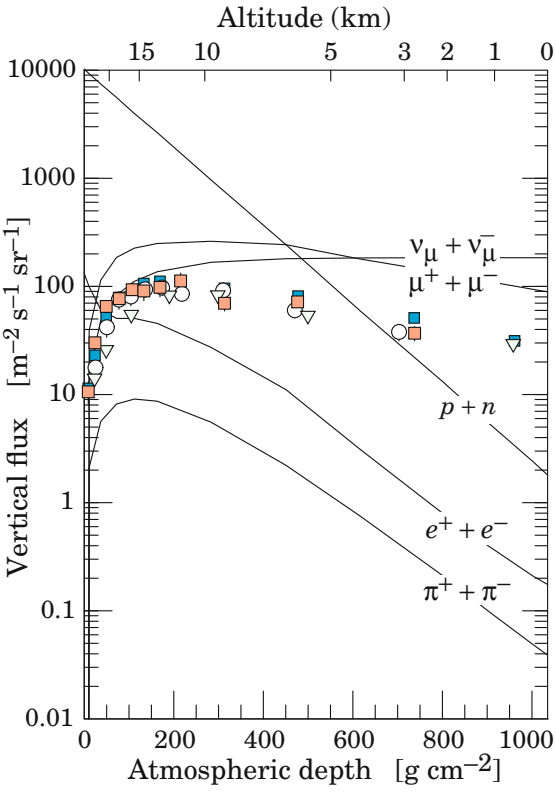


Fig. 2.5 The estimated vertical flux of the main components of a cosmic shower with $E > 1$ GeV in the atmosphere. The points show also measurements of the negative muon flux [10–15]. Reprinted with permission from [9]

Notable examples of nucleon interactions of the cosmic radiation with the atmosphere include also the reactions:

$$n + {}^{14}\mathrm{N} \rightarrow p + {}^{14}\mathrm{C}, \tag{2.1}$$

$$p + {}^{16}\mathrm{O} \rightarrow n + \pi + X. \tag{2.2}$$

The first of these two reactions is exploited in the so-called *radiocarbon* technique used to date organic materials; the ^{14}C is a radioactive element which stays in plants and animals in the form of carbon dioxide (CO_2), with a concentration that matches that of the atmosphere. Since the ^{14}C has a half-life of about 5,730 years, the remaining concentration after a period of time can be related to the epoch a plant or animal stopped exchanging ^{14}C with its surroundings.

Muons and neutrinos are instead the product of the decay chain of charged mesons produced in the upper atmosphere, particularly pions (see (2.2)) and are the most abundant particles from an altitude of about 4,000 m down to sea level. The energy distribution of the vertical muons peaks at around 30 GeV, but spans a wide range from a few GeV to several hundred GeV. The mean energy of the muons at sea level is approximately 4 GeV. The Fig. 2.6 shows the muon and neutrino momentum spectra at an incident angle of $\theta = 0$. Moreover, muons are more abundant than antimuons

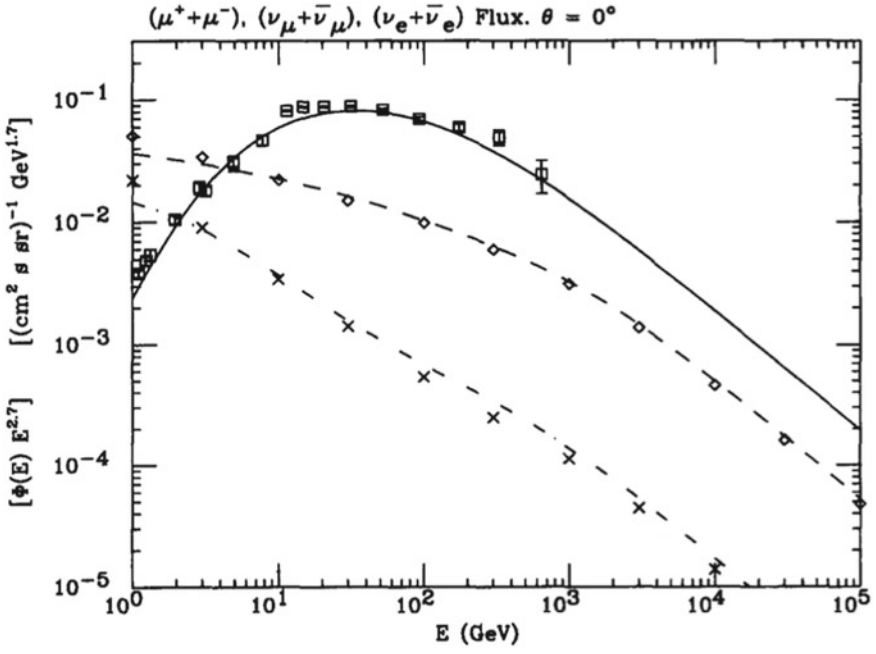


Fig. 2.6 Models (*solid* and *dashed lines*) and measured data of the vertical muon and neutrino fluxes at sea level. Reprinted from [16] with permission from Elsevier

by 20–40% over a wide range of momenta. This charge asymmetry is explained by the dominance of protons over antiprotons in the incident cosmic flux.

One of the recent experiments designed for the study of the cosmic showers reaching the ground is the Pierre Auger Observatory, in Argentina. The detector array covers an area of 3,000 km^2 and is designed to measure ultra-high-energy cosmic rays with energy beyond 10^6 TeV. The array of 1,600 surface detectors is capable of measuring the flux, arrival direction, distribution, and mass composition of cosmic rays.

2.3 Solar Radiation

Solar radiation is the part of cosmic radiation that is originated from our Sun. It is composed of electromagnetic radiation, neutrinos from the core, and massive particles (protons, electrons, etc.) from the Sun's surface. Of these three, the ultraviolet (UV) component of the electromagnetic spectrum is the only damaging component that reaches sea level, only partially filtered by the ozone layers in the atmosphere.

The electromagnetic radiation of the Sun corresponds to a blackbody spectrum with temperature of approximately 5,776 K. The Fig. 2.7 shows the solar spectrum (irradiance) between wavelengths of 1 nm (extreme UV) and 10^4 nm (infrared). The graph shows the spectral irradiance, the characteristic altitude of absorption in the Earth's atmosphere, and the relative and absolute variability during solar cycle variations. Atmospheric absorption can be seen to suppress the dangerous UV component with wavelengths below 300 nm. At sea level the solar spectrum exhibits absorption lines from atmospheric molecules, e.g. from O_3, H_2O and CO_2.

Neutrinos, unlike electromagnetic radiation, are harmless to humans and can be detected on the ground with special neutrino detectors. In the Sun's core, neutrinos are produced through several reactions but the vast majority of neutrinos is produced through the so-called *p-p chain*:

$$p + p \rightarrow {}^{2}_{1}\mathrm{D} + e^{+} + \nu_e, \tag{2.3}$$

releasing neutrinos with energy less than 0.42 MeV. The Table 2.1 shows the neutrino-producing reactions in the Sun and the neutrino's typical energy. More energetic neutrinos are produced by the *p-e-p* reaction ($E_{\nu_e} = 1.442$ MeV):

$$p + e^{-} + p \rightarrow {}^{2}_{1}\mathrm{D} + \nu_e, \tag{2.4}$$

or through the reaction (up to 15 MeV):

$${}^{8}_{5}\mathrm{B} \rightarrow {}^{8}_{4}\mathrm{Be}^{*} + e^{+} + \nu_e. \tag{2.5}$$

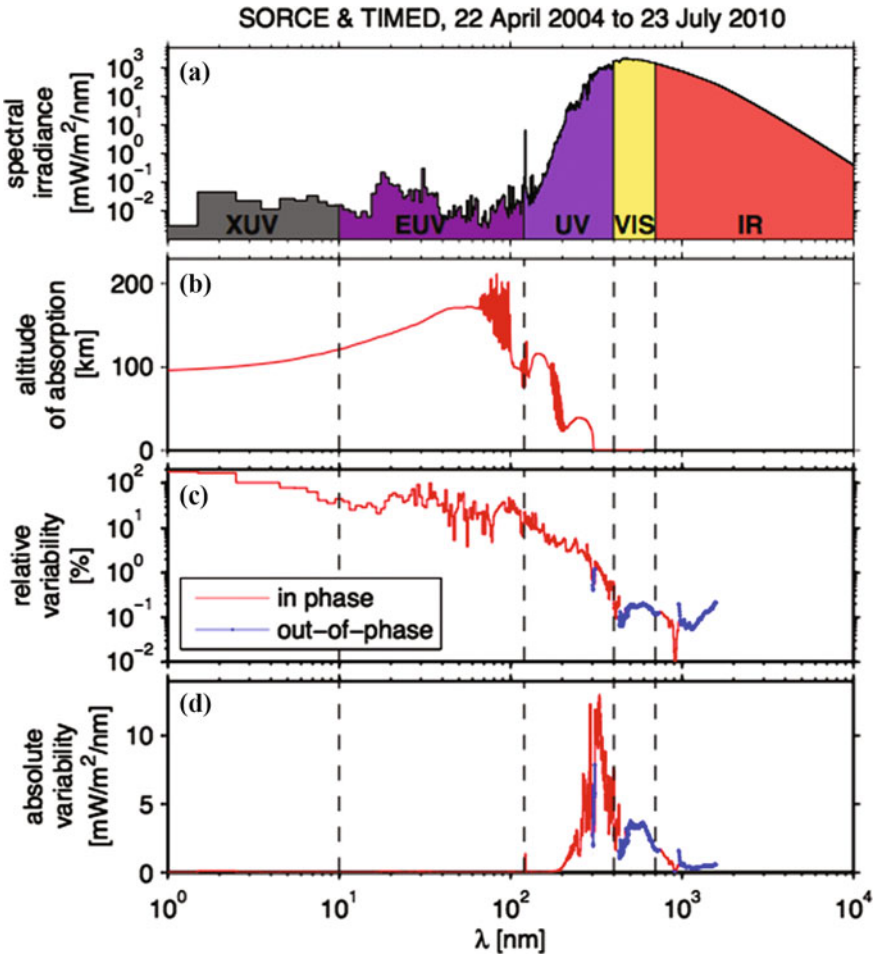


Fig. 2.7 The solar spectral irradiance inferred from the SOlar Radiation and Climate Experiment (SORCE) satellite and the Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED) data. Reprinted with permission from [17]

The plot in Fig. 2.8 shows the flux of solar neutrino from the various reaction sources versus the emission energy. The solar neutrino flux was measured for the first time by the Homestake experiment [18] (South Dakota, USA, 1969–94). The experiment was based on a detector consisting of a tank filled with 615 tons of tetrachloroethylene (C_2Cl_4, a cleaning agent), lowered in the Homestake gold mine 1,478 meters underground. In the tank, the following reaction was being produced by the solar electron-neutrinos: $^{37}\mathrm{Cl} + \nu_e \rightarrow ^{37}\mathrm{Ar} + e$. Since the ^{37}Ar is a radioactive source, a measurement of its concentration could be related to the number of reactions from the neutrinos, hence the neutrino flux. The energy threshold of the ^{37}Cl-^{37}Ar absorption reaction is 0.814 MeV.

Table 2.1 Neutrino-producing reactions in the Sun and the neutrino's typical energy (summary from [18])

Reaction	E_ν [MeV]
Proton–Proton chain:	
Phase 1:	
$p + p \rightarrow {}^2_1\mathrm{D} + e^+ + \nu_e$	≤ 0.42
$p + e^- + p \rightarrow {}^2_1\mathrm{D} + \nu_e$,	1.442
$p + {}^2_1\mathrm{D} \rightarrow {}^3_2\mathrm{He} + \gamma$	–
${}^3_2\mathrm{He} + {}^3_2\mathrm{He} \rightarrow {}^4_2\mathrm{He} + e^+ + \nu_e$	≤ 18.8
Phase 2:	
${}^3_2\mathrm{He} + {}^4_2\mathrm{He} \rightarrow {}^7_4\mathrm{Be} + \gamma$	-
${}^7_4\mathrm{Be} + e^- \rightarrow {}^7_3\mathrm{Li} + \nu_e$	0.861 (90%)
${}^7_3\mathrm{Li} + p \rightarrow {}^4_2\mathrm{He} + {}^4_2\mathrm{He}$	–
Phase 3:	
${}^7_4\mathrm{Be} + p \rightarrow {}^8_5\mathrm{B} + \gamma$	–
${}^8_5\mathrm{B} \rightarrow {}^8_4\mathrm{Be}^* + e^+ + \nu_e$	≤ 15
${}^8_4\mathrm{Be}^* \rightarrow {}^4_2\mathrm{He} + {}^4_2\mathrm{He} + \gamma$	≤ 15
Carbon-Nitrogen cycle:	
$p + {}^{12}_6\mathrm{C} \rightarrow {}^{13}_7\mathrm{N} + \gamma$	–
${}^{13}_7\mathrm{N} \rightarrow {}^{13}_6\mathrm{C} + e^+ + \nu_e$	≤ 1.20
$p + {}^{13}_6\mathrm{C} \rightarrow {}^{14}_7\mathrm{N} + \gamma$	–
$p + {}^{14}_7\mathrm{N} \rightarrow {}^{15}_8\mathrm{O} + \gamma$	–
${}^{15}_7\mathrm{O} \rightarrow {}^{15}_7\mathrm{N} + e^+ + \nu_e$	≤ 1.73
$p + {}^{15}_7\mathrm{N} \rightarrow {}^{12}_6\mathrm{C} + {}^4_2\mathrm{He}$	–

The Homestake experiment detected fewer than one third of the neutrinos expected from the solar model prediction of its time [20]. Over its lifespan of 25 years, the experiment in fact measured a solar neutrino-induced ^{37}Ar production rate of 2.56 ± 0.23 SNU (one Solar Neutrino Unit, SNU, corresponds to one interaction per 10^{36} target atoms per second). Once compared with the expected 9.3 ± 1.3 SNU [21], the Homestake result provided the first evidence of solar neutrino disappearance, a phenomena now understood to be due to neutrino oscillations (spontaneous changing from one type to another). The neutrino detectors will be discussed in detail in Chap. 11. The first detection of neutrinos and X-ray from space was recognised and celebrated with the 2002 Nobel prize in Physics for R. Davis Jr., N. Koshiba, and R. Giacconi.

Massive particles (electrons, protons and nuclei) are also accelerated from the surface of the Sun in every direction, referred to as the *solar wind*. Luckily for us, the Earth's magnetic field bends away these charged particles quite effectively, except at the geomagnetic poles, where they penetrate and interact with the atmosphere creating visible auroras. The composition and dynamics of the solar wind is best studied with instruments onboard satellites. The flux and speeds of these particles

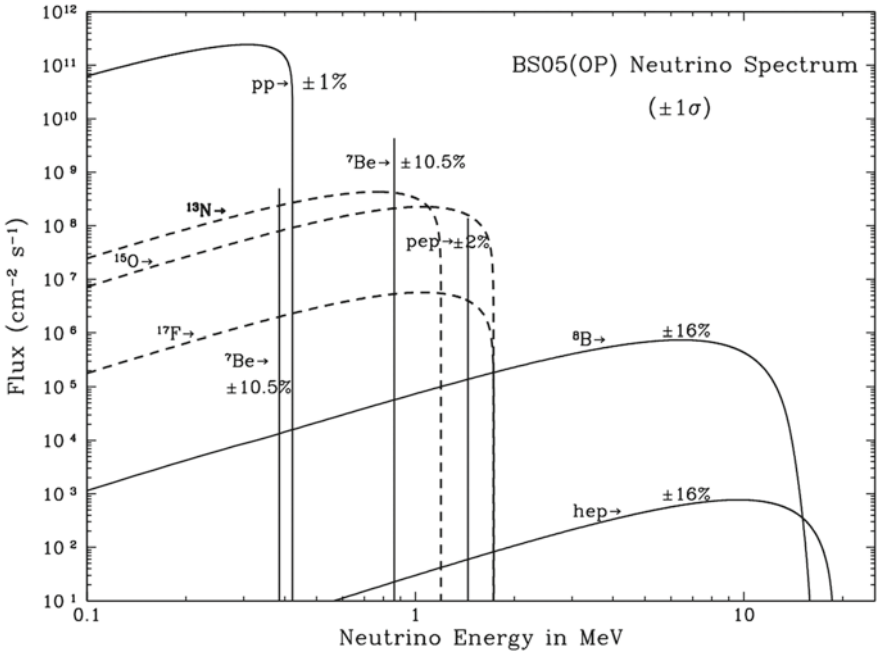


Fig. 2.8 The solar neutrino energy spectrum according to one specific solar model. © AAS. Reprinted from [19] with permission

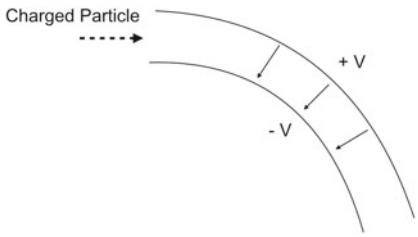


Fig. 2.9 Schematic diagram of an electrostatic analyser for electrons and ions

has been measured by various instruments, one of which was the Solar Wind Observations Over the Poles of the Sun (SWOOPS) electrostatic analyser, part of the Solar Wind Plasma Experiment onboard the National Aeronautics and Space Administration (NASA) and European Space Agency (ESA) Ulysses spacecraft (1990–2009). The SWOOPS instrument provided measurements of the solar electrons, hydrogen, helium, and heavy-ion components at latitudes from the equator to near the solar poles [22, 23]. An electrostatic analyser is an instrument used to measure the flux of ions and electrons of a specific energy. The simplest configuration (see Fig. 2.9) consists of curved parallel plates held at high and different potentials which create

a radial electric field. Only electrons and ions of a specific energy can perform the desired trajectory and exit the plates, according to the simple equation:

$$\frac{mv^2}{R} = qE, \tag{2.6}$$

where m, q and v are the particle mass, electric charge and speed, respectively, and E and R are the electric field within the plates and their radius of curvature. Particles that do not perform the curved trajectory collide with the walls of the plates, which are typically copper plated and blackened to avoid UV reflections. If the energy of the charged particles is indicated by $\mathscr{E}$, an electrostatic analyser filters particles with a specific $\mathscr{E}/q$. The diagram in Fig. 2.10 shows the sensor of the Solar Wind Ion Composition Spectrometer (SWICS) onboard the Advanced Composition Explorer (ACE) spacecraft (in operation since 1997). The electrostatic deflection system is visible on the right hand side. The schematics for the ion optics of the SWOOPS instrument can instead be found in [24].

The plot in Fig. 2.11 shows the measured speed of the solar wind as a function of latitude during a solar minimum. The speeds range between about 300–400 km/s

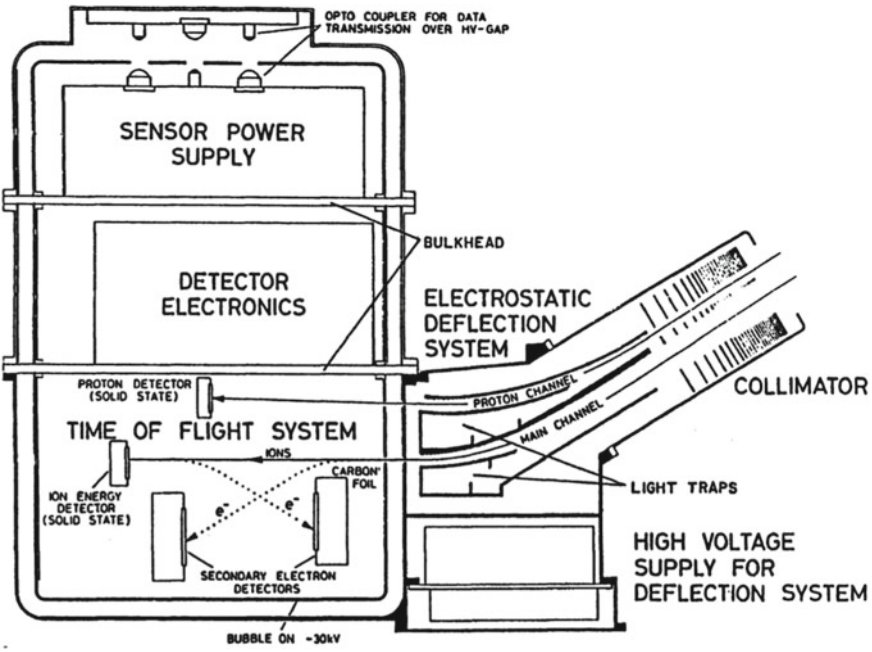


Fig. 2.10 Cross-section of the SWICS sensor onboard the ACE robotic spacecraft. The collimator and two-channel deflection system are visible on the right hand side. Printed from [25] with permission

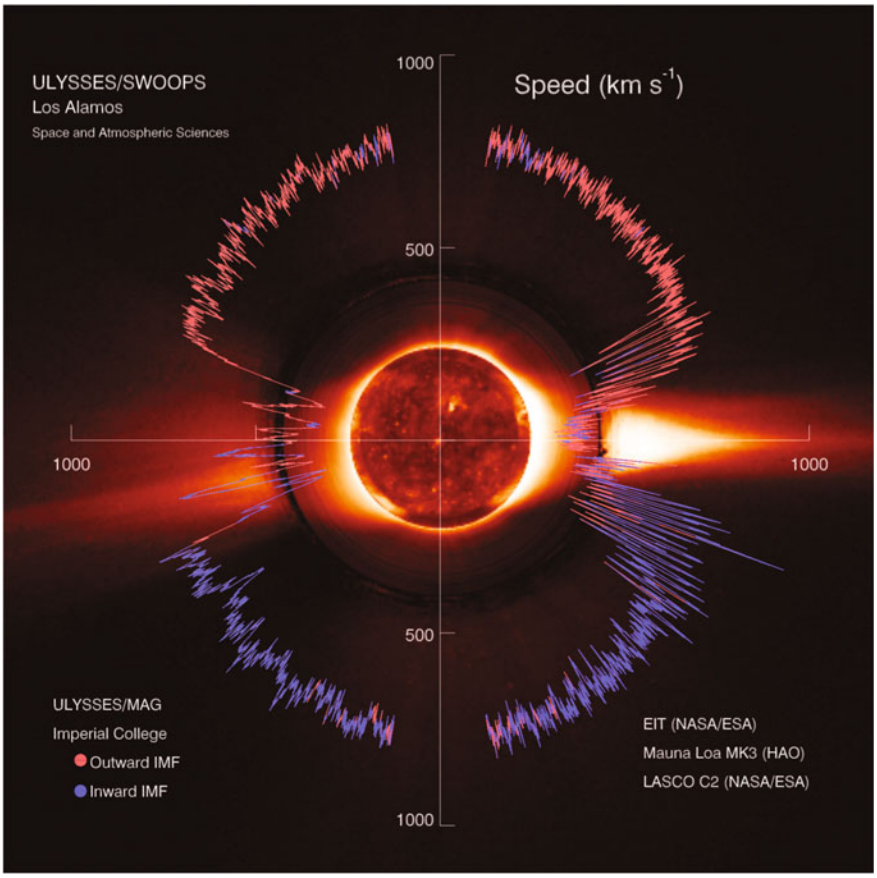


Fig. 2.11 The speed of the solar wind measured by the SWOOPS instrument onboard the Ulysses satellite. Printed from [26] with permission

at the equator to about 700 km/s elsewhere. Such speeds, clearly non relativistic, for ions correspond to a kinetic energy of up to a few MeV. The Ulysses spacecraft completed three orbits of the Sun during solar minimum and maximum cycles [24].

2.4 Natural Radioactivity

Natural radioactivity is commonly present in nature and is due to transitions between energy levels of nuclei or caused by nuclear fragmentation or from the re-adjustment in the electron cloud of atoms.

A nucleus is characterised by the number of protons and neutrons that it contains, so it is commonly indicated by letters and numbers showing the atomic number

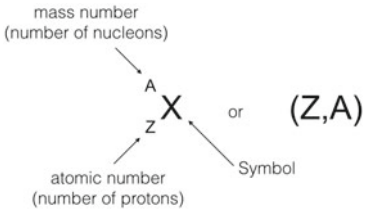


Fig. 2.12 Scheme for indicating atomic nuclei through a symbol, atomic number and mass number

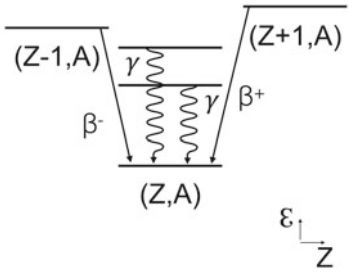


Fig. 2.13 Nuclear energy-level and transitions diagram

(the number of protons) and the mass number (the number of nucleons), as shown in the diagram of Fig. 2.12. For instance, a nucleus of Deuterium (D), containing 1 proton and 1 neutron would be indicated as ${}^{2}_{1}\mathrm{H}$.

The most common types of nuclear emissions are referred to as β, γ and α decays. The transitions between energy levels are often visualised through energy level diagrams as shown in Fig. 2.13. The diagram pictures nuclei with varying Z on the horizontal direction and the energy levels of the nuclei, $\mathscr{E}$, on the vertical direction. The vertical arrows between levels show transitions associated with a particular type of radiation.

The Table 2.2 summarises the most common types of radiation produced by processes of natural radioactivity, with their common name, origin, mass and typical energy. A description is given in the following list.

Beta radiation:
The name beta (β) radiation refers to fast electrons or positrons which result from the weak-interaction decay of neutrons or protons within a nucleus:

$$(Z, A) \rightarrow (Z+1, A) + e^- + \bar{\nu}_e \qquad (n \rightarrow p + e^- + \bar{\nu}_e), \tag{2.7}$$

$$(Z+1, A) \rightarrow (Z, A) + e^+ + \nu_e \qquad (p \rightarrow n + e^+ + \nu_e). \tag{2.8}$$

The spectrum of energy of the emitted β electron or positron is continuous, because the available energy is shared with the outgoing neutrino, with a maximum value (called the spectrum *end-point*) corresponding to the difference between the energy levels, and is typically of the order of keV to MeV. An example of a pure β^- emitter is the ${}^{14}\mathrm{C}$, already encountered in the description of cosmic radiation in Sect. 2.2. Other notable examples include ${}^{60}\mathrm{Co}$ and ${}^{137}\mathrm{Cs}$. A list of some β emitters with their half-life and typical emission energy is shown in Table 2.3. Many β emitters are used in

Table 2.2 Characteristic types of radiation in natural radioactivity

Name	Particle type	Origin	Charge [e]	Mass [MeV/c^2]	Energy distribution
α	$2p2n$ (nucleus)	Nuclear	$+2$	3.73	Discrete [MeV]
$\beta^{\pm}$	$e^{\pm}$	Nuclear	± 1	0.51	Continuous [keV–MeV]
γ	EM radiation [$\lambda < 10^{-3}$ nm]	Nuclear	0	0	Discrete [keV–MeV]
Conversion	e^-	Nuclear	-1	0.51	Discrete [high keV]
Fission fragment	nuclei	Nuclear	$\simeq 20$	80–160	Continuous [30–150 MeV]
Neutron	n	Nuclear	0	939.6	Continuous or discrete [keV–MeV]
X-ray	EM radiation [$\lambda \sim 0.01 - 10$ nm]	Atomic shell	0	0	Discrete [eV–keV]
Auger	e^-	Atomic shell	-1	0.51	Discrete [eV–keV]

Table 2.3 A list of β emitters ordered in decreasing end-point energy. The data is from [27]

Isotope	Half-life [days]	End-point energy [keV]	Average energy [keV]
^{66}Cu	0.0035	2641	1112
^{90}Y	2.67	2280.1	933.7
^{32}P	14.3	1710.7	695.0
^{89}Sr	50.6	1500.1	587.1
^{39}Ar	$98.2 \cdot 10^3$	565.5	218.8
^{10}Be	$5.5 \cdot 10^8$	556.0	202.6
^{115}In	$1.6 \cdot 10^{17}$	497	153.2
^{121}Sn	1.1	390.6	115.8
^{113}Cd	$2.9 \cdot 10^{18}$	322.0	92.6
^{99}Tc	$7.7 \cdot 10^7$	293.5	84.6
^{45}Ca	162.6	255.8	76.9
^{33}P	25.3	248.5	76.4
^{135}Cs	$8.4 \cdot 10^8$	268.7	75.7
^{35}S	87.4	167.3	48.7
^{14}C	$2.1 \cdot 10^6$	156.5	49.5
^{79}Se	$1.1 \cdot 10^8$	151.0	55.8
^{63}Ni	$3.7 \cdot 10^4$	66.9	17.4
^{107}Pd	$2.4 \cdot 10^9$	34.3	9.3
^{3}H	4498.8	18.6	5.7
^{187}Re	$1.6 \cdot 10^{13}$	2.5	0.6

medical diagnosis, imaging and treatment. Most β transitions are also accompanied by a subsequent γ emission (see next paragraph).

Gamma radiation and X-ray:
The transition between nuclear energy levels leads typically to the emission of gamma (γ) rays (see Fig. 2.13). The emitted photons are mono-energetic, with energy equal to the difference in the corresponding nuclear energy levels; energy values are in the keV to MeV range. The X-rays can instead be produced in the rearrangement of the atomic shell. Their typical energy is in the eV to keV range, with a discrete spectrum. In addition, electromagnetic radiation in the form of *bremsstrahlung* (described in more detail in Chap. 10) may be produced when β electrons interact with the surrounding material and slow down. The resulting photons have a continuous energy spectrum that can extend up to the energy of the parent electron, although most of the emission is in the keV region.

Alpha radiation:
The alpha (α) radiation consists of emission of so-called α particles, i.e. ^{4}He nuclei consisting of 2 protons and 2 neutrons. The electric charge is therefore $+2e$. The emission of α is most common in reactions involving heavy nuclei, e.g. through the process:

$$(Z, A) \rightarrow (Z - 2, A - 4) + \alpha. \tag{2.9}$$

The energy spectrum of the α particles is mono-energetic and relatively high-energy, typically in the 2–10 MeV range, because of its nuclear origin. Given the high $+2e$ electric charge, α particles loose energy rapidly when interacting electromagnetically with matter, and travel typically only a few centimeters. Examples of α emitters include uranium, thorium, radium and other transuranic elements.

Annihilation radiation:
Annihilation radiation refers to the emission of photons from the process:

$$e^+ + e^- \rightarrow 2\gamma. \tag{2.10}$$

Electrons are readily found in the outer shell of atoms, while positrons might be emitted in some nuclear decays. For instance, a positron emitter is ^{22}Na, which when enclosed in an absorbing material gives rise to annihilation processes. The two photons are distinguishable from other forms of γ radiation because of their distinctive energy of 511 keV (each equals the electron mass which has been converted into energy). In the center of mass frame, the photons are also emitted back-to-back to conserve momentum.

Internal Conversion:
The internal conversion is the name for the mono-energetic emission of β^- from nuclear de-excitation. While de-excitation through γ emission is most common, sometimes a nucleus can de-excite through a kinetic-energy transfer to an orbital electron. The typical electron energies range from the high keV to MeV.

Electron Capture and Auger electrons:
This process refers to the capture of an inner electron by a nucleus through the reaction:

$$p + e^- \rightarrow n + \nu_e. \tag{2.11}$$

The subsequent filling of the inner-shell vacancy is accompanied by a characteristic X-ray photon emission, and sometimes by the ejection of one of the outer electrons (called an *Auger* electron), in cases where the latter acquires sufficient kinetic energy directly from the de-excitation of the atom. The Auger electrons have a discrete energy spectrum with a few keV energy, therefore softer than β and conversion electrons, and are favoured only in low-Z elements for which the binding energies are small.

Spontaneous Fission and Neutrons:
The spontaneous split of a nucleus is the only natural process that can produce energetic particles heavier than the α. Due to the nuclear potential energy barrier, this process is naturally favoured only for some transuranic isotopes of very large mass number. A common element which undergoes spontaneous fission is Californium ^{252}Cf, which is the heaviest element to occur naturally on Earth. Each of the two fission fragments is emitted back to back with kinetic energies of around 100 MeV. Neutrons of a few MeV can be emitted either as result of spontaneous fission or as part of a chain of nuclear reactions.

Glossary

α **radiation** Emission of α particles from nuclear decays

β **radiation** Fast electrons or positrons which result from the weak-interaction decay of neutrons or protons within a nucleus

β **emitters** Atoms emitting β radiation

γ **radiation** Emission of γ rays from the transition between nuclear energy levels

Annihilation radiation Emission of photons from the annihilation of an electron and positron pairs

Auger electrons Emission of one of the outer atomic electrons following an electron capture by the nucleus

Cosmic Microwave Background (CMB) Low energy photons found everywhere in the Universe which are remnant of the early moments of the Universe. The spectrum is consistent with a blackbody emission at a temperature of 2.7 K

Cosmic radiation Flux of particles originating from the galaxies in the Universe; it includes all stable particles and nuclei (particularly light ones, e.g. Helium)

Cosmic shower Shower of particles originating in the Earh's atmosphere as a result of cosmic radiation colliding with the air molecules; it includes mainly electrons, nucleons, pions, muons and neutrinos

Electrostatic analyser Instrument widely used in Space to measure the flux of ions and electrons of a specific energy

Internal conversion Mono-energetic emission of β^- from nuclear de-excitation

Natural radioactivity Emission of particles due to transitions between energy levels of nuclei or caused by nuclear fragmentation or from the re-adjustment in the electron cloud of atoms

p-p chain Nuclear reaction taking place aboundandly in the Sun's core which produces electron neutrinos with energy less than 0.42 MeV:

$$p + p \rightarrow {}^{2}_{1}\mathrm{D} + e^{+} + \nu_e,$$

p-e-p chain Nuclear reaction taking place aboundandly in the Sun's core which produces electron neutrinos with energy of 1.442 MeV

$$p + e^{-} + p \rightarrow {}^{2}_{1}\mathrm{D} + \nu_e,$$

Solar radiation Cosmic radiation which is originating from the Sun; it includes primarily electromagnetic radiation, neutrinos from the core, and massive particles (protons, electrons, etc.) from the Sun's surface

Solar wind Massive particles (electrons, protons and nuclei) which are accelerated from the surface of the Sun

Spontaneous fission The spontaneous split of a nucleus, naturally favoured only for some transuranic isotopes of very large mass number.

References

1. NASA image, Remix Project, Image no. GPN-2003-00013, https://www.flickr.com/photos/remix-man/3947459724/in/photolist-71PLfq-6gbrmg/. Printed under CC-BY 2.0 licence. Accessed 28 Nov 2016
2. A.A. Penzias, R. Wilson, Astrophys. J. **142**, 419 (1965)
3. R.H. Dicke et al., Astrophys. J. **142**, 414 (1965)
4. J.C. Mather et al., Astrophys. J. **354**, L37–L40 (1990). doi:10.1086/185717
5. J.C. Mather et al., Astrophys. J. **420**, 439–444 (1994). doi:10.1086/173574
6. J.C. Mather et al., Opt. Eng. **21**, 769 (1982)
7. J.C. Mather et al., Appl. Opt. **23**, 3181 (1984)
8. Image no. PIA16874 (2013), http://photojournal.jpl.nasa.gov/catalog/PIA16874. Accessed 28 Nov 2016. Courtesy of NASA/JPL-Caltech/ESA
9. C. Amsler et al., (Particle Data Group), Phys. Lett. B **667**, 1 (2008)

10. R. Bellotti et al., Phys. Rev. D **53**, 35 (1996)
11. R. Bellotti et al., Phys. Rev. D **60**, 052002 (1999)
12. M. Boezio et al., Phys. Rev. D **62**, 032007 (2000)
13. M. Boezio et al., Phys. Rev. D **67**, 072003 (2003)
14. S. Coutu et al., Phys. Rev. D **62**, 032001 (2000)
15. S. Haino et al., Phys. Lett. B **594**, 35 (2004)
16. P. Lipari, Astropart. Phys. **1**, 195–227 (1993)
17. I. Ermolli et al., Recent variability of the solar spectral irradiance and its impact on climate modelling. Atmos. Chem. Phys. **13**, 39453977 (2013). doi:10.5194/acp-13-3945-2013. Reprinted under CC-BY 3.0 license
18. B.T. Cleveland et al., ApJ **496**, 505 (1998). doi:10.1086/305343
19. J.N. Bahcall, A.M. Serenelli, S. Basu, Astrophys. J. Lett. **621**, L85 (2005). doi:10.1086/428929
20. J.N. Bahcall, N.A. Bahcall, G. Shaviv, Phys. Rev. Lett. **20**, 1209 (1968)
21. J.N. Bahcall, M.H. Pinsonneault, Rev. Mod. Phys. **67**, 781 (1995)
22. K.P. Wenzel, R.G. Marsden, B. Battrick, The international solar polar mission: its scientific investigations, European Space Agency ESA-SP-1050 ; SP-1050 (1983), p. 57
23. S.J. Bame, D.J. McComas, B.L. Barraclough, J.L. Phillips, K.J. Sofaly, J.C. Chavez, B.E. Goldstein, R.K. Sakurai, Astron. Astrophys. Supplement Ser. **92**(2), 237–265 (1992)
24. D.J. McComas, R.W. Ebert, H.A. Elliott, B.E. Goldstein, J.T. Gosling, N.A. Schwadron, R.M. Skoug, Geophys. Res. Lett. **35**, L18103 (2008). doi:10.1029/2008GL034896
25. G. Gloeckler et al., Space Sci. Rev. **86**(1), 497–539 (1998)
26. D.J. McComas et al., Ulysses' return to the slow solar wind. Geoph. Res. Lett. **25**(1), 1–4 (1998)
27. National Nuclear Data Center, information extracted from the NuDat 2 database, http://www.nndc.bnl.gov/nudat2/. Accessed 27 Jan 2015

Chapter 3
Dosimetry

Abstract Radiation interacts with matter by *exciting* and *ionising* the atoms that form the molecules or crystal structure of the material encountered. The physics of atomic excitation and ionisation is discussed in some detail in Part II of this book, particularly in the context of exploiting such mechanisms for detection and research purposes. This chapter focuses instead on the extent of measuring radiation for the purpose of quantifying its effects, including those on biological tissue. The chapter introduces the basic quantities to describe radiation and its impact, with the latter sections dedicated to the biological effects and the recommended radiological dose limits.

3.1 Flux, Activity and the Radioactive Decay Law

Consider a number of particles (N_{inc}) incident onto a surface of area A in a time interval ΔT (see Fig. 3.1), and let the particles be distributed uniformly in space and time. The *flux* (ϕ) (also *fluence rate*) is defined as the ratio:

$$\phi = \frac{N_{inc}}{A\ \Delta T}, \tag{3.1}$$

that is the number of incident particles per unit area, per unit time. The flux is measured in $\mathrm{m}^{-2}\mathrm{s}^{-1}$. The *fluence* ($\Phi$) is the particle flux integrated over a period of time, i.e.

$$\Phi = \frac{N_{inc}}{A}, \tag{3.2}$$

The *activity* of a radioactive source is defined as the number of decays per second and in the SI is measured in becquerel (Bq):

$$1\ \text{becquerel (Bq)} = \frac{1\ \text{decay}}{\text{s}}, \tag{3.3}$$

L. Cerrito, *Radiation and Detectors*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-53181-6_3

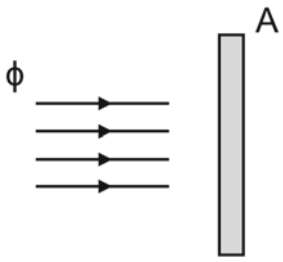


Fig. 3.1 Flux of particles incident onto a surface A

Table 3.1 Activity of common radioactive sources

Source	Activity (Bq)
1 banana	15
1 kg of coffee	1,000
1 adult human (75 kg)	4,900
1 smoke detector (americium)	30,000
1 kg uranium	25 million

with the curie also sometimes found:

$$1 \text{ curie (Ci)} = 3.7 \times 10^{10} \text{ Bq}. \tag{3.4}$$

These are both called units of disintegration, with the curie originally defined as the activity of one gram of pure radium-226. The activity of a short list of materials is shown in Table 3.1, as an example. For instance, one banana shows an activity of 15 Bq, non negligible for a widespread consumer fruit, but a rather small value when compared to 1 kg of coffee (1000 Bq).

The radioactive decay law is a simple exponential expression that describes the time change of the average number of decay centres in a radioactive substance. In a decay process characterised by a constant probability λ, starting with N decay centres, the number of decays ΔN that occurred after a time ΔT, which also corresponds to the reduction of decay centres by an amount ΔN, can be written as:

$$\Delta N = -\lambda N \Delta T. \tag{3.5}$$

For infinitesimal time intervals (3.5) is $\mathrm{d}N = -\lambda N \mathrm{d}t$, which solved by simple integration gives:

$$N(t) = N_0\, e^{-\lambda t} = N_0\, e^{-t/\tau}, \tag{3.6}$$

where N_0 is the number of decay centers at time $t = 0$ and $\tau = 1/\lambda$. The exponential decrease of the number of decay centres is governed by the constant τ, called the *mean lifetime* (in seconds), while λ is measured in inverse seconds. The value of τ

represents the time required for the number of decay centres to drop by a factor e. A similar quantity which is often used is the *half-life*, that is the time required to half the sample; it is easy to see by re-arranging (3.6) that the half-life is equal to $\tau \ln 2$, or approximately 0.7τ.

3.2 Radiological Units

To the extent of measuring the effects of radiation, including those on biological tissue, the atomic and nuclear excitation and ionisation are the primary factors to consider. Dosimetric units are chosen to quantify either the amount of ionisation that occurs, or the amount of energy deposited in the material.

Ionisation based units:
The *Ion dose* D_I (also called *Exposure*) is defined as the charge Q liberated in a volume V of air, per unit density ρ_{air}:

$$D_I = \frac{Q}{V\rho_{air}} \qquad \text{(exposure)}. \tag{3.7}$$

The unit of measurement for the ion dose D_I is the roentgen (R), and it corresponds to:

$$1\ \text{R} = 2.58 \times 10^{-4}\ \text{C/kg air}. \tag{3.8}$$

The roentgen is actually a legacy unit defined operationally as the charge liberated by X-rays in one kilogram of air at standard pressure. It is therefore mostly used and still routinely employed in environments where X-rays and gamma rays are to be found, such as in research and energy installations (see for instance the type of radiation that might be found in a particle physics research laboratory such as Fermilab, shown in Table 3.2), and values are often given in milli-roentgen (mR). The exposure rate at time t is the ratio:

$$\dot{D}_I(t) = \frac{\mathrm{d}D_I}{\mathrm{d}t} \qquad \text{(exposure rate)}, \tag{3.9}$$

Table 3.2 Sources of radiation that might be found in a particle research laboratory such as Fermilab (*source* Fermi National Accelerator Laboratory, USA)

Environment	α	β	Muon	γ-/X-rays	Neutrons
Accelerators (except Cockcroft-Walton)		•	•	•	•
Klystrons, Septa, Beam Separators, and Cockcroft-Walton				•	
Activation Products		•		•	
Radioactive Sources	•	•		•	•

which can itself be used to express the exposure occurring in a time interval between t_0 and t_1:

$$D_I = \int_{t_0}^{t_1} \dot{D}_I(t)\mathrm{d}t. \tag{3.10}$$

There is a simple formula that allows for the estimate of the gamma exposure rate $\dot{D}_I(t)$ in air at a distance d from a given emitter, under simplifying assumptions:

- the source can be approximated by a point source
- only photons originated directly from the source contribute to the rate, and not those from scattering in surrounding material

The exposure rate under such assumptions is given by:

$$\dot{D}_I(t) = \Gamma \frac{\alpha}{d^2}, \tag{3.11}$$

where α is the activity and Γ is the exposure rate constant characterising the radionuclide of interest. Table 3.3 lists a number of values for Γ calculated for gamma-ray energy greater than 10 keV.

Example: Estimate the exposure rate when working at an average distance of 50 cm from a source of ^{22}Na with an activity of 3.7 MBq.

The exposure rate is given by (3.11) as:

$$\dot{D}_I(t) = \Gamma \frac{\alpha}{d^2} = 3.8 \times 10^{-2} \frac{(\mathrm{R\ m^2})}{(\mathrm{h\ MBq})} \frac{3.7\ \mathrm{MBq}}{0.25\ \mathrm{m^2}} = 0.56\ \mathrm{R/h} \tag{3.12}$$

Energy based units:

More robust definitions of radiation are based on the energy deposited in the material encountered. The *Energy dose* (also called the *Absorbed dose*) is defined as the mean energy (W) absorbed by a material of density ρ per unit volume V (or also energy imparted to the material per unit mass):

$$D = \frac{W}{V\rho}. \tag{3.13}$$

Table 3.3 Unshielded gamma-ray exposure constant for gamma-ray energy greater than 10 keV. Values from [1] have been converted in units of exposure

Nuclide	$\Gamma((\mathrm{R\ m^2})/(\mathrm{h\ MBq}))$
^{57}Co	4.4×10^{-3}
^{60}Co	3.9×10^{-2}
^{137}Cs	1.1×10^{-2}
^{131}I	8.2×10^{-3}
^{22}Na	3.8×10^{-2}
^{226}Ra	4.4×10^{-4}

In the SI, the unit for the energy dose measurements is the called gray (Gy):

$$1\ \text{Gy} = 1\ \text{J/kg} \qquad \text{(absorbed dose)}. \tag{3.14}$$

The absorbed dose is a rigorous and measurable physical quantity and is used for all types of radiation. The gray is found, particularly in the United States, alongside a different unit called the *rad*:

$$1\ \text{rad} = 100\ \text{erg/g} = 10^2 \times 10^{-4}\ \text{J/kg} = 10^{-2}\text{J/kg} = 0.01\ \text{Gy}. \tag{3.15}$$

Unlike the roentgen, the gray applies to all types of radiation and all types of materials. For typical gamma and X-rays, 1 roentgen corresponds to approximately 0.01 Gy or 1 rad of absorbed dose. Other quantities might be defined for specific situations, for instance the *kerma*, K is defined to express the energy per unit mass transferred only by uncharged particles (e.g. photons and neutrons) to all charged particles in a material of unit mass [3].

Neither the ion dose, nor the absorbed dose comprehensively account for the potential effects of radiation on the human body.

Example: Determine a relationship between the energy dose and the ion dose of 1 R in air, assuming that 33.7 eV is the mean energy necessary to form an electron-ion pair.

Since the electron charge is 1.6×10^{-19} C, an ion dose of 1 R (from (3.8)) corresponds to a number N_{pair} of electron-ion pairs per kilogram of air equal to:

$$N_{pair} = \frac{2.58 \times 10^{-4}\ \text{C/kg} - \text{air}}{1.6 \times 10^{-19}\ \text{C}} \simeq 1.6 \times 10^{15}/\text{kg} - \text{air}. \tag{3.16}$$

Given the energy of 33.7 eV per pair, the total energy needed to liberate N_{pair} electron-ions is therefore:

$$D|_{1R} = 1.6^{15} \times 33.7\ \text{eV/kg} - \text{air} \tag{3.17}$$

$$= 5.4 \times 10^{16}\ \text{eV/kg} - \text{air} \simeq 0.9 \times 10^{-2}\ \text{Gy} = 0.9\ \text{rad}. \tag{3.18}$$

Therefore for typical γ and X-rays, 1 R corresponds to approximately 1 rad or 0.01 Gy.

Radiological protection units:

The sole information on the energy released in a biological material is not sufficient to indicate the effect of the radiation on that material. In order to give a more comprehensive indication, the *Equivalent radiation dose* has been introduced in relation to the effects of radiation on the human body. It is defined as:

$$H_T = w_R\ D_T, \tag{3.19}$$

where D_T is the absorbed dose in the volume of a specific biological tissue or organ (T), and w_R is a weighting coefficient that depends on the type of radiation R. The weighting factor w_R classifies the biological hazard of each type of radiation. If more types of radiation R are involved, the equivalent radiation dose is the sum of the equivalent dose of each radiation type:

$$H_T = \sum_R w_R \, D_{T,R} \qquad \text{(equivalent dose)}. \tag{3.20}$$

The SI unit for the equivalent dose is called sievert (Sv), while a legacy unit is the *roentgen equivalent man* (rem), with the correspondence of 1 Sv=100 rem. Typical radiation weighting factors (see Table 3.4) are 1 for electrons and photons, 2 for protons and charged pions, 20 for α, fission fragments and heavy ions, and 2–20 for neutrons depending on their energy, as shown in Fig. 3.2.

In addition to the weighting factors for the type of radiation, different biological tissues respond differently to the same amount and type of radiation. Therefore, tissue weighting factors w_T are also used to account for the effect of radiation on different parts of the body. The *Effective dose* is defined from the equivalent dose as the weighted sum of tissue equivalent doses:

$$E = \sum_T w_T \, H_T = \sum_T w_T \sum_R w_R \, D_R \qquad \text{(effective dose)}, \tag{3.21}$$

Table 3.4 Recommended radiation weighting factors [2]

Radiation type	Weighting factor, w_R
Photons	1
Electrons and muons	1
Protons and charged pions	2
α, fission fragments and heavy ions	20
Neutrons	see Fig. 3.2

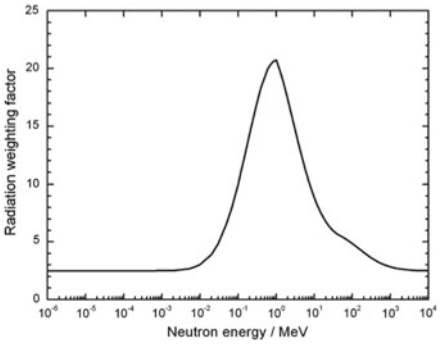


Fig. 3.2 Radiation weighting factor, w_R, for neutrons as a function of the neutron energy. Reprinted with permission from [2]

Table 3.5 Recommended tissue weighting factors [2]

Tissue	w_T	$\sum_T w_T$
Bone-marrow (red), Colon, Lung, Stomach,	0.12	0.72
Breast, Remainder tissues[a]		
Gonads	0.08	0.08
Bladder, Oesophagus, Liver, Thyroid	0.04	0.16
Bone surface, Brain, Salivary glands, Skin	0.01	0.04
	Total	1.00

[a]Remainder tissues: Adrenals, Extrathoracic (ET) region, Gall bladder, Heart, Kidneys, Lymphatic nodes, Muscle, Oral mucosa, Pancreas, Prostate (male), Small intestine, Spleen, Thymus, Uterus/cervix (female)

where the sums run over all tissues T involved and all radiation types R. The unit of effective dose is also the sievert (Sv), corresponding to $\mathrm{J\,kg^{-1}}$, as for the absorbed dose and equivalent dose. The recommended tissues weighting factors are listed in Table 3.5 [2].

It should be noted that a number of assumptions form the basis for the practical application of the radiological protection units defined here. In particular, the average dose over a volume of tissue is used to represent an inherently stochastic process, and it is assumed that for sufficiently low doses the summing of weighted mean doses in different organs, as well as the addition of doses due to internal and external exposure, bear a correlation of sufficient accuracy to a stochastic description of the particular interactions and effects taking place.

Example: Neutrinos interact in the human body through the process:

$$\nu_e + N \rightarrow e^- + N', \tag{3.22}$$

where the radiation damage is then caused by the outgoing electron. Estimate the annual dose for a person of 80 kg of mass under the assumptions that 50% of the neutrino energy is transferred to the electron, the typical energy of a solar neutrino is 100 keV and the number of interactions of solar neutrinos in the human body is approximately 100 per year.

Given the above, the energy transferred to the person's body in one year is therefore:

$$\Delta \mathrm{E} = 50\ \mathrm{keV} \times 100 = 5\ \mathrm{MeV/year}, \tag{3.23}$$

and since $1\ \mathrm{MeV} = 1.6 \times 10^{-13}$ the absorbed dose is:

$$\mathrm{D} = \frac{\Delta \mathrm{E}}{\Delta \mathrm{V}\rho} = \frac{5\ \mathrm{MeV/year}}{80\ \mathrm{kg}} = \frac{5 \times 1.6 \times 10^{-13}\mathrm{J/year}}{80\ \mathrm{kg}} \simeq 10^{-14}\mathrm{Gy/year}. \tag{3.24}$$

which is indeed a very small amount. The fact that the neutrino's electric charge is zero, and their practically null mass, implies that the probability of neutrino interaction with matter is extremely small, and therefore neutrinos are always ignored in radiological physics.

3.3 Radiation Doses in Life and the Environment

Specific operational quantities are usually employed to estimate internal and external exposure to radiation, and to provide an estimate of the radioprotection quantities connected to an exposure of people. This is because the equivalent and effective dose cannot be measured directly in a tissue. The definition of such operational quantities goes beyond the scope of this book but it can be found in [2]. The natural radiation sources that are responsible for an average non-negligible radiation dose to humans are called sources of background radiation. They are grouped into three categories: Radiation from Space, Terrestrial Radiation and Radionuclides present in the body.

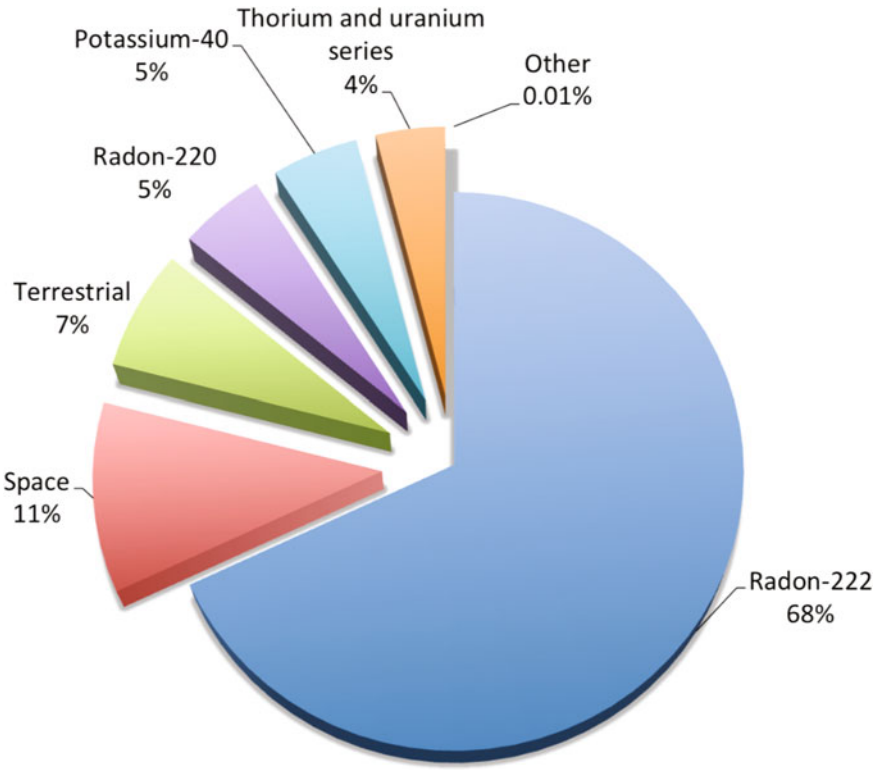


Fig. 3.3 Dose distribution of ubiquitous background sources in the United States. Data from [4]

The radiation from space is the result of the interaction between cosmic rays and the Earth's atmosphere, producing radionuclides which include ^{3}H, ^{7}Be, ^{14}C and ^{22}Na. The terrestrial radiation is due to the primordial radionuclides which are still present in the rocks, and include mainly the decay products of uranium and thorium. Radionuclides present in the body trace their origin to both cosmic and terrestrial sources which then enter the body through food, water and air. The most important is radon gas which is constantly inhaled, and which is directly linked to the uranium and thorium concentrations in the soil.

The average total dose (that is the *effective dose*, i.e. the amount of energy per kg absorbed by the body) per person due to the natural sources of background radiation described above is 2–3 mSv per year, depending on the geographical location on Earth. About 80% of this is due to radon gas, with the remaining split between other terrestrial radiation and cosmic rays. The pie chart in Fig. 3.3 shows a breakdown of the ubiquitous background in the United States. Other locations will display moderately different compositions.

Aside from the radiation background to which everyone is exposed, by far the major common human-made source of radiation is from medical applications. For instance, a Computed-Tomography (CT) scan to the abdoment and pelvis is responsible for a typical effective dose of 10 mSv, or 20 mSv if repeated with contrast material; this is the equivalent of 7 years of background radiation. Alternatively, flying once a month on an intra-European route amounts to about 10–15% of the yearly effective dose from background radiation. The typical, estimated effective

Table 3.6 Estimates of effective dose from common sources. Data from [5–9]

Source	Annual effective dose (mSv)	Fraction of annual dose (%)
Cosmic radiation	*0.39*	*16*
Terrestrial radiation (external)	*0.48*	*20*
Inhalation	*1.26*	*52*
Ingestion	*0.29*	*12*
Total natural sources	2.4 ([a])	100
Smoke alarms	0.00007	0.003
Radioluminous wristwatch containing ^{3}H	0.01	0.4
Flight on route Frankfurth–Rome	0.003–0.006	0.1–0.3
Flight on route Frankfurth–San Francisco	0.045–0.110	2–5
Commercial flight at 33,000 feet (lat. 60° N)	0.0024–0.0036/h	0.10–0.15/h
Peaceful nuclear power activities	$\leq$0.028	$\leq$1.2
CT scan (Abdomen and Pelvis)	10–15	400–600
Dental X-ray	0.005	0.2
PET/CT scan	25	1000
US average annual from medical exposure	3	125

([a]): Worldwide average

dose absorbed by each person from some common sources is shown in Table 3.6. The total yearly average dose per person, including man-made doses such as medical procedures is thought to be about 6.1 mSv.

3.4 Biological Effects of Radiation

The basis of our knowledge on the human effects of radiation include the medical history of early radiation workers and the survivors of atomic bombs, the impact on populations of radiation accidents and the monitoring of radiation therapy. The energy released by radiation into a biological tissue might cause tissue damage, for instance UV radiation from the sun can cause skin damage. However, gamma and X-rays can penetrate the skin. Without going into the details of the biological reaction of cells to radiation, some of which can be found in [2], it suffices mentioning that radiation can cause direct ionization of DNA or ionization of other molecules. Some of this damage can be repaired by the body and for those cases there might not be visible effects. Factors that influence the biological effects of radiation include: (i) the total dose, (ii) the dose rate, (iii) the type of radiation, (iv) the area and the location of the body exposed, (v) the cell sensitivity and (vi) the individual's sensitivity. The International Commission on Radiation Units and Measurements (ICRU) classifies the effects of radiation exposure into two groups:

- deterministic effects (i.e. tissue reactions) due to death or malfunction of cells following a high radiation dose;
- stochastic effects (i.e. cancer and heritable mutation and deseases) due to permanent damage in the DNA of cells.

Tissue or organ reactions occur in the presence of acute radiation doses, either as a single dose or from protracted exposure (although the latter tends to be less damaging), and generally take place above a certain dose threshold. The tissue reactions can occur either early or much later after irradiation and the severity of the damage, caused ultimately by the death of a critical number of cells, increases with the dose. The data available so far indicates that for doses up to around 100 mGy there is no visible cell functioning impairment in any tissue. The actual tissue thresholds vary; for instance, Table 3.7 lists threshold doses for a number of tissues. Red bone marrow, reproductive cells and the lens of the eye, or otherwise tissues with rapidly dividing cells, are the most sensitive tissues. Death occurs for all-body absorbed doses of 1 Gy or more, although if the same dose is delivered with a rate of 0.2 Gy per hour the thresholds may increase by 50%.

Cancer and other heritable effects, unlike tissue reactions, can be induced even at very low doses. Therefore there is no threshold, although the frequency of such stochastic effects is expected to increase as the dose is increased. Studies show clear evidence of higher cancer rates from doses of about 100 mSv or more, while the occurence of heritable deseases has a less clear evidence base (coming mostly from experimental animals) but it is nevertheless expected to be a factor. For doses

Table 3.7 Estimates of threshold for tissue reactions to single and multiple radiation doses. Data from [10–13]

Tissue	Effect	Threshold dose (single) (Gy)	Threshold dose rate (Gy/y)	Threshold total dose (multiple) (Gy)
Embryo	Malformations	0.1	–	–
Testes	Temporary sterility	0.15	0.4	–
Bone marrow	Low hematopoiesis	0.5	>0.4	–
Lens of eye	Opacities	0.5–2.0	>0.1	5.0
Ovaries	Permanent sterility	2.5–6.0	>0.2	6.0
Testes	Permanent sterility	3.5–6.0	2.0	–
Lens of eye	Cataract	5.0	>0.15	>8
All body	Death	1–2	–	–

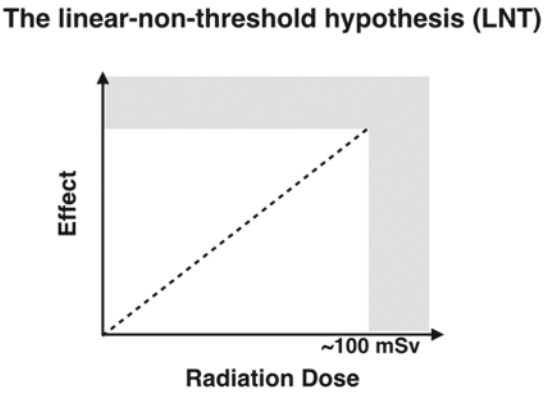


Fig. 3.4 The linear-non-threshold hypothesis for the stochastic effects of low-doses and low dose-rates radiation

below 100 mSv it is plausible that the incidence of stochastic effects is directly proportional to the radiation dose, and this forms the so-called *linear-non-threshold* (LNT) hypothesis (see Fig. 3.4), which is largely used for the drafting of radiological protection recommendations. The assumption relates, linearly, the stochastic effects with low-doses and low dose-rates radiation. The risk coefficients for stochastic effects due to the exposure to low dose and low dose rate radiation are reported in [2] to be around 5.5% per 1 Sv of dose. This value varies somewhat by a few units amongst different populations and risk models [14]. This means an incidence of cancer of 0.05% per 10 mSv of exposure. The occurence of heritable effects is estimated to have a risk coefficient of around 0.2% per 1 Sv.

3.5 Recommended Dose Limits

All humans are exposed to ionising radiation in one way or another, from natural or artificial sources. Dose limits are recommendations based on our knowledge of the detrimental impact of excessive radiation on biological materials. The International Commission on Radiological Protection (ICRP) recommends the separation of the exposures into three categories: occupational, medical, and the general public. Similarly, the exposed individuals are classed as workers, patients, and members of the public. A dose higher than 100 mSv as a single exposure event or accumulated over one year causes a significant risk of cancer and an increased probability of tissue reactions. Therefore, this is considered a reference level of exposure acceptable only for extreme situations. The recommended dose limit for workers in planned situations are expressed as 20 mSv per year, averaged over a 5 year period, with the

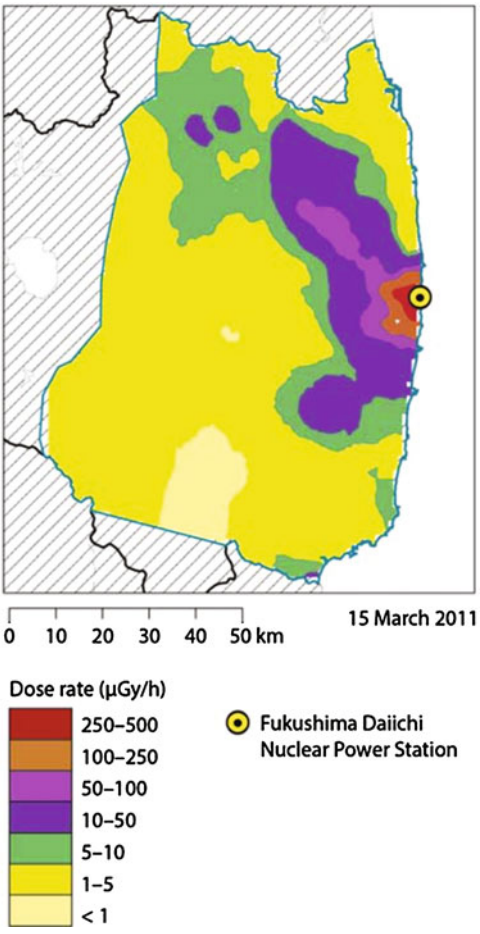


Fig. 3.5 Estimation of the weighted absorbed dose rates for a large mammal in the vicinity of the Fukushima-Daiichi power station.
Courtesy © UNSCEAR, reprinted from [15] with permission

additional provision of not exceeding 50 mSv in any single year. The recommended effective dose limit for the general public in planned situations is of 1 mSv per year. These do not apply in emergency exposure situations where a higher exposure might be necessary.

Accidents do however occur; in March 2011 the Fukhushima-Daiichi nuclear power installation in Japan suffered major damage following a magnitude 9.0 earthquake and tsunami. About 78,000 people living within a 20 km radius of the power plant were immediately evacuated as the plant leaked in the atmosphere primarily iodine-131 and caesium-137 (two largely beta and gamma emitters with half-life of 8 days and 137 years, respectively). The Fig. 3.5 shows an estimation of the weighted absorbed dose rates to a large mammal in the vicinity of the Fukushima-Daiichi power station, four days after the accident. The peak value is in the range 250–500 μGy/h, and if a similar level were considered for the public then the dose limit of 1 mSv above background, which is the recommended threshold for the general public, would be reached within two hours of such exposure.

On average, it is estimated that members of the public within the 20 km evacuation zone were exposed to less than 10 mSv during or before the evacuation [15]. People living in the city of Fukushima have been estimated to have received on average 4 mSv in the year following the accident. Only about about 170 workers were thought to have received doses of more than 100 mSv over the first 19 months after the accident. No radiation-related death or acute events of tissue reactions were recorded in the general public or workers.

Glossary

Absorbed dose See Energy dose

Activity For a radioactive source is defined as the number of disintegrations per second and is measured (in the SI) in becquerel (Bq)

$$1\ \text{Bq} = 1\ \text{decay/s}$$

Average total dose See Effective dose

Becquerel (Bq) The SI unit of measurement for radioactive disintegration (activity), corresponding to one disintegration per second

Curie (Ci) Non-SI unit of measurement for radioactive disintegration (activity), corresponding to 3.7×10^{10} Bq

Deterministic effects (also Tissue reactions) Reaction of biological tissue to an acute radiation dose, owing to the death or malfunction of a critical number of cells

Effective dose (E) (also Average total dose) Tissue-weighted sum of tissue equivalent doses:

$$E = \sum_T w_T \, H_T,$$

where w_T is the tissue-weighting coefficients and H_T is the per-tissue equivalent dose. Effective dose is measured in sievert (Sv)

Energy dose (also Absorbed dose) Mean energy (W) absorbed from radiation by a material per unit volume (V) per unit density (ρ), or otherwise per unit mass m

$$D = \frac{W}{V\rho} = \frac{W}{m}.$$

The energy dose is the fundamental dose quantity and in the SI is measured in gray (Gy), defined as 1 Gy=J/kg

Equivalent radiation dose (H_T) The SI unit of measurement for the absorbed dose in a specific tissue or organ. It is defined by means of tissue-specific weight coefficients w_T:

$$H_T = w_T \, D_T,$$

and measured in sieverts (Sv)

Exposure See Ion dose

Exposure rate Time derivative of the exposure

$$\dot{D}_I(t) = dD_I/dt$$

Flux, ϕ The number of particles (N_{inc}) incident onto a surface per unit area (A) per unit time interval (ΔT):

$$\phi = \frac{N_{inc}}{A \, \Delta T}$$

Fluence rate See Flux

Fluence, Φ The number of particles (N_{inc}) incident onto a surface per unit area (A)

$$\Phi = \frac{N_{inc}}{A}$$

Gray (Gy) The SI unit of measurement for the absorbed dose, corresponding to 1 J/kg

Half-life In a radioactive decay, the time necessary for the number of decay centres to drop by a factor of two. Equals the mean lifetime τ multiplied by the natural logarithm of 2, or approximately 0.7τ

Ion dose, D_I (also Exposure) The electric charge Q liberated by radiation in a volume V of air per unit density ρ_{air}

$$D_I = \frac{Q}{V\rho_{air}}.$$

The unit of measurement is the roentgen (R)

LNT (linear-non-threshold) hypothesis Assumption that relates linearly the incidence of stochastic effects in biological tissue to the amount of absorbed effective dose

Mean lifetime, τ In a radioactive decay, the time necessary for the number of decay centres to drop by a factor of e. Related to the decay probability λ by $\tau = 1/\lambda$

Rad Non-SI unit of measurement for the absorbed dose, corresponding to 0.01 Gy

Radioactive decay law Exponential mathematical expression that describes the number of decay centres $N(t)$ at time t following a disintegration process whose characteristic probability is λ:

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

where N_0 is the number of decay centers at time $t = 0$

Radon Radioactive gas accounting for nearly 80% of the ubiquitous radioactive background dose to humans

Roentgen, R Non-SI unit of measurement for the Ion dose, or Exposure,
1 R=2.58 10^{-4} C/kg-air

Rem (roentgen equivalent man) Non-SI unit of measurement for the equivalent radiation dose. 100 rem = 1 Sv

Sievert (Sv) The SI unit of measurement for the equivalent radiation dose. One Sv corresponds to 1 J/kg

Stochastic effects Reaction of biological tissue to a radiation dose causing permanent damage to the DNA of cells. Such reaction can take the form of cancer or heritable mutations

Threshold dose Effective radiation dose above which tissue reactions are expected to take place in biological tissue

Tissue reactions See Deterministic effects

Ubiquitous background radiation Natural radiation responsible for non-negligible dose to all humans. It consists of cosmic and terrestrial external radiation and radionuclides present in the body

References

1. L.M. Unger, D.K. Trubey, Speficic Gamma-Ray Dose Constants for Nuclides Important to Dosimetry and Radiological Assessment, Oak Ridge Laboratory ORNL/RSIC-45/R1 (1982), http://web.ornl.gov/info/reports/1982/3445603573381.pdf. Accessed 14 Sep 2015
2. International Commission on Radiation Units and Measurements (ICRP), The 2007 Recommendations of the International Commission on Radiological Protection. ICRP Publication 103. Ann. ICRP **37** (2–4)
3. A.F. Herbert, *Introduction to Radiological Physics and Radiation Dosimetry* (Wiley-VCH, New York, 1986)
4. National Council on Radiation Protection and Measurements (NCRP), Publication 160, Fig 1.1
5. Exposures of the Public and Workers from Various Sources of Radiation, Annex B to Volume I of the 2008 United Nations Scientific Committee on the Effects of Atomic Radiation Report to the General Assembly, Sources and Effects of Ionizing Radiation, Table 12, available on the UNSCEAR 2008 Report vol. I, www.unscear.org/unscear/en/publications/2008_1.html. Accessed 27 Aug 2015
6. European Commission, Radiation Protection 88, Recommendations for the Implementation of Title VII of the European Basic Safety Standards Directive (BSS) Concerning Significant Increase in Exposure Due to Natural Radiation Sources (1997)
7. R. Smith-Bindman et al., Arch. Int. Med. **169**(22), 2078–2086 (2009). doi:10.1001/archinternmed.2009.427
8. A.L. Jones, W.B. Oatway, J.S. Hughes et al., Review of trends in the UK population dose. J. Radiol. Prot. **27**(4), 381–390 (2007)
9. F.A. Mettler Jr., B.R. Thomadsen, M. Bhargavan et al., Medical radiation exposure in the U.S. in 2006: preliminary results. Health Phys. **95**(5), 502–507 (2008)
10. International Commission on Radiation Units and Measurements (ICRP), Publication 103. Ann. ICRP **37** (2–4) (2007), Annex A, data from Table A.3.1
11. United Nations Scientific Committee on the Effects of Atomic Radiation. Sources, Effects and Risks of Ionizing Radiation, *Report to the General Assembly with Annexes* (United Nations, New York, 1988)
12. Non-stochastic effects of irradiation. ICRP Publication 41. Ann. ICRP **14**(3)
13. A.A. Edwards, D.C. Lloyd, Risk from deterministic effects of ionising radiation. Doc. NRPB **7**(3) (1996)
14. United Nations Scientific Committee on the Effects of Atomic Radiation, Effects of Ionizing Radiation, UNSCEAR 2006 Report to the General Assembly (2006)
15. United Nations Scientific Committee on the Effects of Atomic Radiation, UNSCEAR 2013 Report to the General Assembly, Scientific Annex A, Levels and effects of radiation exposure due to the nuclear accident after the 2011 great east-Japan earthquake and tsunami (2013), Figure F34

Chapter 4
Relativistic Kinematics and Collisions

Abstract Particles move at speeds that might easily approach the speed of light, and therefore classical mechanics is not sufficient to model their motion in space. Relativistic kinematics is needed to correctly account for, amongst many quantities, the time and length intervals, the momentum, and the energy of the particles. This chapter presents the basic equations needed to understand and perform simple calculations of relativistic kinematics. Additionally, collisions, which underpin radiation-matter interactions, are introduced in the context of special relativity. The invariant mass is discussed as a particle discovery tool, and the chapter ends with the calculation of Compton scattering and the formal definition of cross section.

4.1 Motion at Classical and Relativistic Speeds

At speeds much less than the speed of light in vacuum, which is approximately $c \simeq 300,000\,\text{km/s}$, the motion of an object in our three-dimensional space can be described by a three-dimensional position vector $\mathbf{x}$ with respect to a chosen reference frame S and expressed as a function of a time coordinate t, i.e.: $\mathbf{x}(t)$. The velocity, $\mathbf{v}(t) = \dot{\mathbf{x}}(t)$, and acceleration, $\mathbf{a}(t) = \ddot{\mathbf{x}}(t)$, are the single- and double-time derivatives of the position vector, respectively. The mass (or alternatively the momentum) is defined by Newton's second law,

$$\mathbf{a}(t) = \frac{1}{m}\mathbf{F}(t), \tag{4.1}$$

where m (i.e. the object's mass) gives the proportionality between the force to which the object is being subject to and its acceleration. The momentum is defined as:

$$\mathbf{p}(t) = m\mathbf{v}(t), \tag{4.2}$$

and with momentum the most general expression of Newton's second law can be expressed by,

$$\mathbf{F}(t) = \dot{\mathbf{p}}(t). \tag{4.3}$$

L. Cerrito, *Radiation and Detectors*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-53181-6_4

The motion of an object under the application of a force allows for the definition of the kinetic energy (T) associated with the motion of the mass m,

$$T = \frac{1}{2}m\mathbf{v}^2 = \frac{\mathbf{p}^2}{2m}, \tag{4.4}$$

where $\mathbf{v}^2$ indicates the scalar product $\mathbf{v} \cdot \mathbf{v} = v^2$ and therefore $\mathbf{p}^2$ and $\mathbf{v}^2$ are both scalars. In (4.4), and in what follows, the explicit time dependence has been dropped but it is implied.

In addition to the kinematics of a single object, the description of the collision between two objects of mass m_1 and m_2, respectively, can be treated using Newton's third law, from which it follows the conservation of the total momentum

$$\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}_1' + \mathbf{p}_2', \tag{4.5}$$

where the primed vectors indicate the momenta of m_1 and m_2 *after* the collision. If the collision is elastic, then the total kinetic energy is also conserved,

$$T_1 + T_2 = T_1' + T_2'. \tag{4.6}$$

All that has been summarised up to this point is valid for motion with non-relativistic speeds.

Relativistic Transformations

In the case of speeds which are non-negligible compared to c, the special relativity is needed to correctly describe the kinematics of the objects and their motion under the influence of a force. Elementary and composite particles are all seen to behave according to relativistic kinematics equations when moving at speeds near the speed of light. The vector parameter $\boldsymbol{\beta} = \mathbf{v}/c$ describes the ratio between the particle's velocity $\mathbf{v}$ to the speed of light and its modulus is bound between 0 (towards non-relativistic speeds) and 1 (towards relativistic speeds).

Time intervals can be significantly dilated when switching from one reference frame to another. If $\Delta\tau$ is a time interval elapsed in a coordinate system where a particle is at rest, the corresponding time interval Δt observed in a different reference frame which is moving with speed v with respect to the first reference frame is:

$$\Delta t = \frac{1}{1-\beta^2}\Delta\tau = \gamma\Delta\tau, \tag{4.7}$$

where $\gamma = 1/\sqrt{1-\beta^2}$ and is bound between 1 (non-relativistic speeds) and infinity (relativistic speeds). Hence, Δt is always larger than $\Delta\tau$ and this effect is called *time dilation*. This result may be obtained applying so-called *Lorentz* transformation of the position and time coordinates between two inertial frames [1, 2].

Lengths measured in reference frames which are in relative motion with each other, and where the motion is parallel to the measurement direction, are also affected. If

$\Delta\ell$ is the proper length, i.e. the length measured in the reference frame where the particle is at rest, then the length Δl measured in any other frame moving with respect to the first is:

$$\Delta l = \sqrt{1-\beta^2}\,\Delta\ell = \frac{1}{\gamma}\Delta\ell, \tag{4.8}$$

and so it is always shorter than the proper length, that is, lengths are contracted and this effect is called *length contraction*. The (4.7) and (4.8) are called Lorentz transformations of time intervals and lengths.

Example: In its rest frame a charged pion has a mean lifetime of 2.6×10^{-8} s. If the pion is moving with a relativistic γ equal to 70 with respect to the laboratory frame, determine the corresponding mean lifetime as measured in the laboratory. What is the mean path length of the pion before decaying?

With a relativistic $\gamma = 70$, the decay time of the pion in the laboratory frame is given by:

$$\Delta t = \gamma\,\Delta\tau = 70 \times 2.6 \times 10^{-8} = 1.8\,\mu\text{s}.$$

Notice how the travelling pion is seen to decay much later than its lifetime would imply. With $\gamma = 70$ the corresponding speed is $\beta = v/c = 0.99990$. The path length covered in 1.8 μs at such speed corresponds to:

$$\ell = 0.9999\,c\,\Delta t = 0.9999 \times 3 \times 10^8 \frac{\text{m}}{\text{s}} \times 1.8 \times 10^{-6}\text{s} \simeq 540\ \text{m} \tag{4.9}$$

The plot depicted in Fig. 4.1 shows γ as a function of β between 0 and 1.

In general, motion at speeds with $\beta < 0.3$ ($\gamma < 1.05$) can be described in the classical regime, while cases with larger β values need to be treated in the relativistic regime.

Not only lengths and time intervals but also mass, momentum, and energy need the proper relativistic description. Assuming that m_0 is the known mass of a particle at rest (called the *rest mass*), then the relativistic version of momentum in a moving reference frame is the following [3]:

$$\mathbf{p} = \frac{m_0\mathbf{v}}{\sqrt{1-\beta^2}} = \gamma m_0\mathbf{v} = m\mathbf{v} \qquad \text{(relativistic momentum)}. \tag{4.10}$$

The product γm_0 is called the relativistic mass m, and notice that depending on γ its value can grow to be infinitely large. This expression is sometimes interpreted by saying that as the speed of the particle increases, its "mass" (meaning the relativistic mass m) grows with it.

The relativistic kinetic energy (T) can be calculated with the work-energy theorem [2] and its expression is:

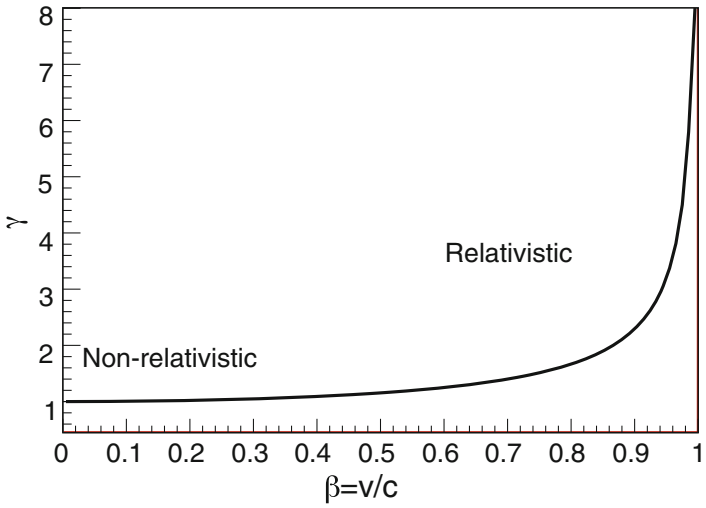


Fig. 4.1 Values of γ as a function of β, with an indication of the relativistic and non-relativistic regimes

$$T = \frac{m_0 c^2}{\sqrt{1-\beta^2}} - m_0 c^2 = (\gamma - 1) m_0 c^2 \qquad \text{(relativistic kinetic energy)}. \tag{4.11}$$

Notice that the first term of (4.11) grows with the speed of the particle and is unbound while the second term is a constant value. Also, by taking Taylor's expansion of $1/\sqrt{1-\beta^2}$ for small β values, the (4.11) is approximated by its classical counterpart of (4.4).

The (4.11) can also be rewritten by rearranging it and defining the first term on the right hand side as the total energy E,

$$E = T + m_0 c^2 = \frac{m_0 c^2}{\sqrt{1-\beta^2}} = \gamma m_0 c^2 = mc^2 \qquad \text{(relativistic total energy)}, \tag{4.12}$$

where the constant term $m_0 c^2$ is interpreted as the energy of the particle at rest, which is the energy when $T = 0$, and is therefore called the *rest energy*.

The (4.12) can be expressed directly as a function of momentum $\mathbf{p}$ rather than velocity $\mathbf{v}$. To do this, it is sufficient to square and invert (4.10):

$$p^2 = \frac{1}{1-\beta^2} m_0^2 v^2 = \frac{1}{1 - v^2/c^2} m_0^2 v^2 = \frac{c^2}{c^2 - v^2} m_0^2 v^2 \tag{4.13}$$

obtaining an expression for v^2/c^2:

$$p^2c^2 - v^2p^2 = c^2m_0^2v^2$$
$$v^2(p^2 + c^2m_0^2) = p^2c^2$$
$$\frac{v^2}{c^2} = \frac{p^2}{p^2 + m_0^2c^2},$$

which can be substituted into (4.12) to give:

$$E^2 = \frac{m_0^2c^4}{1 - \frac{p^2}{p^2+c^2m_0^2}} = (\mathbf{p}c)^2 + \left(m_0c^2\right)^2. \tag{4.14}$$

The equation obtained,

$$E = \sqrt{(\mathbf{p}c)^2 + \left(m_0c^2\right)^2} \qquad \text{(relativistic total energy)}, \tag{4.15}$$

describes once again the total energy as the sum of a rest energy term and a momentum term. The (4.15) can instructively be visualised by plotting the energy versus momentum. The plot in Fig. 4.2 shows the total energy of a proton, whose rest mass if 938 MeV/c^2, as a function of its momentum. Notice that the total energy equals the rest mass until the momentum approaches values corresponding to the energy of the rest mass. The kinetic energy shown in the plot derives from (4.11), indicating that the kinetic energy increases linearly, as long as it is small compared to the rest energy, then it approximates the total energy for values much larger than 1 GeV.

If the particle has no momentum, then its total energy is simply given by the rest energy:

$$E = m_0c^2. \tag{4.16}$$

If the particle has no mass, as is the case for the photon for example, then its energy is entirely given by the momentum component.

Example: A muon is travelling with an energy of 10 GeV. The muon lifetime is 2 μs and its mass is 105 MeV/c^2. Determine the length it travels before decaying.

With a total energy of 10 GeV the muon is travelling with a relativistic γ of:

$$\gamma = \frac{E}{m_0c^2} = \frac{10\,\text{GeV}}{0.105\,\text{GeV}} \simeq 95.$$

The lifetime in the laboratory frame is dilated to:

$$\Delta t = \gamma\Delta\tau = 95 \times 2\mu\text{s} = 190\,\mu\text{s},$$

and $\beta \simeq 1$. The path length is therefore:

$$\ell = c\Delta t = 3 \times 10^8\,\frac{\text{m}}{\text{s}} \times 190 \times 10^{-6}\,\text{s} = 57\,\text{km}$$

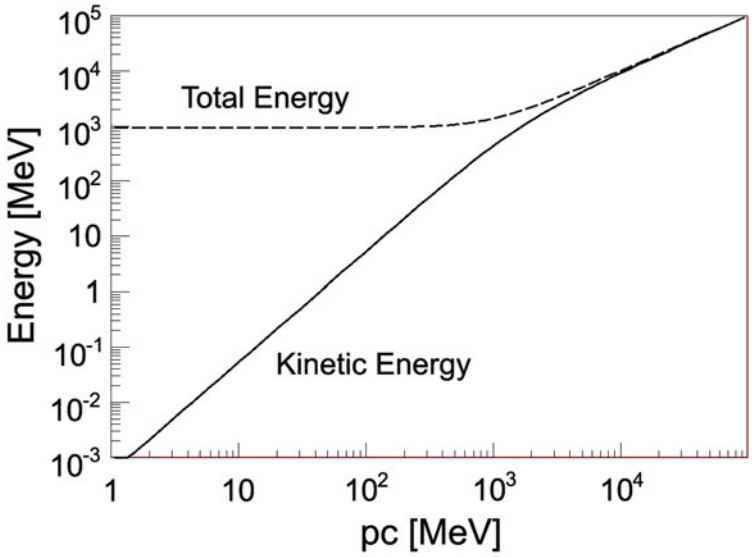


Fig. 4.2 Total and kinetic energy of a proton as a function of the proton momentum

Time dilation is the very reason why muons produced in collisions of cosmic rays with the Earth's atmosphere are able to travel as far as sea level, despite being produced at up to tens of kilometers of altitude.

Non-relativistic and Ultra-relativistic Limits

The relativistic relationships between energy, kinetic energy and momentum, shown in (4.11) and (4.12), can often be simplified owing to motion either in the non-relativistic case (for which $v << c$) or the ultra-relativistic condition, ($v \simeq c$). The non-relativistic approximation is applicable to particles of mass m_0 whose kinetic energy satisfies the condition:

$$T << m_0 c^2, \tag{4.17}$$

which can be derived from $v << c$ by writing, using Taylor's expansion of γ:

$$E = \gamma m_0 c^2 \simeq m_0 c^2 + \frac{1}{2} m_0 v^2. \tag{4.18}$$

This yields the classical expression for kinetic energy, $T = p^2/2m_0$.

The ultra-relativistic approximation is applicable to particles of mass m_0 whose kinetic energy satisfies:

$$T >> m_0 c^2, \tag{4.19}$$

which derives from (4.11) with $\gamma >> 1$. In this case, $E = T = pc$.

Four-Vectors and Natural Units

The relativistic equations shown so far can be significantly simplified by using four-vectors and natural units ($\hbar = 1$, $c = 1$). The notation:

$$\underline{q} = (E/c, \mathbf{p}) \tag{4.20}$$

indicates the *energy-momentum* four-vector of a particle. The first component is the energy, while the remaining three are the components of the momentum vector. The modulus of the energy-momentum vector is defined as:

$$\underline{q}^2 = \frac{E^2}{c^2} - |\mathbf{p}|^2, \tag{4.21}$$

and can be readily recognised from (4.15) as equal to $m_0^2 c^4$; this is indeed an invariant under Lorentz transformations. An additional operation, the dot product $(\underline{q}_1 \cdot \underline{q}_2)$ is useful in equations dealing with four-vectors, and is defined as

$$\underline{q}_1 \cdot \underline{q}_2 = \frac{E_1 E_2}{c^2} - \mathbf{p_1} \cdot \mathbf{p_2}, \tag{4.22}$$

where $\mathbf{p_1} \cdot \mathbf{p_2}$ is the ordinary dot product of the three-dimensional momentum vectors. It is easy to verify that the square of a four-vector is the dot product of the four-vector with itself.

Finally, when using natural units the following energy-momentum simplified relationships hold:

$$\begin{aligned} E &= \gamma m_0, \\ \underline{q} &= (E, \mathbf{p}), \\ \underline{q}^2 &= E^2 - |\mathbf{p}|^2 = m_0^2, \\ T &= E - m_0 = m_0(\gamma - 1). \end{aligned}$$

Relativistic Collisions

Consider the collision between particles a and b. The collision process can lead to a generic set of new particles:

$$a + b \rightarrow c + d + e + \ldots \tag{4.23}$$

this is called a particle *formation* process, where the new particles c, d, e etc. are formed from the collision. If the collision leads to the same particles as the initial state, then the collision is said to be *elastic*:

$$a + b \rightarrow a + b. \tag{4.24}$$

As in the case of non-relativistic collisions, energy and momentum are separately conserved in the collision process, that is they are the same before and after the collision if measured with respect to the same reference frame.

4.2 Mass of a Set of Particles

Consider a set of n particles with energies E_i and individual momenta $\mathbf{p}_i$ ($i = 1, \ldots, n$). The four-momentum ($\underline{P}$) vector of this system as a whole is built from the individual four-momentum vectors as:

$$\underline{P} = \sum_{i=1}^{n} \underline{p}_i = \left(\sum_{i=1}^{n} E_i, \sum_{i=1}^{n} \mathbf{p}_i \right), \tag{4.25}$$

i.e. the energy component is the algebraic sum of the energies of each particle, and the momentum component is the vector sum of the individual momenta. The quantity $\underline{P}^2$ is therefore an invariant under Lorentz transformations, and is called the invariant mass of the system of particles:

$$\underline{P}^2 = M^2 = \left(\sum_{i=1}^{n} E_i \right)^2 - \left(\sum_{i=1}^{n} \mathbf{p}_i \right)^2. \tag{4.26}$$

If the momenta are measured in the center of mass frame of this set of particles, for which by definition of center of mass $\sum_{i=1}^{n} \mathbf{p}_i = 0$, then the invariant mass is simply given by the sum of the energies:

$$M = \sum_{i=1}^{n} E_i \qquad \text{(in the center of mass frame)} \tag{4.27}$$

meaning that the invariant mass of a set of particles is also the total energy in the center of mass reference frame. Moreover, not only the invariant mass M is an invariant under Lorentz transformation, but it is also conserved in the production or decay of particles because total energy and momentum are separately conserved in reactions of production or decay.

Consider now two particles moving off in different directions as shown in Fig. 4.3. The square of the invariant mass is

$$M^2 = (E_1 + E_2)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2 \tag{4.28}$$

$$= E_1^2 + E_2^2 + 2E_1E_2 - (|\mathbf{p}_1|^2 + |\mathbf{p}_2|^2 + 2|\mathbf{p}_1||\mathbf{p}_2| cos\theta) \tag{4.29}$$

$$= m_1^2 + m_2^2 + 2E_1E_2 - 2|\mathbf{p}_1||\mathbf{p}_2| cos\theta, \tag{4.30}$$

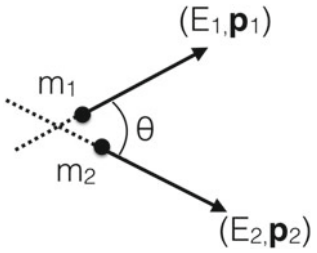


Fig. 4.3 Sketch of two particles moving away in different directions

notice how this is not equal to the simple sum of the masses of the two individual particles. An example of this situation is given by the decay process of the Z boson $Z \rightarrow e + e$. In this case, since particles one and two are both electrons $m_1 = m_2 = m_e$ and

$$M^2 = 2m_e^2 + 2E_1E_2 - 2|\mathbf{p}_1||\mathbf{p}_2|cos\theta, \tag{4.31}$$

and in the rest frame of the Z boson one obtains (either by setting $\theta = \pi$ in (4.31) or also $\mathbf{p}_1 + \mathbf{p}_2 = 0$ in (4.28), both due to momentum conservation),

$$M = 2E. \tag{4.32}$$

Therefore, in the rest frame of the Z boson, the electrons will fly off in opposite directions and with energy equals to half the boson mass.

Example: The Higgs boson might decay to two photons through the process:

$$H \rightarrow \gamma + \gamma. \tag{4.33}$$

The general reconstruction of the invariant mass of a set of two particles follows the steps from (4.28), however the photons have no rest mass, so $m_1 = m_2 = m_\gamma = 0$, and $E_{1,2} = |\mathbf{p}_{1,2}|$ therefore (4.28) becomes:

$$M_H^2 = 2E_{\gamma 1}E_{\gamma 2} - 2E_{\gamma 1}E_{\gamma 2}cos\theta = 2E_{\gamma 1}E_{\gamma 2}(1 - cos\theta) \tag{4.34}$$

$$= 4E_{\gamma 1}E_{\gamma 2}sin^2(\theta/2). \tag{4.35}$$

The reconstruction of the invariant mass of unstable particles through the kinematics of their decay products is the basis that led the discovery of numerous new particles. For instance, Fig. 4.4 shows the distribution of invariant mass of a large number of pairs of photons seen by the ATLAS experiment [4]. The upper panel reports the distribution overlaid to an exponential function, the typical functional dependence for unrelated photon pairs from background sources in these collisions. The peak at values of $M_{\gamma\gamma}$ around 125 GeV corresponds to the decay of Higgs bosons. A similar invariant mass reconstruction can also be performed when the final state contains more than two particles, as in the case of

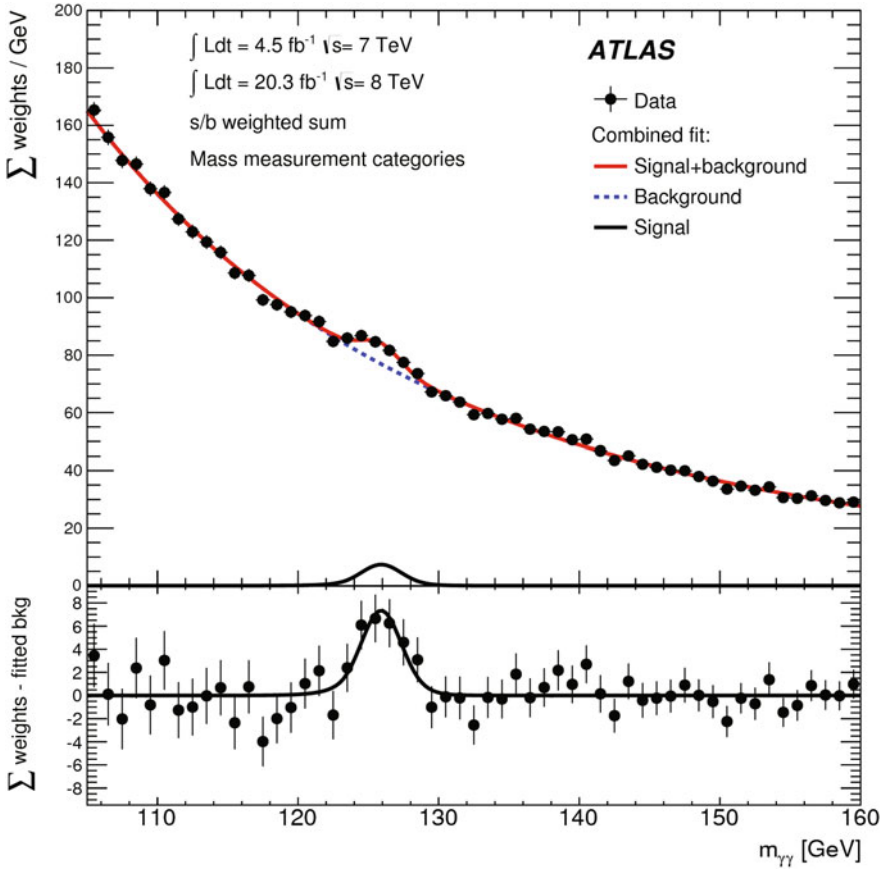


Fig. 4.4 The distribution of invariant mass of two photon candidates as seen by the ATLAS experiment. Reprinted from [4] with permission

$$H \rightarrow \ell + \ell + \ell' + \ell', \tag{4.36}$$

which has four final state particles. The invariant mass of the four leptons must equate the mass of the Higgs boson if the four leptons are the decay product of the Higgs. The distribution in Fig. 4.5 shows the invariant mass of four leptons as seen by the CMS experiment [5]. The peak at around 125 GeV corresponds to the Higgs boson.

4.3 Particle Formation in Collisions

The collision of two particles, particles a and b, can be used to create new particles by converting the energy available from the collision into mass.

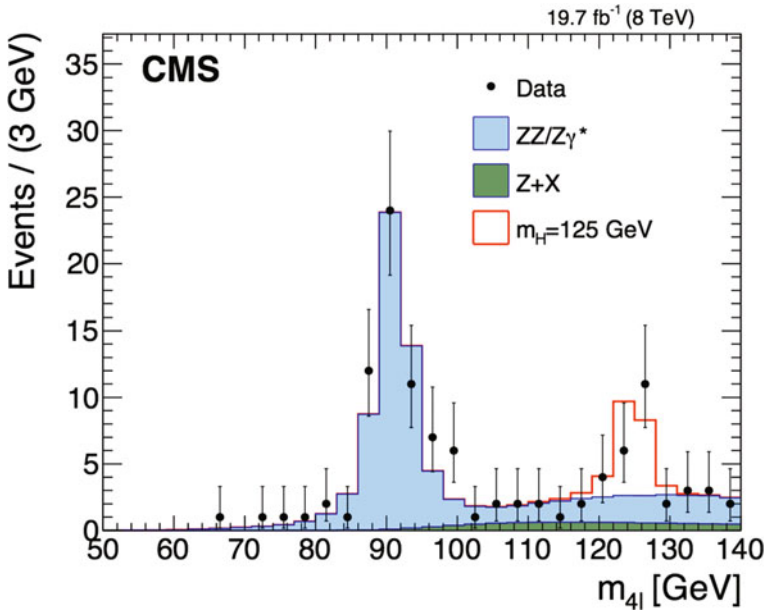


Fig. 4.5 The distribution of invariant mass of four lepton candidates as seen by the CMS experiment. Reprinted from [5] with permission

Fig. 4.6 Collision of two particles a and b, either with b stationary (*left*) or head-on (*right*)

In the *laboratory* (*lab*) frame, defined such that one of the particles is initially at rest ($|\mathbf{p}_b| \equiv 0$ and $E_b = m_b$, see the left diagram of Fig. 4.6), the four-vector of this set of particles is:

$$\underline{P} = (E_a + E_b; \mathbf{p}_a)\,, \tag{4.37}$$

and the invariant mass is:

$$M^2 = (E_a + E_b)^2 - |\mathbf{p}_a|^2 = E_a^2 + m_b^2 + 2E_a m_b - |\mathbf{p}_a|^2 = m_a^2 + m_b^2 + 2E_a m_b. \tag{4.38}$$

This is a simple result stating that the invariant mass of a and b, that is also the energy in the center of mass frame, is linearly proportional to the square root of the energy of a, E_a. For masses negligible compared to the energy of a ($E_a >> m_a,\ m_b$), then (4.38) simplifies to

$$M \simeq \sqrt{2E_a m_b}. \tag{4.39}$$

For example, for a proton with an energy of 200 GeV incident on a target proton at rest, the invariant mass or also the center of mass energy is equal to:

$$M = \sqrt{1 + 1 + 400}\,\text{GeV} \simeq 20.0\,\text{GeV}, \tag{4.40}$$

and notice that this is a very small amount compared to the moving proton energy. The implication is that much of the initial proton's energy is being retained as kinetic energy of the final state particles. This effect can be seen explicitly by writing the *threshold* energy for particle formation in the case of a moving particle a striking a stationary partible b. From (4.38) and considering that $\underline{P}^2$ is the energy in the center of mass of the final state particles:

$$E_{CM}^2 = (E_a + E_b)^2 - |\mathbf{p}_a|^2 = m_a^2 + m_b^2 + 2E_a m_b. \tag{4.41}$$

The energy E_{CM} is given by both the kinetic energy and the invariant mass of the final state particles:

$$E_{CM} = \sum (T_f + m_f), \tag{4.42}$$

but the minimum energy required for the production of particles with mass m_f in the final state is achieved when all these particles are at rest, i.e. $T_f = 0$. Therefore, at threshold, $E_{CM,th} = \sum m_f$ and substituting this into (4.41) gives

$$E_{th,a} = \frac{\left(\sum m_f\right)^2 - m_a^2 - m_b^2}{2m_b}. \tag{4.43}$$

This result is general; suppose to have the intention to produce antiprotons by colliding protons (with mass indicated by m_p). The process $p + p \to \bar{p}$ does not conserve particle numbers, but $p + p \to p + p + p + \bar{p}$ is possible as long as the initial moving proton energy is equal to at least the threshold energy of:

$$E_{th,p} = \frac{(4m_p)^2 - (2m_p)^2}{2m_p} = 6\,m_p, \tag{4.44}$$

which is 5.6 GeV. The antiproton was indeed discovered using a proton beam with an energy of 6.2 GeV. The plot in Fig. 4.7 shows the ratio of masses between the newly produced negative-charge particle and the proton, at the Bevatron accelerator in 1955 [6].

More efficient particle production can be achieved with a different approach, that is if the collision is operated colliding two particles with equal and opposite momenta (see the right diagram of Fig. 4.6). In this case the lab frame coincides with the center of mass frame (for which $\mathbf{p}_b + \mathbf{p}_a \equiv 0$ and therefore $\mathbf{p}_b = -\mathbf{p}_a$), then the four-vector of the two particle system is:

$$\underline{P} = (E_a + E_b; 0)\,, \tag{4.45}$$

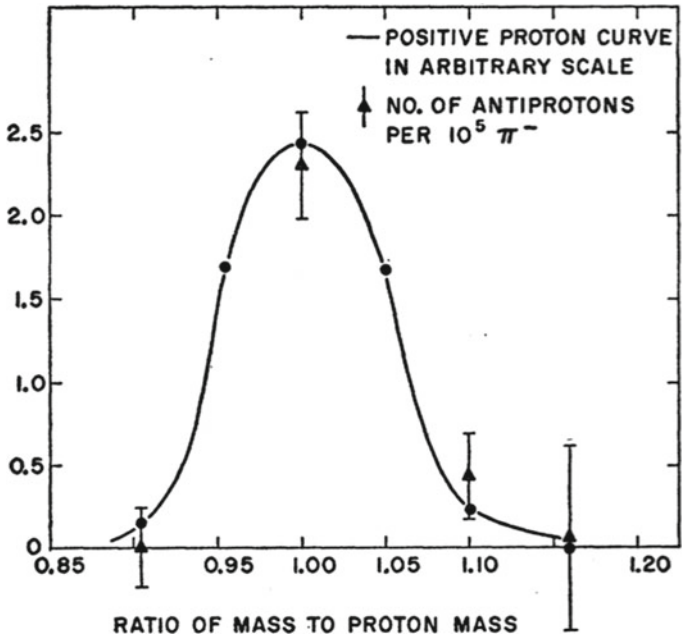


Fig. 4.7 Ratio of the antiproton to proton masses (*triangles*) and the experimental resolution (*solid curve*) at the Bevatron accelerator (1955). The vertical scale indicates the number of antiprotons produced per 10^5 negative pions. Reprinted from [6] with permission

and the invariant mass is:

$$M^2 = (E_a + E_b)^2 - 0 \Rightarrow M = E_a + E_b. \tag{4.46}$$

If the particles a and b have the same mass, or also if the initial energies are much larger than the particle masses, then $E_a = E_b = E^*$ and

$$M = 2E^*. \tag{4.47}$$

Considering now two protons with energy 200 GeV that are made to collide head-on, then the invariant mass is equal to:

$$M = 2 \times 200\,\text{GeV} = 400\,\text{GeV}, \tag{4.48}$$

and all of the initial protons energy is retained as energy in the center of mass.

4.4 Compton Scattering

The elastic collision of two particles is a process called two-body scattering. The Compton Scattering, described in Chap. 10 in the context of the interaction between photons and matter, is the name for the scattering of light off particles and a case of two body scattering. It is a useful example of kinematic calculation and is therefore introduced here as a worked exercise to determine the transferred energy between the photon and the electron. The diagram in Fig. 4.8 indicates an incoming photon colliding with a stationary electron. In treating this case, the time spent in the interaction between the two particles is considered small and is neglected. The four-momentum balance of the particles before and after the collision is

$$\underline{p}_{\gamma} + \underline{p}_{e} = \underline{p}_{\gamma'} + \underline{p}_{e'}. \tag{4.49}$$

The equation can be solved for the four-momentum of the electron after the collision

$$\underline{p}_{e'} = \underline{p}_{\gamma} + \underline{p}_{e} - \underline{p}_{\gamma'}, \tag{4.50}$$

and squaring (4.50) gives (since also $\underline{p}^2_{e'} = m_e$, $\underline{p}^2_{e} = m_e$, and $\underline{p}^2_{\gamma} = \underline{p}^2_{\gamma'} = 0$)

$$m_e = m_e + 2\underline{p}_{\gamma} \cdot \underline{p}_{e} - 2\underline{p}_{\gamma'} \cdot \underline{p}_{\gamma} - 2\underline{p}_{e} \cdot \underline{p}_{\gamma'} \tag{4.51}$$

In (4.51), it has been used the dot product of two four-vectors, already defined in (4.22).

The last three terms on the right hand side of (4.51) contain the multiplication of the energy components of the four-vectors and of the momenta, but $\mathbf{p}_e = 0$, $E_e = m_e$ as the initial electron is stationary and $|\mathbf{p}_{\gamma}| = E_{\gamma}$ so (4.51) simplifies to:

$$0 = E_{\gamma} E_e - \left(E_{\gamma'} E_{\gamma} - \mathbf{p}_{\gamma'} \cdot \mathbf{p}_{\gamma}\right) - E_e E_{\gamma'} \tag{4.52}$$

$$0 = E_{\gamma} E_e - \left(E_{\gamma'} E_{\gamma} - |\mathbf{p}_{\gamma}||\mathbf{p}_{\gamma'}| cos\theta\right) - E_e E_{\gamma'} \tag{4.53}$$

$$0 = E_{\gamma} E_e - \left(E_{\gamma'} E_{\gamma} - E_{\gamma'} E_{\gamma} cos\theta\right) - E_e E_{\gamma'} \tag{4.54}$$

$$0 = E_{\gamma} m_e - E_{\gamma'} E_{\gamma} (1 - cos\theta) - m_e E_{\gamma'} \tag{4.55}$$

$$E_{\gamma} m_e = E_{\gamma'} \left[m_e + E_{\gamma}(1 - cos\theta)\right] \tag{4.56}$$

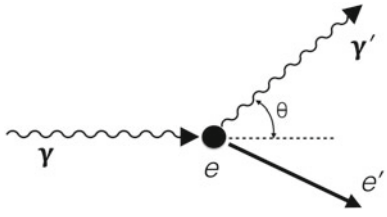


Fig. 4.8 Diagram of Compton scattering

This leads to a remarkable equation for the outgoing photons energy:

$$E_{\gamma'} = \frac{E_\gamma}{\left[1 + \frac{E_\gamma}{m_e}(1 - cos\theta)\right]}, \tag{4.57}$$

which relates $E_{\gamma'}$ to the incoming photon energy E_γ and the scattering angle θ. This process produces therefore variable energy electrons and photons. The minimum and maximum values for the outgoing photon energy are immediately derived from (4.57). The maximum energy is only obtained when $(1 - cos\theta)$ is smallest, that is for $\theta = 0$, where the photon is completely forward and effectively there is no interaction (the photon energy is unchanged compared to its initial value). For meaningful scattering, the final photon energy is always lower than the initial value, and the minimum value is when $(1 - cos\theta)$ is largest, that is for $\theta = \pi$. In the latter case, the photon is emitted in the direction opposite to the incoming photon, and its energy is:

$$E_{\gamma'}^{min} = \frac{E_\gamma}{\left[1 + \frac{2E_\gamma}{m_e}\right]}. \tag{4.58}$$

The corresponding maximum energy *transferred* to the outgoing electron, E_T^{max}, is

$$E_T^{max} = E_\gamma - E_{\gamma'}^{min} = E_\gamma - \frac{E_\gamma}{\left[1 + \frac{2E_\gamma}{m_e}\right]} = \frac{E_\gamma}{\frac{m_e}{2E_\gamma} + 1} \tag{4.59}$$

and this maximum energy value of the outgoing electron is called the *Compton Edge*.

Example: Two photons are produced through the annihilation of an electron and a positron, which are initially at rest. One of the photons goes on to interact with a nearby electron via Compton Scattering. What is the maximum energy of the scattered electron?

The production of the two photons occurs through the process:

$$e^+ + e^- \rightarrow \gamma + \gamma. \tag{4.60}$$

Notice that although the production of a single photon through the process $e^+ + e^- \rightarrow \gamma$ is allowed according to conservation of all quantum numbers, this process does not conserve simultaneously energy and momentum. To see this, write the energy and momentum conservation in the center of mass frame of the e^+ and e^- pair:

$$E_{e-} + E_{e+} = E_\gamma \tag{4.61}$$

$$\mathbf{0} = \mathbf{p}_\gamma \tag{4.62}$$

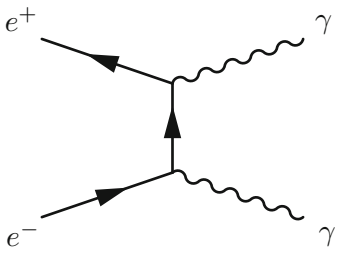


Fig. 4.9 Leading order diagram of electron-positron annihilation

which implies $E_\gamma = |\mathbf{p}_\gamma| = 0$ and therefore the first equation cannot be satisfied, given that the minimum energy of each electron is the electron rest mass.

The electron-positron annihilation to two photons (see Fig. 4.9) is, on the other hand, allowed by energy and momentum conservation. If the electron and positron are initially at rest ($E_{e-} = m_e$, $E_{e+} = m_e$):

$$E_{e-} + E_{e+} = 2m_e = E_{\gamma 1} + E_{\gamma 2} \tag{4.63}$$

$$\mathbf{0} = \mathbf{p}_{\gamma 1} + \mathbf{p}_{\gamma 2} \tag{4.64}$$

therefore $E_{\gamma 1} = E_{\gamma 2} = E_\gamma$ and finally:

$$E_\gamma = m_e = 511\,\text{keV} \tag{4.65}$$

When photons of this energy collide with a nearby electron via Compton scattering, the maximum energy transferred is given by inserting $E_\gamma = m_e$ into (4.59), leading to

$$E_T^{max} = \frac{2}{3} m_e = 341\,\text{keV}. \tag{4.66}$$

4.5 Cross Section

The cross section, generally indicated by the greek symbol σ, of a physical process is a quantity whose definition is derived from, and intimately connected to, the probability of occurrence of the interaction that leads to that process. For example, consider a collimated beam of individual particles hitting a target. The typical processes that can derive from the interaction between the beam and the target are particle creation or elastic scattering. The probability of occurrence of the interaction, which can be modelled through the probability of the removal of beam particles as they traverse the target, is calculable and it will normally depend on the number of incident particles and on the size of the target. Therefore, given a flux of particles a, ϕ_a (the flux has been introduced in Sect. 3.1) hitting a target b in a time interval dt, the number of scattered particles can be written as:

$$dN_{sc} = A_b \phi_a dt, \tag{4.67}$$

where A_b is a coefficient which describes the target. Because flux is measured in inverse length squared and inverse time, then A_b in the SI is measured in square meters. Also, since $dN_{sc}/(\phi dt)$ is the number of scattered particles divided by the number of incident particles and is therefore a probability, then also A_b can be interpreted as a probability. If the target is composed of many (N_b) centers, then $A_b = \sigma_b N_b = \sigma_b n_b dV$, where n_b is the density of the target centers per unit volume. The quantity σ_b is called the *cross section* of the target centers, it has units of area and is measured more conveniently in barn

$$1 \text{ barn (b)} = 10^{-28} \text{ m}^2 \tag{4.68}$$

Substituting the target coefficient A_b back into (4.67) gives:

$$\mathrm{d}N_{sc} = \sigma_b n_b \phi_a \mathrm{d}V \mathrm{d}t \tag{4.69}$$

Often the interaction between particles leads to the scattering of final state particles which depends on the emission angle. For this purpose, the *differential* cross section is defined for scattering into a direction at mean polar angle θ and azimuthal angle ϕ with respect to the incident beam direction, and within a solid angle $\mathrm{d}\Omega$:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{\phi_a} \frac{\mathrm{d}N_{sc}}{\mathrm{d}\Omega}, \tag{4.70}$$

which is the number of particles per unit incident flux which are scattered in a particular direction within a solid angle $\mathrm{d}\Omega$. The total cross section is recovered by integration:

$$\sigma = \int_0^{\Omega} \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \mathrm{d}\Omega. \tag{4.71}$$

Finally, the number of particles *removed* from the incident beam by the scattering process as the beam passes through the target centers, b, is obtained by writing the flux term as $\phi_a \mathrm{d}V\mathrm{d}t = N_{inc}\mathrm{d}x$, where $\mathrm{d}x$ is an element of distance in the direction of motion of the incident particles, then (4.69) becomes:

$$\mathrm{d}N_{sc} = -\mathrm{d}N_{inc} = \sigma_b n_b N_{inc} \mathrm{d}x, \tag{4.72}$$

which integrated leads to

$$N_{inc}(x) = N_{0,inc} e^{-x/(\sigma_b n_b)}. \tag{4.73}$$

The number of particles surviving the transit through the target centers is therefore given by the initial number of particles multiplied by an exponential term. The constant in the exponent:

$$\lambda_{coll} = \frac{1}{\sigma_b n_b}, \tag{4.74}$$

is given the name of *collision length* or also *mean free path*, and is the distance over which the incident number of particles diminishes by a factor of e.

Glossary

Barn The SI unit of measurement for the cross section, $1\ \text{b} = 10^{-28}\,\text{m}^2$

Collision length The distance over which the number of incident particles interacting with scattering centers, uniformly distributed in a volume, are reduced by a factor of $e \simeq 2.7$

Compton scattering Elastic scattering of light off particles

Cross section Effective area or scattering probability off a target center, measured in barn

Elastic collision (relativistic) A collision that leads to the same particles as the initial state

Energy-momentum vector A four-vector containing the energy and momentum components of a particle

$$\underline{q} = (E/c, \mathbf{p})$$

Invariant mass For a set of particles, the same as *Mass of a set of particles*

Kinetic Energy (classical) Scalar quantity (T) defined by:

$$T = \frac{1}{2} m \mathbf{v}^2 = \frac{\mathbf{p}^2}{2m},$$

where m is the mass, $\mathbf{v}$ is velocity and $\mathbf{p}$ is momentum

Length contraction Relativistic effect seeing lengths change between two reference frames moving with relative speed v parallel to the measurement direction. The equation:

$$\Delta l = \frac{1}{\gamma} \Delta \ell$$

relates the length $\Delta\ell$ measured at rest with the length $\Delta\ell$ measured in the moving frame

Mass of a set of particles (also Invariant mass) The Lorentz-invariant quantity M obtained from the modulus of the energy-momentum vector of the particles system:

$$\underline{P}^2 = M^2 = \left(\sum_{i=1}^{n} E_i\right)^2 - \left(\sum_{i=1}^{n} \mathbf{p}_i\right)^2,$$

where i is the particle index and E_i, $\mathbf{p}_i$ are the particles energy and momentum, respectively

Mean free path See *Collision length*

Momentum (classical) Three-dimensional vector obtained from the product of mass m and velocity $\mathbf{v}$:

$$\mathbf{p} = m\mathbf{v}$$

Relativistic β The ratio

$$\beta = v/c,$$

between the an object speed, v, and the speed of light, c, and is bound to values between 0 and 1

Relativistic γ The quantity defined as

$$\gamma = 1/\sqrt{1 - \boldsymbol{\beta}^2},$$

taking values between 1 and infinity (ultra-relativistic)

Relativistic kinetic energy For a particle moving with relativistic γ, the kinetic energy is given by the expression

$$T = (\gamma - 1)\, m_0 c^2,$$

where m_0 is the rest mass and c is the speed of light

Relativistic momentum For a particle moving with relativistic γ, the momentum is given by the product:

$$\mathbf{p} = \gamma m_0 \mathbf{v},$$

where m_0 is the rest mass and $\mathbf{v}$ is the velocity

Relativistic total energy For a particle moving with relativistic γ, the total energy is given by the expression

$$E = T + m_0c^2 = \gamma m_0c^2 = \sqrt{(\mathbf{p}c)^2 + \left(m_0c^2\right)^2}$$

where m_0 is the rest mass, c is the speed of light, $\mathbf{p}$ the relativistic momentum

Rest energy The energy of a particle at rest, also written as

$$E = m_0c^2$$

Rest mass The ratio between the energy of a particle at rest and the constant factor c^2

Time dilation Relativistic effect seeing time intervals change between two reference frames moving with relative speed v. The equation:

$$\Delta t = \gamma \Delta \tau$$

relates the time interval $\Delta\tau$ measured at rest with the time interval Δt seen by a frame moving with relativistic speed γ

References

1. R. Hagedorn, *Relativistic Kinematics: A Guide to the Kinematic Problems of High-Energy Physics* (Benjamin, New York, 1964)
2. H.D. Young, R.A. Freedman, *University Physics with Modern Physics*, 13th edn. (Addison-Wesley, Reading, 2013)
3. W.G.V. Rosser, *Introductory Relativity* (Butterworths, London, 1967)
4. ATLAS Collaboration, Phys. Rev. D. **90**, 052004 (2014) doi:10.1103/PhysRevD.90.052004
5. CMS Collaboration, J. High Energy Phys. **04**, 005 (2016). doi:10.1007/JHEP04(2016)005
6. O. Chamberlain, E. Segrè, C. Wiegand, T. Ypsilantis, Phys. Rev. **100**, 947 (1955)

Chapter 5
Elements of Accelerator Physics

Abstract The physics and technology of accelerating charges is a fascinating field. The ingenuity and refinement behind much of today's accelerators is astonishing. This chapter introduces some of the elements and the language of accelerator physics, from the principles of accelerating structures to the ideas behind beam transport and focussing.

5.1 Cockcroft–Walton and Van de Graaff Accelerators

Between 1906 and 1909, E. Rutherford working with H. Geiger and E. Marsden, showed that α particles accelerated towards gold nuclei can be backscattered by 180° [1], and this could not be explained by the Thomsons model of atoms. Rutherford used α particles as a particle *beam*, with energy of some million electronvolts. However, he could not control this beam, and had the idea of artificially accelerating charged particles to be used as an experimental tool.

In 1928, Rutherford encouraged J.D. Cockcroft to design an accelerator and in 1932, only four years later, the first artificially-accelerated particle beam was produced. Cockcroft utilised protons of 0.4 MeV to split a Lithium nucleus [2, 3] and that machine was the basis for the *Cockcroft–Walton* generator. The working principle is to use an electric field between two terminals to accelerate particles. The kinetic energy ε imparted is

$$\varepsilon = q\Delta V, \tag{5.1}$$

where q is the particle's electric charge and ΔV is the potential difference between the terminals. In order to have a sufficiently high accelerating voltage, Cockcroft used a rectifier circuit (see Fig. 5.1 for an example), whose baseline design is still used today. The particle source was a hydrogen discharge tube, and the accelerator consisted of an evacuated glass tube, with a Lithium foil arranged as the target. The calculation of the energy acquired by the particles is easy, owing to the definition of electronvolt. A charge of $1e$ accelerated through a potential difference of 0.4 MV acquires an energy of 0.4 MeV. A diagram reproducing one of the original Cockcroft–Walton accelerators is shown in Fig. 5.2. The main limitation of this device is the

L. Cerrito, *Radiation and Detectors*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-53181-6_5

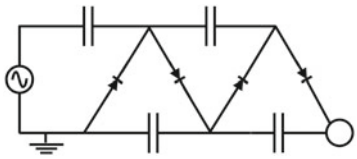


Fig. 5.1 A two-stage rectifier circuit forming the basis for Cockcroft–Walton accelerators

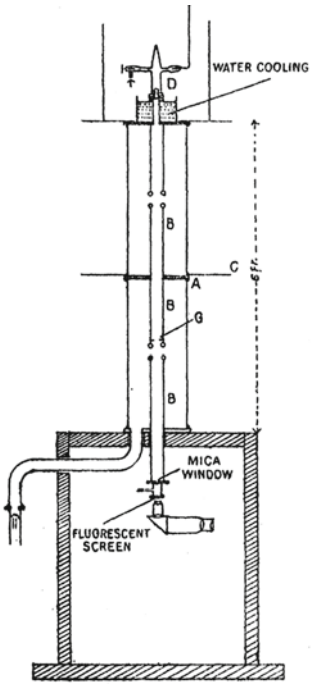


Fig. 5.2 Schematics of one of the original Cockcroft–Walton accelerators. Reprinted from [3] with permission

discharge due to high voltage breakdowns, which places a maximum voltage of about 1 MV, depending on the pressure and type of gas. The Cockcroft–Walton accelerators are still used today, typically as the initial stage of more complex accelerator chains; for instance a 750 kV Cockcroft–Walton accelerator was part of the Fermi National Accelerator Laboratory (FNAL or Fermilab) Tevatron pre-accelerator chain until the collider decommissioning in 2012.

In the early 1930s, the *Van de Graaff* accelerators were also invented [4]. These are electrostatic machines which create a high voltage by mechanical transport of charges. Electric charge is transported by a moving belt and accumulated on a spherical hollow conductor (see Fig. 5.3 for a diagram). The electric field created is then used to accelerate ions towards a target. The terminal potential can reach 12–15 MV using SF_6 gas at high pressure to suppress discharge. In order to increase the effective accelerating field, Van de Graaff accelerators have been used with the so-called *tandem* principle [5]. A Tandem machine consists of two Van de Graaff accelerators

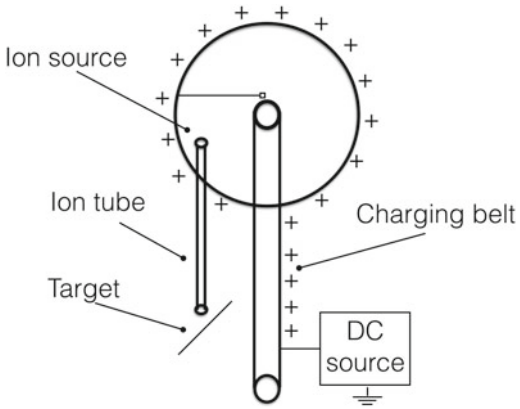


Fig. 5.3 Schematic representation of a Van de Graaff accelerator

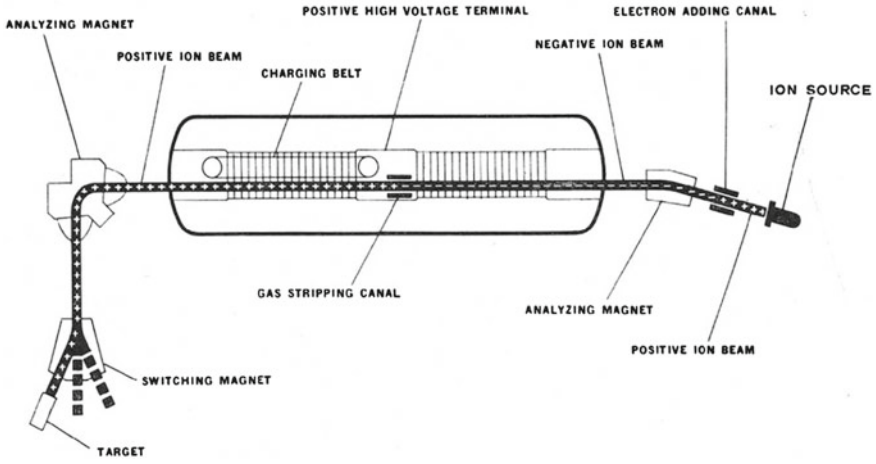


Fig. 5.4 Schematic representation of a Tandem Van der Graaff. Reprinted from [5] with permission

positioned in line one after the other (see Fig. 5.4). Maximum electrostatic potential is located at the center of the tandem, so that negative ions from one end are accelerated towards the other side. As the ions pass through the center they go through an electron stripping target and emerge as positive ions. With positive charge, they are now accelerated further towards the ground potential end of the machine. In practice, the tandem principle allows the accelerating field to be applied twice, so that modern tandem machines can reach potentials of around 28 MV. The photograph in Fig. 5.5 shows a tandem Van de Graaff at the US Brookhaven National Laboratory, completed in 1970 and used until 2012 as part of the Relativistic Heavy Ion Collider ion source. The facility consists of two 15 million volt electrostatic accelerators, each about 24 m long, aligned end-to-end.

Fig. 5.5 The Brookhaven National Laboratory's Tandem Van de Graaff. © Courtesy of Brookhaven National Laboratory. Printed from [6] with permission

5.2 Linear and Radio Frequency Accelerators

In 1928, R. Widerøe had the idea of applying an accelerating voltage to a particle beam several times in sequence using an oscillating potential. The scheme is shown in Fig. 5.6. The particles travelling from the source are subject to an accelerating electric field when travelling between one cavity and the next, and the electric field is reversed by the time the particles reach the following gap.

The particles are shielded from seeing the decelerating field whilst travelling within the cavities, so that the energy acquired after n cavities is:

$$E_n = n \times q \times U_0 \sin\Psi_s, \tag{5.2}$$

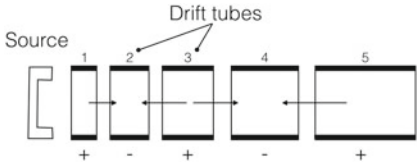


Fig. 5.6 Schematic representation of a radio frequency linac

where q is the accelerated charge, U_0 is the peak voltage of the radio frequency (RF) system and Ψ_s is the synchronous phase between the particle and the RF. In order to avoid the decelerating part of the oscillating voltage, the particles need to be shielded for a time $T_{RF}/2$, where T_{RF} is the period of the RF system. For particles moving with speed v_n inside the n-th tube, the length of the n-th tube is therefore:

Fig. 5.7 Inside view of the Alvarez-type LINAC at the Fermi National Accelerator Laboratory [7]. Credit Fermilab

$$\ell_n = v_n \frac{T_{RF}}{2} = \frac{\beta \lambda_{RF}}{2}, \tag{5.3}$$

where $\beta = v_n/c$ and λ_{RF} is the free space wavelength of the RF. In the non-relativistic regime, the speed is related to the kinetic energy such that:

$$v_n = \sqrt{\frac{2E_n}{m}}, \tag{5.4}$$

which by substitution into (5.3) gives:

$$\ell_i = \sqrt{\frac{2E_n}{m}} \frac{T_{RF}}{2} = \sqrt{\frac{2nqU_{\sin}\Psi_s}{m}} \frac{1}{2\nu_{RF}} = \frac{1}{\nu_{RF}} \sqrt{\frac{nqU_0 \sin\Psi_s}{2m}}, \tag{5.5}$$

where ν_{RF} is the frequency of the RF system. Notice that the synchronous condition in such a structure requires that the length of the cavities grows proportionally to the square root of n, while in the relativistic regime the length of the tubes becomes constant.

The idea of RF accelerators was taken up again much later, in the mid-1940s, by L.W. Alvarez who designed the first proton linear accelerator (LINAC) at the University of California. Alvarez surrounded the structure with a vessel to keep the RF as a standing wave and to reduce the amount of energy radiated when using the high frequencies needed by faster particles. Alvaretz-type linac RF accelerators are still today the standard for proton and ion pre-accelerators, with energies between 50 and 200 MeV (corresponding for protons to β between 0.2 and 0.57) and with frequencies of around 200 MHz. The photograph in Fig. 5.7 shows the drift tubes of the Alvaretz-type low-energy LINAC at FNAL.

The diagram in Fig. 5.8 shows the accelerating structure of the new CERN Linac 4, which will deliver an hydrogen ion beam of 160 MeV. Further details on RF linear accelerators can be found in [8].

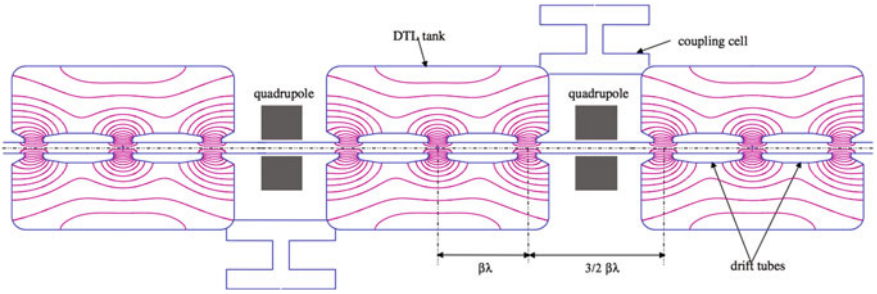


Fig. 5.8 Scheme of a Cell-Coupled Drift Tube Linac at CERN's Linac 4. Reprinted from [9] with permission

5.3 Cyclotrons and Betatrons

The working principle for accelerating charges in a *cyclotron* consists of *bending* a LINAC, that is sending the charges on a spiral trajectory while applying repeatedly the same accelerating electric field. A diagram is shown in Fig. 5.9. The space between the semi-circles is the accelerating gap, while a constant magnetic field is coming out of the page and perpendicular to it and generates a centripetal Lorentz force. The electric charge is therefore accelerated by an electric field at each passage through the gap.

The basic equations of motion for this system, assuming a classical regime, are the Lorentz force ($\mathbf{F} = q\mathbf{v} \times \mathbf{B}$) and the condition for circular motion, such that

$$q v B = \frac{m v^2}{R}, \tag{5.6}$$

where q, m are the particle electric charge and mass, respectively, and $\mathbf{v}$ is the particle velocity which is perpendicular to the magnetic field $\mathbf{B}$. Solving (5.6) for the product of B and R gives:

$$B\,R = \frac{m v^2}{q v} = \frac{p}{q}, \tag{5.7}$$

so one can see that the particle trajectory has an increasing radius for increasing momentum p. Moreover, since the velocity is related to the period T through

$$v = \frac{2\pi R}{T}, \tag{5.8}$$

then (5.6) becomes

$$B\,R = \frac{2\pi R}{T}\frac{m}{q}, \tag{5.9}$$

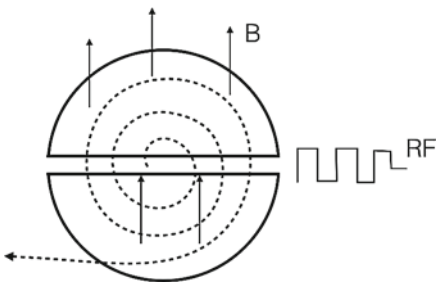


Fig. 5.9 Schematic representation of a cyclotron. The B field is oriented perpendicular to and coming out of the page

hence

$$\frac{Bq}{m} = \omega_z, \tag{5.10}$$

where ω_z is the revolution angular frequency. The (5.10) shows that for non-relativistic particles the radio frequency for such a machine is independent on the particle momentum and therefore $\omega_z =$ constant. In addition, the inverse proportionality between ω_z and the mass also works only for non-relativistic particles. On the other hand, machines for which $\omega_z \neq$ constant are called *synchrocyclotrons*. It is also clear that large momentum requires huge magnetic fields, and therefore large magnets. Cyclotrons are used as proton and ion accelerators for energies up to about 60 MeV. It is easy to check that a proton with kinetic energy of 60 MeV has $\beta \simeq 0.3$ and is therefore at the boundary of where relativistic effects become important. Synchrocyclotrons are used for accelerating protons to higher energies, typically several hundred MeV.

A *betatron* is a machine optimised for electrons and developed in 1940 at the University of Illinois. The idea is to produce an accelerating electric field through a time-varying magnetic field in a device sketched in Fig. 5.10. A magnetic field B guides the particles around a circular path, while a time-varying magnetic field B_a generates an accelerating electric field.

The condition for circular orbit is:

$$\frac{mv^2}{R} = evB \Rightarrow p = eBR, \tag{5.11}$$

where R is the radius of the orbit, e the electron charge and v is the speed of the electrons. Notice that (5.11), expressed in terms of the momentum p, is valid also for relativistic particles.

The induced electric field E that accelerates the electrons is given by:

$$2\pi RE = -\frac{\mathrm{d}\phi(B_a)}{\mathrm{d}t}, \tag{5.12}$$

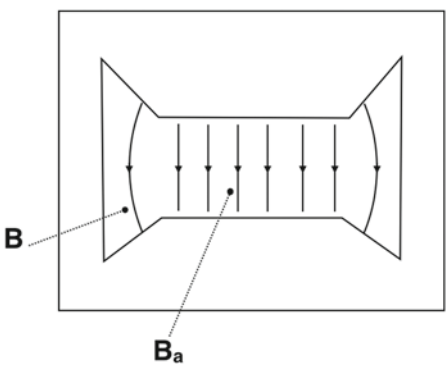


Fig. 5.10 Schematic representation of a betatron

where $\phi(B_a)$ indicates the flux of B_a through the surface bounded by the electron's orbit. On the other hand, the flux can be written as $\phi(B_a) = \pi R^2 B_a$ and therefore:

$$\frac{\mathrm{d}\phi(B_a)}{\mathrm{d}t} = \pi R^2 \frac{\mathrm{d}B_a}{\mathrm{d}t}. \tag{5.13}$$

Combining (5.12) and (5.13) yields:

$$2\pi R\, E(t) = -\pi R^2 \frac{\mathrm{d}B_a}{\mathrm{d}t}, \tag{5.14}$$

$$\Rightarrow \quad E(t) = -\frac{R}{2}\frac{\mathrm{d}B_a}{\mathrm{d}t}. \tag{5.15}$$

On the other hand, using Newton's second law:

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -eE(t), \tag{5.16}$$

and substituting (5.15) in (5.16) yields:

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{1}{2}eR\frac{\mathrm{d}B_a}{\mathrm{d}t}, \tag{5.17}$$

hence:

$$p = \frac{1}{2}eR\, B_a. \tag{5.18}$$

The (5.18) is a relationship between the momentum and the accelerating magnetic field. Comparing (5.18) with the condition for circular orbit of (5.11), gives a characteristic relationship of betatrons between the bending and the accelerating magnetic fields:

$$B = 0.5\, B_a. \tag{5.19}$$

Moreover, notice that increasing speeds require higher magnetic fields; these machines can be used for accelerating electrons to energies of 300–500 MeV.

5.4 Synchrotrons and Colliders

Synchrotrons are accelerator machines which overcome much of the limitations of the earlier accelerators. All the most powerful accelerators of the last decades (e.g. the Tevatron at the Fermi National Accelerator Laboratory, the Relativistic Heavy Ion Collider at Brookhaven National Laboratory, the Hadron-Electron Ring Accelerator at the Deutsches Elektronen-Synchrotron laboratory, the Large Hadron

Collider at the European Organisation for Nuclear Research, just to name a few) are based on the synchrotron principle. In synchrotrons, the bending magnetic field is increased as the particle's velocity increases. There is no massive magnet but a ring of individual magnets. The acceleration is provided by electric fields within a cylindrical resonator. Finally, the RF is synchronised with the particles as they speed along the ring. Designing a synchrotron involves therefore three main tasks: (i) define a circular orbit for the moving charges; (ii) place magnets on that orbit to guide and focus the particles; (iii) provide the acceleration while the particles move on the designed orbit. Colliders are machines based on the synchrotron principles which store and collide two beams circulating in opposite directions for the purpose of subatomic research.

Synchrotrons have also given their name to a type of radiation (*synchrotron radiation*) which is a well known by-product of accelerating and bending electric charges. The power emitted by a circulating beam through synchrotron radiation is given, in SI units, by [10]:

$$P_S = \frac{e^2 c}{6\pi\varepsilon_0}\frac{1}{R^2}\left(\frac{E}{m_0 c^2}\right)^4, \tag{5.20}$$

where e is the electron charge, m_0, E are the mass and energy of the accelerated particles, respectively, and R is the radius of the orbit. Notice that in (5.20) the emitted power is inversely proportional to the fourth-power of the particle's mass, and is therefore vastly different between light particles, such as electrons, and heavy particles, such as protons ($m_p \simeq 2000 \times m_e$). Moreover, considering the energy dependence in the numerator of (5.20), the power increases with γ^4, and therefore in increasing the energy it might be necessary to increase the design radius to limit the radiated power within reasonable values. The energy loss from high energy electrons, per turn, in a circular machine is:

$$\Delta E = \frac{e}{3\varepsilon_0 R}\left(\frac{E}{m_0 c^2}\right)^4, \tag{5.21}$$

and a rather useful form to calculate values from (5.21) is:

$$\Delta E\,(\text{keV}) = 88.5 \times \frac{E\,(\text{GeV})^4}{R\,(\text{m})}. \tag{5.22}$$

Finally, the frequency of the emitted light is given by:

$$\omega_S = \frac{3c\gamma^3}{2R} \tag{5.23}$$

Synchrotron light is extensively used for materials, biological, and medical research and synchrotrons are specifically built today as facilities for such research. Intense beams of synchrotron light can be produced highly polarised with narrow

Fig. 5.11 Aerial view of the UK Diamond Light Source. © Diamond Light Source, printed with permission from [11]

collimation and in a broad range of wavelengths. The Fig. 5.11 shows the Diamond Synchrotron Light accelerator, a synchrotron facility with a circumference of 562 m where a 3.0 GeV circulating beam produces synchrotron light. The Diamond Light Source is currently in operation at Harwell Science Centre in the United Kingdom.

5.5 Beam Transport

In a storage ring such as that of the Large Hadron Collider, a single proton moving at nearly the speed of light and circulating for several hours (which is the typical beam lifetime before the store is dumped) will travel a distance in the machine corresponding to approximately 10^{10}–10^{11} km. To give an idea of such distance, this corresponds to many times the distance between the Sun and Neptune (4.5×10^{9} km). Therefore, the particles in the beam must be guided on a well defined orbit (the *design orbit*) and focussed to keep each trajectory within the vacuum chamber. The basic idea in order to keep a proton on a circular orbit is to use a Lorentz force:

$$F_L = q\,(\mathbf{v}\times\mathbf{B})\,. \tag{5.24}$$

Note that bending protons, with speeds approaching c, by using only an electric field would require electric fields far exceeding one million volts per meter, and therefore this is not a viable option. With $\mathbf{B}$ and $\mathbf{v}$ perpendicular to each other, and $q = e$, (5.24) is simply $F_L = evB$. The condition for circular motion is satisfied by a centripetal force:

$$F_c = \frac{\gamma m_0 v^2}{\rho}, \tag{5.25}$$

where m_0 is the rest mass, ρ is the radius of the circular orbit and $p = \gamma m_0 v$ is the particle momentum. Combining (5.24) and (5.25) yields

$$B\rho = \frac{p}{e}. \tag{5.26}$$

The product $B\rho$ is called the beam rigidity, and notice that particles circulating with higher momentum require stronger B in order for them to maintain the same orbit.

The bending of particles is typically achieved placing dipole magnets on the orbit. However, the magnetic field required by modern accelerators is only achieved using coils with large currents. For instance, the LHC uses dipole magnets operating with a current of 11,850 A. The power loss due to such high currents can be calculated according to Ohms law, and it is prohibitively high, such that only superconducting materials can effectively be used for such magnets. The LHC dipole magnets use Niobium-Titanium cables (7600 km of them) operating at the superconducting temperature of 1.9 K. The choice of temperature, below the peak of maximal thermal conductivity, ensures that small fluctuations place the system in maximal thermal dissipation, increasing stability. The LHC dipole magnets, whose cross sectional view is shown in Fig. 5.12, provide a field of 8.3 T uniform within 1 part in 10,000, are approximately 15 m long and weight 27 tonnes.

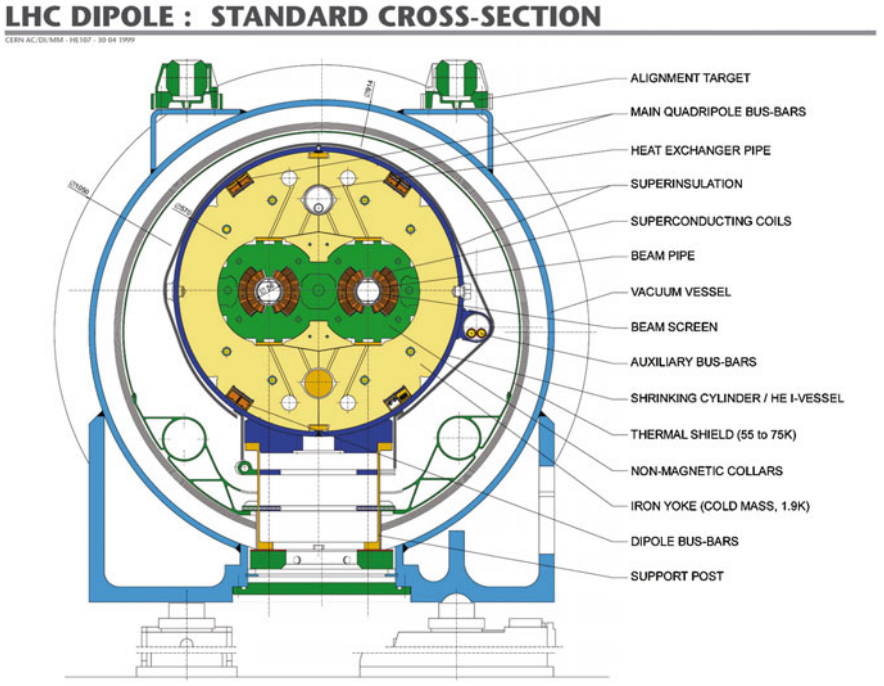


Fig. 5.12 Cross sectional diagram of a Large Hadron Collider dipole magnet [12]. Credit © CERN, printed with permission

Example: Determine the bending radius of the LHC superconducting magnets ($B = 8.3\,\mathrm{T}$, $p = 7\,\mathrm{TeV}/c$).

The bending radius can be calculated from (5.26)

$$\rho = \frac{7000 \times 10^9\,\mathrm{eV}/c}{e\,8.3\,\mathrm{Vs/m^2}} \simeq 2.8\,\mathrm{km}. \tag{5.27}$$

Notice that the corresponding circumference is 17.6 km which is only two thirds of the 27 km ring of the LHC. Indeed, not all the space in the ring can be taken by bending magnets, typically one third being necessary for RF systems, focusing, diagnostics and detectors.

Example: Calculate the equivalent electric field required by a proton travelling at near the speed of light in a circular orbit under a magnetic field of 1 T.

From (5.24) the equivalent electric field necessary for the same centripetal force would be:

$$F_c = qvB = qE \tag{5.28}$$

$$\Rightarrow E = 3 \times 10^8\,\frac{\mathrm{m}}{\mathrm{s}}\,\frac{1\mathrm{Vs}}{\mathrm{m^2}} = 300\,\frac{\mathrm{MV}}{\mathrm{m}} \tag{5.29}$$

5.6 Transverse Focusing

Once the design orbit is defined, and dipole magnets are used to give the required bending, the next task in a synchrotron machine is to keep the particles at or near the design orbit. In order to keep a beam circulating multiple times on the design orbit it is essential that there is a restoring force bringing the beam back to the orbit under the conditions of small transverse displacements. The simplest expression for a restoring force is a linear dependence on the displacement, as in the case of the simple harmonic oscillator for a mass attached to a spring of constant k, whose equation has the form:

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + kx = 0, \tag{5.30}$$

where m is the mass and x is the coordinate of displacement. The solution to (5.30) is a function $x(t) = A\cos(\omega t + \phi)$ so that a displaced mass oscillates about the equilibrium point. A similar situation can be created for the beam of particles if they are subject to a restoring force that increases as a function of the distance from the design orbit. Let x-y be the coordinates for the deviation of the particle from the central orbit, x in the horizontal direction and y on the vertical direction, with s the coordinate of advancement along the ideal path, as shown in Fig. 5.13. Given x the particle deviation from the design orbit in the horizontal plane and perpendicularly to

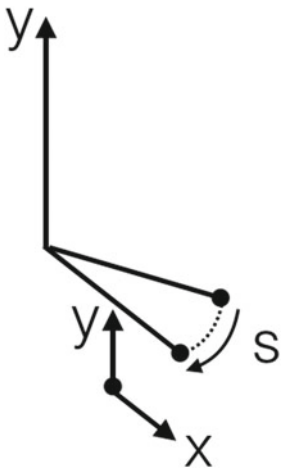


Fig. 5.13 Coordinates in the horizontal (x) and vertical (y) displacement form the design orbit (the *dashed line*)

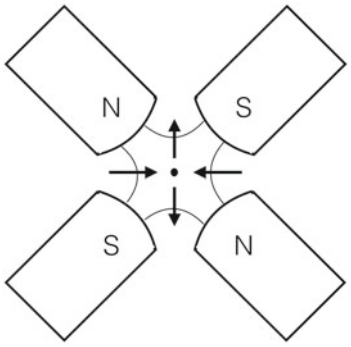


Fig. 5.14 Sketch of a quadrupole magnet. The *arrows* indicate the force on a positive charge

the orbit, the first derivative $\mathrm{d}x/\mathrm{d}s$ is the angle of the particle trajectory with respect to the ideal path line.

A restoring force for the particles can be obtained from the Lorentz force with an appropriate magnetic field:

$$F(x) = qvB_y(x). \tag{5.31}$$

It turns out that the field with the required features $B(x)$ is given by quadrupole magnets (see Fig. 5.14) since they give a linearly increasing magnetic field in the x-y directions

$$B_y = gx, \tag{5.32}$$
$$B_x = gy, \tag{5.33}$$

where g is the restoring constant (the field strength gradient), and the same parameter g can be chosen for the horizontal and vertical restoring fields. As shown in Fig. 5.14,

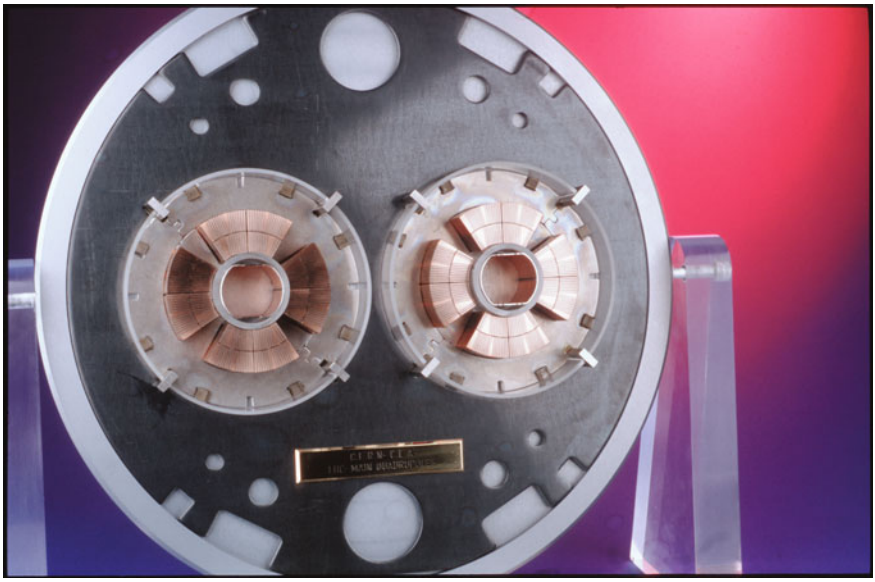


Fig. 5.15 The Large Hadron Collider quadrupole magnets. Credit © CERN, printed from [13] with permission

a positive charge exiting the page will be subject to a restoring force in the horizontal plane towards the center of the magnet.

The photograph in Fig. 5.15 shows the Large Hadron Collider quadrupole magnets. Typical values for g are about 25 T/m (normal conducting) to 220 T/m (superconducting). If we indicate with ρ and k the following ratios:

$$\frac{1}{\rho} = \frac{B}{p/q} \qquad \text{dipole bending} \tag{5.34}$$

$$k = \frac{g}{p/q} \qquad \text{quadrupole focusing} \tag{5.35}$$

then the equation for horizontal focusing from dipole and quadrupole magnets is written as:

$$\frac{\mathrm{d}^2 x}{\mathrm{d}s^2} + x\left(\frac{1}{\rho^2} + k\right) = 0. \tag{5.36}$$

Notice that both the dipole and the quadrupole magnets will have a focusing effect in the horizontal (bending) plane. The equation for vertical motion is instead:

$$\frac{\mathrm{d}^2 y}{\mathrm{d}s^2} - ky = 0, \tag{5.37}$$

and this has no dipole component, but most importantly the quadrupole field is now defocusing (the minus sign). In fact in this configuration the quadrupole field has changed sign and gives no longer a restoring force. The solution to (5.37) is not sinusoidal but an hyperbolic function. However, a rotation by 90° of the quadrupole magnet can give a focusing effect on the y direction, and defocusing in the x, so the equation for the vertical motion with a focusing quadrupole magnet appropriately rotated is

$$\frac{\mathrm{d}^2 y}{\mathrm{d}s^2} + ky = 0. \tag{5.38}$$

In the construction of storage rings, magnets are therefore arranged into a lattice, that is a sequence of magnets performing a sequence of different tasks, e.g. bending, horizontal focusing, vertical focusing, etc., and the solution to the particle motion is derived from the solution to the combined matrix of the lattice. The Fig. 5.16 shows the UK Diamond beam line, with visible magnets arranged in a lattice of dipoles, quadrupoles and sextupoles.

In a strong focusing machine, typical values for displacements from orbit are of the order of millimeters and $\mathrm{d}x/\mathrm{d}s$ is of the order of millirad. The *beam* is therefore the name for the overlapping of many particle trajectories, and the *tune* is the number of

Fig. 5.16 The lattice of the UK Diamond Light Source storage ring, showing the pattern of dipole (in *green*), quadrupole (in *red*) and sextupole (in *yellow*) magnets (coloured online). © Diamond Light Source printed with permission

oscillations, about the design orbit, per turn. Also, the restoring force is not constant, i.e. there is a varying focussing strength, $k(s) \neq \text{const}$, so the equation becomes

$$\frac{\mathrm{d}^2 x}{\mathrm{d}s^2} + k(s)x = 0. \tag{5.39}$$

which is known as Hill's equation, but the solution can still be written in a sinusoidal form:

$$x(s) = \sqrt{\varepsilon}\sqrt{\beta(s)}\cos\left(\Psi(s) + \phi\right), \tag{5.40}$$

where ε is called the beam emittance and β is a function of the lattice (not to be confused with the particle β). The cosine term shows an oscillatory path about the design orbit, with the parameters ε and β describing the transverse size of the beam. The matrix treatment of the lattice and more details on beam emittance can be found in [14].

5.7 Acceleration and Longitudinal Focusing

In synchrotrons, the acceleration of electric charges is done with electric fields:

$$\mathbf{F} = e\mathbf{E}. \tag{5.41}$$

If z is the direction of motion on the design orbit, $\mathbf{E}$ must be aligned with the z direction and the work (W) done by the electric field $\mathbf{E}$ is:

$$\mathrm{d}W = eE_z\mathrm{d}s \Rightarrow W = e\int E_z\mathrm{d}s = eV, \tag{5.42}$$

where V is the voltage applied to the electric charge.

In (5.42), the electric field is synchronous with the accelerating charge. We have seen that in DC accelerators all particles are synchronous with the accelerating potential, by definition. In the Alvarez structure, the synchronisation is achieved by the geometry of the drift tube, the energy gained per gap is given by (5.2) with $n = 1$. In the cyclotron, the synchronisation is done by the spiralling orbit length, which guarantees that the particle is in sync with the radio frequency.

In a synchrotron machine, charges are arranged in bunches and the arrival of each bunch needs to be synchronised with the direction of the electric field. The synchronisation between charges and electric field is not a given of the machine design, it has to be achieved by adjusting the frequency and the phase of the RF. The revolution frequency of the charges is:

$$f = \frac{\beta c}{2\pi R}, \tag{5.43}$$

where β is the relativistic β of the particle and R is the radius of the orbit. The RF frequency f_{RF} must be an integer of f in order to achieve synchronous configuration, $f_{RF} = hf$, where h is called the *harmonic number*. At the LHC for instance, $h = 35{,}640$, so there are h positions around the orbit where a particle can be located in sync and arrive synchronously with the RF. Segments around these points are called *buckets*, and the group of particles in the buckets are called *bunches*. Typically not all buckets are filled in a circulating beam, at the LHC for instance there are 2808 bunches circulating with 25 ns spacing.

The adjustment of the RF allows also for the RF to be used for the purpose of longitudinal focusing. Longitudinal focusing refers to the reduction of the longitudinal spread in an ensamble of electric charges. In a RF cavity the electric field intensity and direction oscillates with a certain frequency (several hundred MHz) and the travelling particles must be synchronous with the electric field direction in the accelerating configuration. Consider first an ensamble of *non*-relativistic particles going through an RF cavity. If the center of the ensamble is placed in sync with the raising component of the electric field sinusoid (see Fig. 5.17), particles arriving later see an higher electric field, and so they speed up, while particles arriving earlier see a lower electric field, so on the next turn they arrive somewhat later. The net effect is a reduction in the spread of particles along the orbit.

The configuration for relativistic particles is however different than what just shown. Relativistic particles get *heavier* under further acceleration, and so a particle that is subject to a higher electric field will travel on the outer side of the bending envelope, and therefore at equal speed it has a longer path length, arriving later in time. The longitudinal focusing configuration is achieved by synchronising charges on the falling side of the sinusoid, as shown in Fig. 5.18. An early particle is subject to a larger electric field, so its mass increases giving it a longer path length from which it will arrive later on the next turn. Between classical and relativistic regimes the RF phase must be quickly switched to avoid de-focusing of the beam.

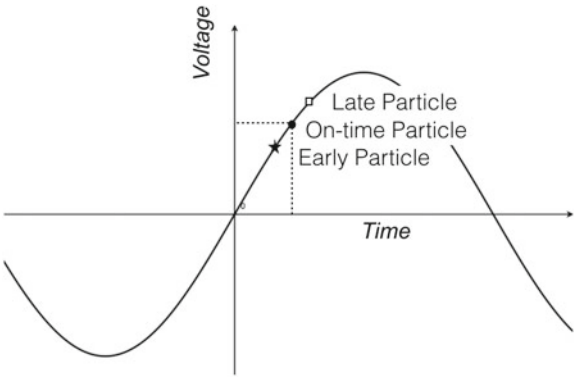


Fig. 5.17 Non-relativistic particles timed with the sinusoidal electric field of an RF cavity. This configuration achieves longitudinal focusing for non-relativistic particles

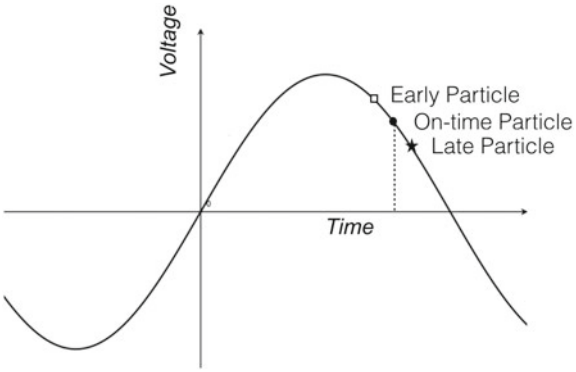


Fig. 5.18 Relativistic particles timed with the sinusoidal electric field of an RF cavity. This configuration achieves longitudinal focusing for relativistic particles

Table 5.1 Large Hadron Colliders machine parameters [15]

Proton energy at collisions (GeV)	7000
Injection energy (GeV)	450
Relativistic gamma	7461
Number of particles per bunch	1.15×10^{11}
Number of bunches	2808
RMS bunch length (cm)	$\simeq$7.55 ($\sim$1 ns)
Synchrotron radiation power (kW)	3.6
Synchrotron radiation energy loss per turn (keV)	6.7
RF frequency (MHz)	400.8
Total RF voltage (MV)	16
Energy gain per turn (keV)	485

The Large Hadron Collider, at CERN's laboratory, is today the most powerful synchrotron in the world. It accelerates and colliders counter rotating beams of protons or lead ions. The main LHC parameters are summarised in Table 5.1 [15]

Example: Calculate the energy gradient of electrons accelerated in vacuum by a traveling plane wave with an electric field of 15 TV/m.

The energy flow of a plane wave is given by the Poynting vector

$$\mathbf{S} = \frac{1}{\mu_0}\mathbf{E} \times \mathbf{B}, \tag{5.44}$$

with $B = E/c$ in vacuum. The Poynting vector gives the power per square meter. If we use the Thomson cross section for the electron:

$$\sigma_T = \frac{8\pi r_e^2}{3} = 6.6 \times 10^{-29}\,\mathrm{m}^2 \tag{5.45}$$

the power transfer to the electron will be:

$$P = S\sigma_T = \frac{E^2\sigma_T}{\mu_0 c} \tag{5.46}$$

On the other hand, if the electron travels with speed v, the accelerating power is given by the average electric field $\bar{E}$ multiplied by the electron charge q and its speed v:

$$P_e = q\bar{E}v. \tag{5.47}$$

Comparing the two expressions for power, (5.46) and (5.47), gives:

$$\frac{E^2\sigma_T}{\mu_0 c} = q\bar{E}v \Rightarrow q\bar{E} = \frac{E^2\sigma_T}{\mu_0 c v}. \tag{5.48}$$

The product $q\bar{E}$ is the energy gradient (measured in N = J/m), and taking for example $v = 0.2c$, inserting the appropriate numbers in (5.48) and converting energy from joules to electronvolts, yields 4.0 MeV/m.

Glossary

Alvaretz-type LINAC A type of LINAC radio frequency accelerator in which the accelerating structure is surrounded by a vessel used to keep the RF as a standing wave

Beam emittance Parameter of the beam that controls the transverse spread

Beam rigidity In a synchrotron, the product $B\rho$ between the bending magnetic field B and radius of the orbit ρ. It equals the ration between the particles momenta p over their electric charge e

$$B\rho = \frac{p}{e}$$

Beam tune Number of oscillations per turn about the design orbit in a storage ring

Betatron Accelerating electron machine based on an induced electric field that accelerates electrons into a circular orbit

Buckets Segments around the design orbit of a synchrotron ring which are synchronous with the accelerating RF

Bunches Ensamble of particles located in the buckets of a storage ring, and therefore synchronous with the accelerating RF

Cockcroft–Walton generator Electrostatic particle accelerator machine based on the Cockcroft rectifier circuit

Cyclotron Accelerator machine featuring an oscillating electric potential and a constant magnetic field used to accelerate charges into a spiral trajectory

LINAC Linear accelerator in which particles are accelerated by a sequence of oscillating electric potentials

Longitudinal focusing The active reduction of the spread of momenta of the particles in the orbit direction

Particle beam Ensamble of particle trajectories around the design orbit

Radio frequency accelerator Particle accelerator based on the application of a sequence of oscillating electric potentials synchronised with the particle position

Synchrocyclotron A cyclotron machine in which the frequency of the oscillating electric potential is adjusted with the accelerating particle momenta

Synchrotron Circular accelerator machine in which both the bending magnetic field and the frequency of the accelerating potentials are adjusted with the particles' velocity along the orbit

Synchrotron light Electromagnetic radiation emitted from charges as they are bent into the circular orbit of a synchrotron

Tandem Van de Graaff accelerator A particular set up of the Van de Graaff accelerator in which two Van de Graaff are positioned in line one after the other, while a stripping target in the middle changes the polarity of the accelerating ions allowing the application of double the voltage of each Van de Graaff

Transverse focusing The active recalling of a circulating beam towards the design orbit, by means of magnets, to reduce the spread of the particles in the plane transverse to the orbit direction. It is mainly achieved through quadrupole magnets appropriately placed along the orbit

Van de Graaff accelerator Electrostatic particle accelerator machine based on the creation of a high voltage by mechanical transport of charges

References

1. H. Geiger, E. Marsden, Proc. R. Soc. **lxxxii**, 495 (1909)
2. J.D. Cockcroft, E.T.S. Walton, Proc. R. Soc. A **137**(831), 229–242 (1932). doi:10.1098/rspa.1932.0133
3. J.D. Cockcroft, E.T.S. Walton, Proc. R. Soc. A **136**(830), 619–630 (1932). doi:10.1098/rspa.1932.0107
4. R.J. Van De Graaff, Phys. Rev. **38**, 1919–1920 (1931)
5. R.J. Van De Graaff, Nucl. Instrum. Methods **8**, 195–202 (1960)
6. https://www.flickr.com/photos/brookhavenlab/4153316575/in/album-72157611346071353/. Accessed 13 Dec 2016
7. Fermilab Creative Services, Image number 95-1039 (1995)
8. T. Wangler, *Principles of RF Linear Accelerators* (Wiley, New York, 2008)

9. L. Arnaudon et al., Linac4 Technical Design Report, CERN-AB-2006-084 ABP/RF (2006), https://cdsweb.cern.ch/record/1004186/. Accessed 9 Dec 2016
10. J. Schwinger, Phys. Rev. **75**, 1912 (1949)
11. UK Diamond Light Source Ltd, Image number 06EC3822
12. Diagram of an LHC dipole magnet, Image number CERN-DI-9906025-01, https://cds.cern.ch/record/40524. Accessed 30 Nov 2016
13. CERN Image Server, Image no. CERN-AC-0001016-01, https://cds.cern.ch/record/40918. Accessed 9 Dec 2016
14. E. Wilson, *Particle Accelerators* (Oxford University Press, Oxford, 2001)
15. The Large Hadron Collider Design Report vol 1, The LHC Main Ring, CERN-2004-003-V1, CERN-2004-003, CERN-2004-003-V-1

Part II
Interaction Mechanisms and Detectors

Part II of the book is dedicated to both the mechanisms of interaction between radiation and matter, and how those principles are exploited in the design and construction of modern detector devices. The material is presented through the description of each mechanism of interaction immediately followed by a review of the detection methods and technologies based on that particular mechanism. Ionisation is introduced first, underpinning a large number of detection techniques. Scintillation, electromagnetic and hadronic cascades, transition radiation, and neutrino interactions follow in the later chapters.

Chapter 6
Ionisation and Multiple Scattering

Abstract The detection of particles is possible through their energy loss in the material of the detector. In other words, in order to be detected a particle must interact with the material of the detector, and in doing so it must transfer energy in some form that can be recognised. The content of this chapter focuses on the energy loss of an heavy charged particle through scattering by atomic electrons, and the consequent ionisation of the medium. The same mechanism underpins the principles of the detectors described in Chaps. 7–9.

6.1 Ionisation: Bohr Classical Derivation

Radiation can interact with matter in several ways, depending on the nature of the radiation, its energy and the interacting medium. Among these, ionisation of atoms is a very significant mechanism for a large number of applications. Heavy charged particles can interact electromagnetically with the electrons they encounter in the absorbing material. A Feynman diagram of this process at tree level is shown in Fig. 6.1, where a particle with mass M and electric charge ze comes in the vicinity of an atomic electron. The energy that the incoming particle loses in this interaction process can be determined from classical and semi-classical arguments.

The Bohr's classical derivation of the energy loss can be summarised as follows [1]: consider the heavy particle of mass M and electric charge ze approaching the electron (whose mass is m_e) with speed v and let the distance of closest approach of the straight trajectory be indicated by b, as shown in Fig. 6.2.
The following assumptions are made:

- the electron is free and initially at rest,
- the movement of the electron is slow during the interaction time, so that its electric field is undisturbed,
- the path of the heavy particle is undisturbed by the interaction with the electron (a consequence of the difference between the respective masses, $M >> m_e$).

L. Cerrito, *Radiation and Detectors*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-53181-6_6

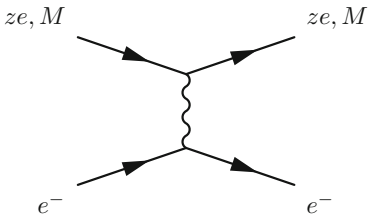


Fig. 6.1 Leading order diagram of the scattering of a charged particle by an electron

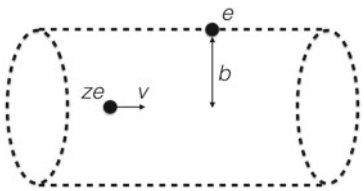


Fig. 6.2 Sketch of the classical approximation for an ionisation process

The momentum (I) transferred to the electron at time t from the electric force $F(t)$ can be written as:

$$I = \int F(t)\mathrm{d}t, \tag{6.1}$$

where the integral runs over the total time of the interaction. From (6.1), if the net contribution to the momentum of the longitudinal component of the electric force is null due to symmetry considerations, writing only the transverse component $E_\perp$ yields:

$$I = \int F(t)\mathrm{d}t = e\int E_\perp \mathrm{d}t = e\int E_\perp \frac{\mathrm{d}t}{\mathrm{d}x}\mathrm{d}x = e\int \frac{E_\perp}{v}\mathrm{d}x, \tag{6.2}$$

where the integration has been changed from time to the spatial variable x.

The electric field $E_\perp$ can be derived using Gauss law applied to the cylinder shown by the dashed lines in Fig. 6.2:

$$\int E_\perp \times 2\pi b\,\mathrm{d}x = \frac{ze}{\epsilon_0} \quad \Rightarrow \quad \int E_\perp \mathrm{d}x = \frac{ze}{2\pi\epsilon_0 b}, \tag{6.3}$$

and substituting this into (6.2) yields:

$$I = \frac{e}{v}\frac{ze}{2\pi\epsilon_0 b} = \frac{ze^2}{2\pi\epsilon_0 b v}. \tag{6.4}$$

The momentum transferred to the electron, shown in (6.4), corresponds to an energy gain for the electron equal to

$$\Delta E = \frac{I^2}{2m_e} = \frac{z^2e^4}{8\pi^2\epsilon_0^2 b^2 v^2 m_e}, \tag{6.5}$$

where the classical relationship between energy and momentum has been used. The calculation just shown allows the determination of the energy lost to one electron in the interaction by the heavy charged particle. If instead of only one electron the interaction involves many electrons distributed over a given volume with volumetric density N_e, then the energy lost from such collisions at a distance b in a thickness dx, i.e. in a volume element $dV = 2\pi\, b\, db\, dx$ is

$$-\,\mathrm{d}E = \Delta E\, N_e \mathrm{d}V = \Delta E\, N_e\, 2\pi\, b\, \mathrm{d}b\, \mathrm{d}x. \tag{6.6}$$

Substituting the expression for ΔE of (6.5) into (6.6) yields the energy loss formula at a generic distance b from the atomic electrons

$$-\,\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{z^2 e^4 N_e}{4\,\pi\,\epsilon_0^2\, v^2\, m_e}\frac{\mathrm{d}b}{b}, \tag{6.7}$$

which integrated between suitable minimum and maximum values of b is written as

$$-\,\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{z^2 e^4 N_e}{4\,\pi\,\epsilon_0^2\, v^2\, m_e}\ln\left(\frac{b_{\max}}{b_{\min}}\right). \tag{6.8}$$

The (6.8) has been obtained using only classical physics, however reasonable values for $b_{\max}$ and $b_{\min}$ can only be assigned with non-classical arguments. Clearly $b_{\min}$ cannot be equal to zero otherwise the ratio in the logarithm would go to infinity, implying an infinite energy transfer while there is only a maximum value of energy, namely the full energy of the heavy particle, that could be transferred to the electron. One could use Compton's wavelength of the electron (λ_e) and invoke Heisenberg uncertainty principle to write:

$$b_{\min} \approx \lambda_e = \frac{h}{p} = \frac{2\pi\hbar}{\gamma m_e v}. \tag{6.9}$$

Regarding $b_{\max}$, the interaction time would classically be $t = b/v$, and accounting for relativity it is $t = b/\gamma v$. This time interval must be small compared to the period of the bound electron $\tau = 1/\nu_e$ and therefore:

$$\frac{b}{\gamma v} < \frac{1}{\nu_e} \;\Rightarrow b_{\max} = \frac{v\gamma}{\nu_e}. \tag{6.10}$$

Using (6.9) and (6.10) together with (6.8) gives Bohr's semi-classical formula for energy loss:

$$-\,\frac{dE}{dx} = \frac{z^2 e^4 N_e}{4\,\pi\,\epsilon_0^2\, v^2\, m_e}\ln\left(\frac{m_e v^2 \gamma^2}{2\pi\hbar\nu_e}\right) \qquad \text{(Bohr formula)}. \tag{6.11}$$

The (6.11) describes to a good approximation the interaction of very heavy particles with matter. It predicts that the energy loss is proportional to $1/v^2$ (or equivalently to $(\beta\gamma)^{-2}$), indicating a rapid drop as a function of the incident particle's energy, then a logarithmic rise with $\ln(v^2)$. The denominator in the logarithm represents the average excitation and ionisation potential of the absorber, approximately 13.5 eV.

6.2 Bethe-Block Formula

A more sophisticated quantum-mechanical formula for the average energy loss in an ionisation process is given by the so-called Bethe-Blocks equation [2]:

$$-\left\langle \frac{\mathrm{d}E}{\mathrm{d}x} \right\rangle = 2\pi N_a r_e^2 m_e c^2 \rho \frac{Z}{A} \frac{z^2}{\beta^2} \left[\ln\left(\frac{2m_e \gamma^2 v^2 W_{\max}}{I^2} \right) - 2\beta^2 \right], \qquad (6.12)$$

where N_a is Avogadro's number, r_e is the classical electron radius ($r_e = e^2/4\pi\epsilon_0 m_e c^2 \simeq 2.8\,\mathrm{fm}$), ρ is the density of the absorber (in g/cm^3), A is the atomic mass of the absorbing material (in g/mol), z is the charge of the incident particle in units of the electron charge, $W_{\max}$ is the maximum energy transfer per collision (in MeV) and I is the mean excitation potential (in eV). The behaviour of (6.12) with β is, as for the Bohr formula of (6.11), of type $1/\beta^2 \ln\left(\text{const}\,\beta^2\gamma^2\right)$ and describes well the experimental energy loss for intermediate Z materials in the region $\beta\gamma$ between 0.1 and 1,000.

The plot in Fig. 6.3 shows the energy loss of positive muons on a copper target over a wide range of the incident momenta. The vertical scale is expressed in units of MeV cm^2/g = MeV/cm g/cm^3, that is the plot shows the average of the distribution of $-(1/\rho)\mathrm{d}E/\mathrm{d}x$ (called the *Mass Stopping Power*). Being this the mean of a one sided distribution, in actual facts the most probable value is much lower than what plotted. On the x-axis, the scale shows the relativistic factor $\beta\gamma$, independent on the incident particle mass, as well as displaying the muon momentum considering that $\beta\gamma = p/Mc$, where M is the particle's mass. At non relativistic speeds, $v/c < 0.9$, the energy loss has a functional dependence of type $1/\beta^2$, decreasing to a minimum for $\beta\gamma \simeq 3$–4 with a value of 1.1–1.8 MeV cm^2/g. Classically, the $1/\beta^2$ behaviour is understood as slower particles feel the electric force of the atomic electrons for a longer time (as in Bohr calculation). The minimum is known as *minimum ionisation*. For muons, the minimum region extends over momentum values of a few GeV/c. For $v/c > 0.9$ the logarithmic term becomes dominant, slowly increasing the energy loss. The relativistic rise for $\beta\gamma > 4$ is understood as the effect of the transversal electric field which increases due to the Lorentz stretching of that electric field component, effectively increasing the interaction cross section. Beyond and below the $\beta\gamma$ range of 0.1–1,000, the Bethe formula is in disagreement with the experimental data.

For values of $\beta\gamma$ smaller than 0.1, the incident particle velocity is no longer much larger than the orbital velocity of the electrons and the energy loss does not grow as predicted by the Bethe equation. So called *shell corrections* must be included in

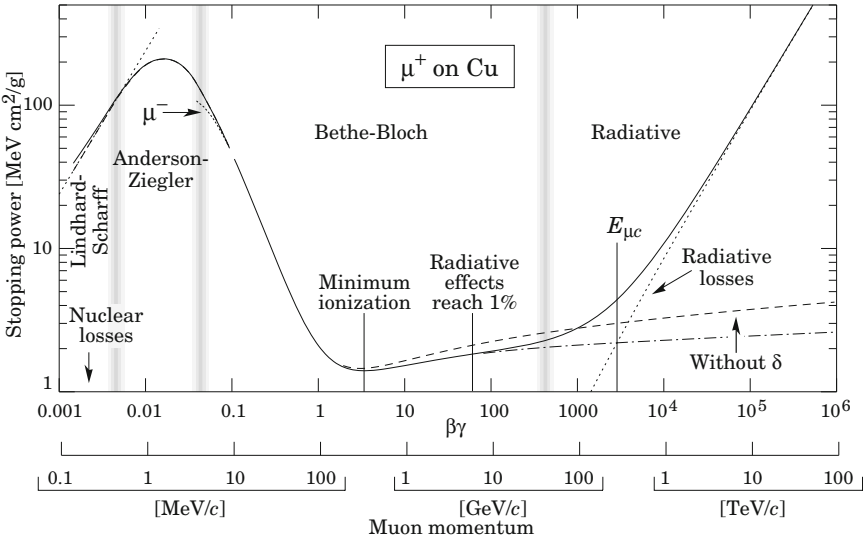


Fig. 6.3 Mean energy loss for positive muons incident onto a Cu target, as predicted by the Bethe-Block formula and its modifications. Reprinted from [3] with permission

(6.12) to describe the low momentum regime. Some of these, which go beyond the scope of this textbook, are discussed in [4].

For values of $\beta\gamma$ greater than 1,000 the medium becomes polarised, suppressing the increase in energy loss due to relativistic effects. However, at high energy and for instance already at values of around 300 GeV/c for muons in copper the radiative energy loss (described in Chap. 10), and not ionisation, becomes the dominant mechanism.

Target dependence For most practical purposes, $\langle \mathrm{d}E/\mathrm{d}x \rangle$ in a given material is a function of β alone. Except for hydrogen, the same heavy particles with the same velocity will have approximately the same rate of energy loss when penetrating different materials. The dependence on different absorbing materials is given in the (6.12) by the ratio Z/A and the graph in Fig. 6.4 shows the calculated curves of energy loss for various targets, from liquid hydrogen to lead. Radiative losses are not included in this simulation.

The minima are all occurring at $\beta\gamma$ between 3 and 4, for Z ranging from 2 to 100, dropping towards 3 for the heavier elements. On the horizontal axis, the labels indicate explicitly the momenta of muons, pions and protons corresponding to the same $\beta\gamma$ scale. Because heavy particles, such as muons, pions or protons are commonly occurring in today's laboratories of high energy physics with $\beta\gamma$ values in the range of 2 to 100, these particles are also called *minimum ionising*.

At the minimum position of the Bethe-Block curve, the rate of energy loss is about 1–2 MeV cm^2/g for most materials. The plot in Fig. 6.5 shows the minimum value of

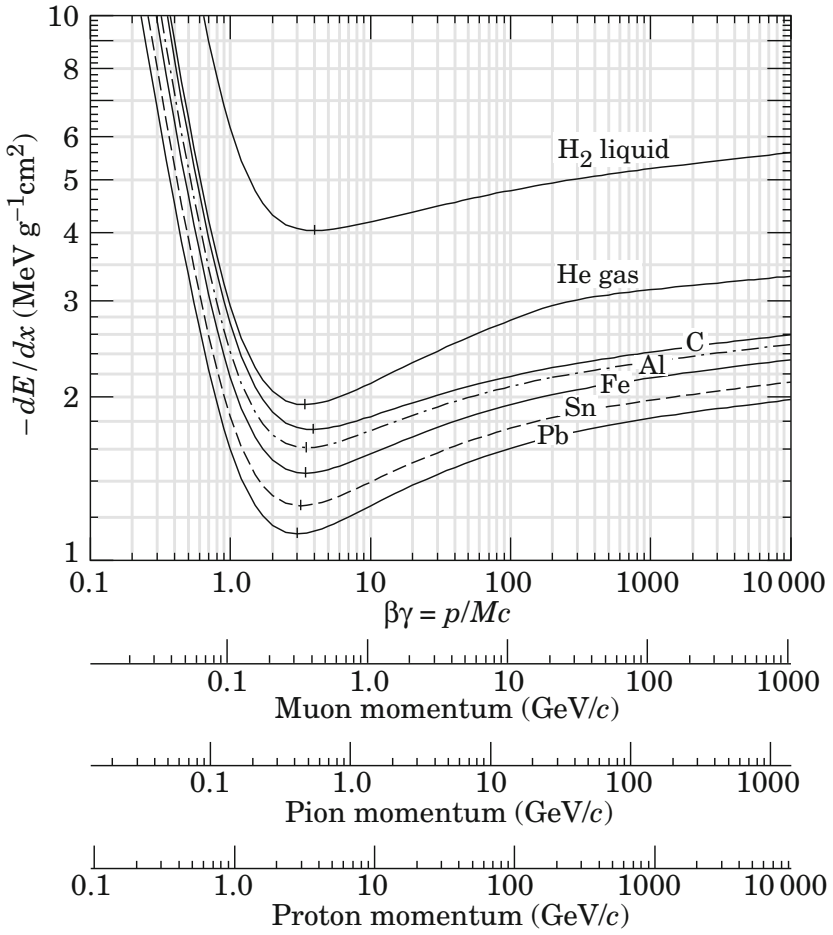


Fig. 6.4 Mean energy loss rate for a number of target elements. Radiative effects are not included. Reprinted from [3] with permission

the energy loss curve as a function of the target Z. The dependence is approximately linear in the natural logarithm of Z.

Example: Determine the energy loss of a muon at minimum ionisation as it traverses a calorimeter corresponding to 200 cm of iron ($\rho = 8\,\mathrm{g/cm^3}$).

A muon with momentum in the region of the minimum ionisation looses energy at a rate given by the values around the minimum point in Fig. 6.3, roughly corresponding to a mass stopping power of 2 MeV cm^2/g. The energy loss can therefore be calculated as:

$$\Delta E = 2\,\mathrm{MeV\,cm^2 g^{-1}} \times 200\,\mathrm{cm} \times 8\,\mathrm{g\,cm^{-3}} \simeq 3.2\ \mathrm{GeV}. \tag{6.13}$$

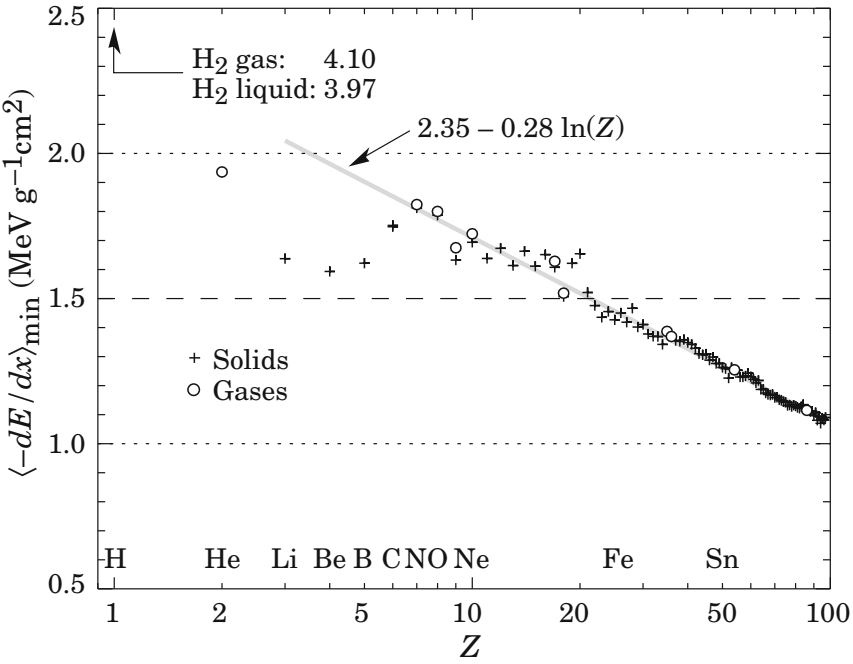


Fig. 6.5 Minimum ionisation point as a function of target Z. Reprinted from [5] with permission

Example: Considering a beam of protons with energy of 600 MeV, determine the thickness of copper (whose density is 8.96 g/cm^3) needed to lower the energy down to 400 MeV.

Assuming a constant rate of energy loss of 3.5 MeV cm^2/g, the thickness needed is:

$$\Delta x = \frac{200\,\text{MeV}}{3.5\,\text{MeV cm}^2\,\text{g}^{-1}\,8.96\,\text{g cm}^{-3}} \simeq 6.4\,\text{cm} \tag{6.14}$$

6.3 Particle Identification Through Energy Loss

The Bethe formula of (6.12) shows that the energy loss of heavy charged particles in a given material is dependent only on the relativistic $\beta\gamma$ of the particle. Since $\beta\gamma$ corresponds to

$$\beta\gamma = \frac{p}{Mc}, \tag{6.15}$$

the same amount of energy loss is obtained at different momenta for heavy particles of different mass. In other words, if the amount of deposited energy is measured

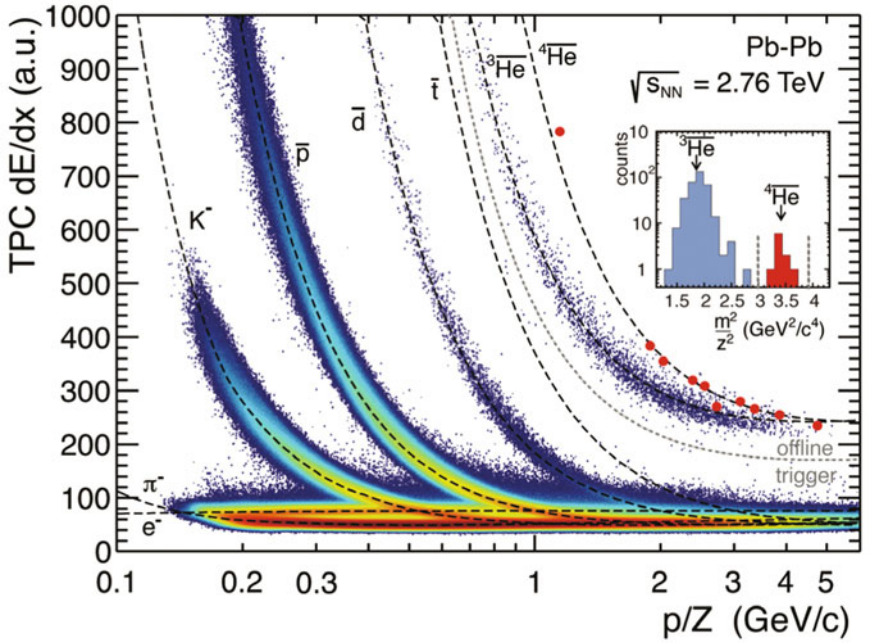


Fig. 6.6 The ionisation signal in the Time Projection Chamber of the ALICE detector versus the magnetic rigidity of the incident particles together with the expected curves for negatively-charged particles. Reprinted from [6] with permission

alongside a measurement of the momentum p of the incident particle, then the mass of that particle can be inferred. An example is shown in Fig. 6.6 for the ALICE experiment at the LHC [6]. The vertical axis reports the ionisation signal in the Time Projection Chamber (more details on the TPC detector are given in Sect. 7.9), which is strictly linked to the amount of energy deposited in the active material of the detector. The horizontal axis shows the momentum p scaled by the electric charge Z of the particle depositing energy in the chamber's active material. The curves resemble the functional behaviour of the Bethe formula with different scalings on the horizontal axis, compare with Fig. 6.4. The combined measurement of the ionisation signal and of the momentum allows therefore inference of the mass of the particle, and is an extensively used technique for particle identification at colliding particle experiments.

6.4 Statistical Distribution of Energy Loss and the Range

The Bethe formula describes the *mean* value of the energy loss of an incident heavy charged particle travelling through a given material. The mean value of a distribution is not in general coincident with the most probable value, and this is the case also

for the statistical distribution of the energy loss. In fact, the energy loss ΔE is the collective effect of the individual energy transfers δE_n from the incident particle during multiple (n) collisions, and is therefore a statistical process:

$$\Delta E = \Sigma_{i=1}^{N} \delta E_n. \tag{6.16}$$

The function that best describes the probability of energy loss in such mechanism is a Landau distribution [7, 8], characterised by an asymmetric unidimensional curve raising from zero to a maximum value and decreasing with a long tail. The tail in this case is due to events with large energy transfers. In general, the curve becomes more symmetric for heavier particles and thicker absorbers, owing to the fact that collisions with large energy transfers are less likely. The curves in Fig. 6.7 show examples of the distribution of energy deposited by a 500 MeV/c pion incident on thin silicon wafers. The most probable energy loss is significantly lower than the mean energy loss, because the rarer high-energy-transfer collisions push the mean value into the tail of the distribution; the mean energy loss is therefore best suited only when interested in energy loss by multiple particles.

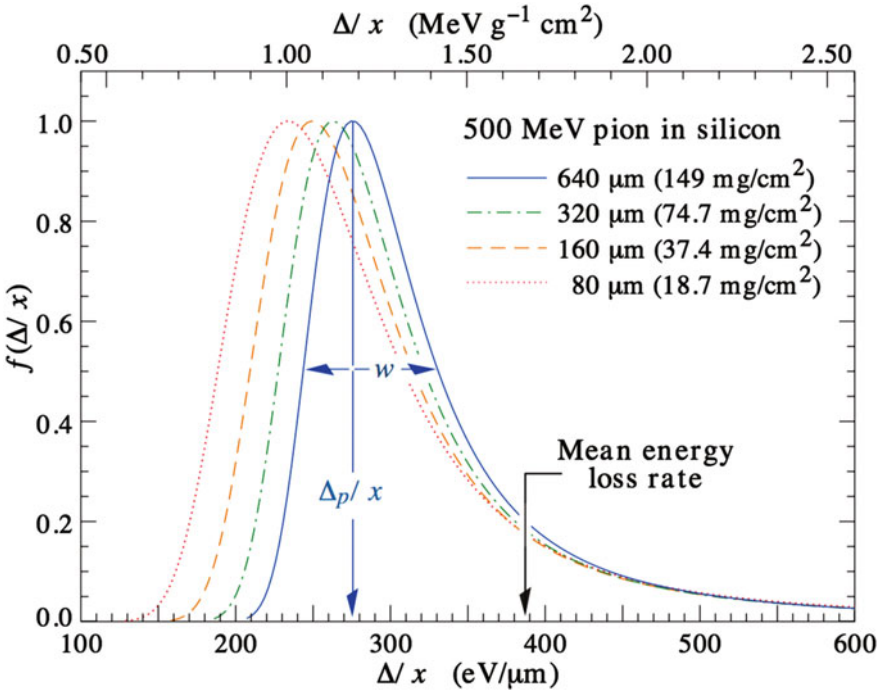


Fig. 6.7 The normalised distribution of energy loss for 500 MeV pions in silicon wafers. The width w is the full width at half maximum. Reprinted from [3] with permission

Table 6.1 Range of heavy charged particles by ionisation in various target materials

Material	$\beta\gamma$	Range, R/M (g cm^{-2} GeV^{-1} c^2)
Lead	1	180
Lead	10	7,000
Carbon	1	100
Carbon	10	5,000
Helium gas	5	900

The $\mathrm{d}E/\mathrm{d}x$ function can be integrated over the full path from the initial energy down to zero to determine the total *range* (R) of the incident particle passing through the target material (assuming that the only energy-loss mechanism is ionisation):

$$R = \int_E^0 \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)^{-1} \mathrm{d}E. \tag{6.17}$$

Since $\mathrm{d}E/\mathrm{d}x$ is a function of β, R is also only a function of β. The Table 6.1 shows approximate values of R/M, where M is the mass of the incident particle, for a few target materials. Naturally, the actual range of a given radiation will be subject to fluctuations because there are fluctuations in the loss of energy.

As an example, with a simple calculation the range of protons with momentum of 1 GeV/c, for which $\beta\gamma \simeq 1.07$, travelling in in lead ($\rho = 11.3\,\mathrm{g/cm^3}$) can be determined as:

$$R = \frac{180\ \mathrm{g\,cm^{-2}GeV^{-1}}c^2 \times 0.938\,\mathrm{GeV}/c^2}{11.3\,\mathrm{g\,cm^{-3}}} \simeq 15\,\mathrm{cm}. \tag{6.18}$$

In practice, the range is only a useful calculation for low-energy hadrons and muons below a few hundred GeV; for greater energies, radiative energy loss becomes significant.

6.5 Bragg Peak

The rate of the energy loss by ionisation, per distance travelled, for heavy charged particles is approximately constant, and slowly varying, over a wide kinematic range of the incident particle. This can be seen in Fig. 6.3 for $\beta\gamma$ between approximately 3 and nearly 1,000, which for a proton corresponds to a wide range of momenta between about 2.8 and 1,000 GeV/c. For $\beta\gamma$ below about 3, the rate of energy loss grows rapidly and is multiple times larger than the minimum ionisation rate. This behaviour implies that for a beam of incident heavy particles, most of the energy loss happens in the kinematic range below the minimum ionising point. A plot of

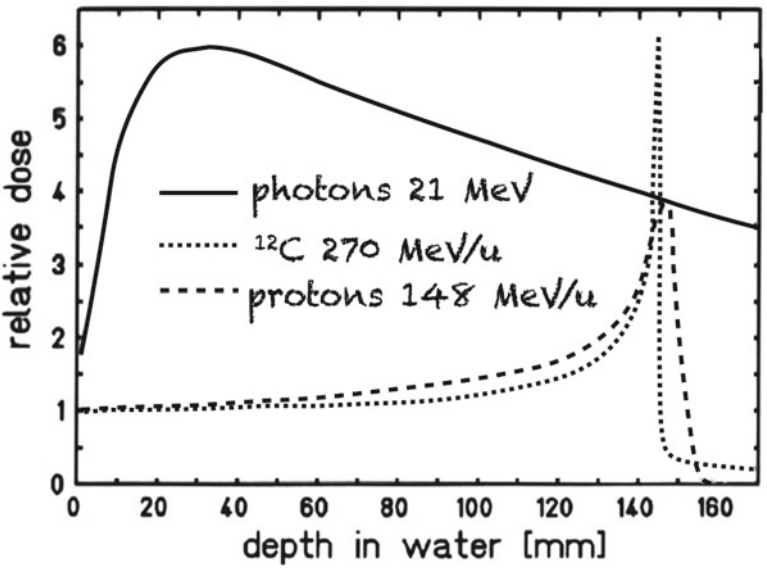


Fig. 6.8 Depth dose of protons, carbon ions and photons in water. Adapted after the original plot courtesy of G. Kraft, Biophysics GSI Darmstadt

the distribution of differential energy loss as a function of the penetration depth would therefore appear approximately constant until the momentum is degraded to $\beta\gamma < 1$, at which point a peak of energy deposition would follow. This is called a *Bragg Curve*, and the peak in such distribution is called the Bragg Peak.

The plot in Fig. 6.8 shows the Bragg peaks for protons and Carbon ions penetrating in water. This feature can be extremely advantageous in nuclear physics and medical applications. On the one hand, the Bragg Peak can be used to calibrate the energy of a beam by directing such beam to a stopping material of known composition, thickness, and peak position to infer the beam energy. On the other hand, the possibility of delivering energy only at a precise depth with ions is used as a cancer therapy technique alternative to X-ray treatment [9]. The photograph in Fig. 6.9 shows the synchrotron of the Italian National Center for Oncological Hadron-therapy in Pavia. The machine, producing beams of 250 MeV protons and 4800 MeV carbon ions, is based on the synchrotron principle. In 2016, the Particle Therapy Co-Operative Group [11] counts 67 such facilities in operation worldwide and a total of 137,000 patients treated since 1954.

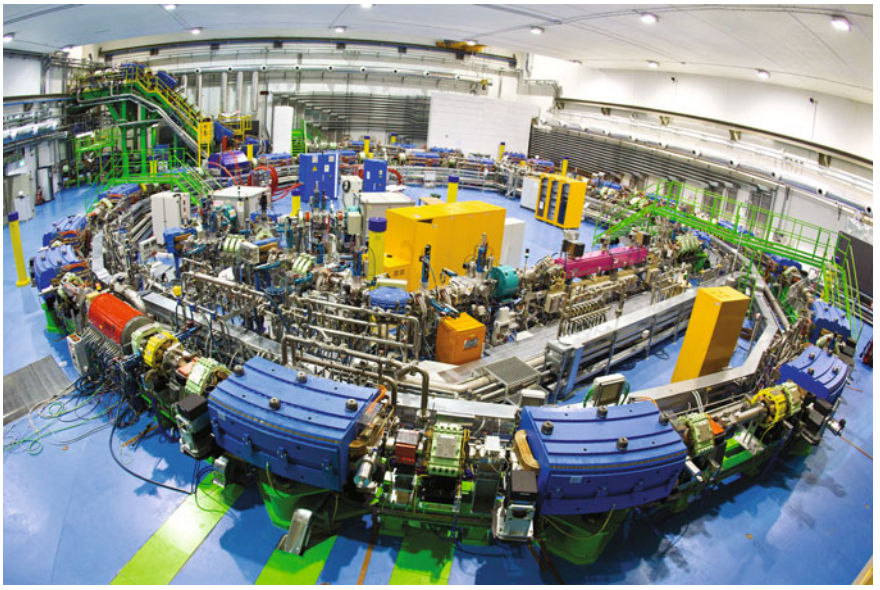


Fig. 6.9 Synchrotron at the Italian National Center for Oncological Hadron-therapy (CNAO). Credit CNAO. From [10] with permission

6.6 Multiple Scattering

The previous sections have presented in some detail the energy loss of a heavy charged particle by the inelastic collision with the atomic electrons of the interacting medium. In addition to this effect, charged particles will also suffer repeated elastic collisions (Coulomb scattering) with the nuclei which will deflect the original direction of motion. For hadronic particles, the strong interaction with the nuclei also contributes to the interaction mechanisms. The electromagnetic elastic collision with nuclei is given by the Rutherford formula, and for multiple repeated collisions (for example more than about 20) this phenomenon is called *multiple Coulomb scattering*.

Consider a particle incident perpendicularly to a material of thickness x as shown in Fig. 6.10. The particle will undergo multiple scattering and emerge at a point placed

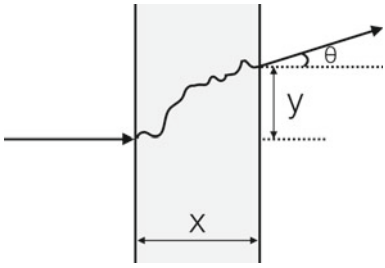


Fig. 6.10 Sketch of a particle entering a medium of thickness x and undergoing multiple Coulomb scattering until emerging at position y and with angle θ with respect to the original trajectory

a certain distance y from the original trajectory, and will continue in a direction making the angle θ_{MS} with the original trajectory. For small angles, θ_{MS} is a random variable described approximately by a Gaussian distribution with width given by the following empirical formula [12]:

$$\Delta\theta_{MS} = \frac{13.6\ \text{MeV}}{\beta cp} z\sqrt{z/X_0}\,[1 + 0.038\ln(x/X_0)]\,, \tag{6.19}$$

where z is the charge number of the incident particle, p and βc are its momentum and speed, respectively, and x/X_0 is the thickness of the scattering medium expressed in radiation lengths (X_0, which will be introduced in Chap. 10). Notice that in (6.19), the angular spread due to multiple scattering is proportional to the charge of the incident particle, to the thickness of the medium, and also to the inverse of the momentum, p. This implies that at small momenta multiple scattering limits the resolution of the momentum measurement of a particle by tracking reconstruction (the momentum resolution is discussed in Chap. 8). A quantity that is also useful in describing multiple scattering effects is the deflection

$$\Delta y_{MS} = \frac{x}{\sqrt{3}}\Delta\theta_{MS}. \tag{6.20}$$

The plot in Fig. 6.11 shows the width of the deflection distribution for muons as they pass through three different detector components, having different thickness values, in the CDF experiment [13]. The deflection is shown as a function of the muon transverse momentum p_T, and exhibits the reciprocal dependence indicated

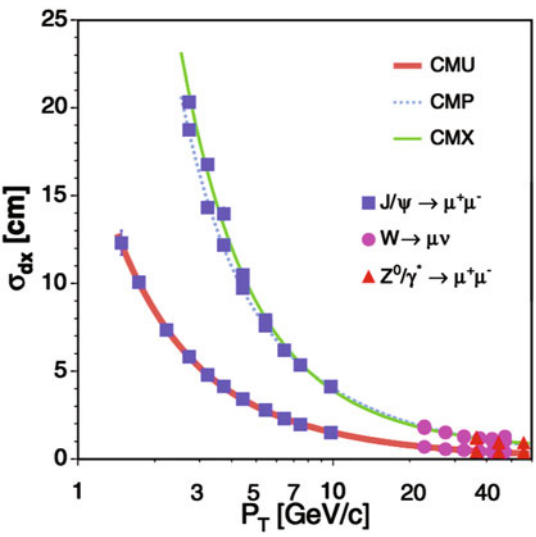


Fig. 6.11 Width of the deflection distribution for muons as they pass through three different detector components in the CDF experiment. Printed from [13] with permission

in (6.19), i.e. softer muons are subject to significant deflections amounting to tens of centimeters.

Example: Muons with momentum of 10 GeV/c go through 100 cm of iron. The value of X_0 for iron is $X_0(\mathrm{Fe}) = 1.76\,\mathrm{cm}$ and therefore $x/X_0 = 57$. The width of the multiple scattering angle is:

$$\Delta\theta_{MS} = \frac{13.6\,\mathrm{MeV}}{10\,\mathrm{GeV}}\sqrt{57}\,(1 + 0.038\ln 57) = 0.01\,\mathrm{rad}. \tag{6.21}$$

The deflection would be therefore of about

$$\Delta y_{\mathrm{MS}} = \frac{100}{\sqrt{3}}0.01 = 0.6\,\mathrm{cm}. \tag{6.22}$$

Glossary

Bethe-Block ionisation formula Quantum-mechanical formula describing the rate of energy loss of a heavy charged particle through an absorbing material by the ionisation mechanism

Bohr ionisation formula Semi-classical approximation of the rate of energy loss of a heavy charged particle through an absorbing by the ionisation mechanism

Bragg curve Distribution of differential energy loss as a function of the penetration depth, in the ionisation mechanism by heavy charged particles

Bragg peak Peak appearing towards the end of the curve of differential energy deposition versus penetration depth, in the ionisation mechanism by heavy charged particles

Energy straggling The spread of energy loss by ionisation away from the mean value, due to the statistical nature of the energy transfer in multiple collisions with target centers

Hadron oncological therapy Use of heavy charged particles, typically protons and ions, in the treatment of tumors by exploiting the localised energy deposition of a Bragg peak at a given depth of tissue

Ionisation interaction The interaction of radiation with the electrons of the absorbing material, leading to the ionisation of some of that material atoms

Mass Stopping Power Rate of energy loss with penetration depth per unit volumetric density of the absorbing material

Minimum ionisation Minimum of the Bethe-Block formula for the rate of energy loss per unit density, corresponding to about 1.5 MeV cm^2/g

Minimum ionising particle Heavy charged particles, such as muons, pions or protons, moving at speeds corresponding to $\beta\gamma$ between approximately 2 and 100, for which the energy loss by ionisation is around its minimum

PID by energy loss Particle identification technique using the measured energy deposition and the momentum of a particle to infer the mass of that particle through the known energy loss rate by ionisation

Range The penetration depth of a particle through an absorbing material

Shell corrections Corrections to the Bethe-Block formula in the low $\beta\gamma$ regime

References

1. W.R. Leo, *Techniques for Nuclear and Particle Physics Experiments* (Springer-Verlag, Berlin, 1993)
2. J. Beringer et al., (Particle Data Group). Phys. Rev. D **86**, 010001 (2012). 2013 partial update for the 2014 edition
3. S. Eidelman et al., (Particle Data Group). Phys. Lett. B **592**, 1 (2004)
4. Stopping Powers and Ranges for Protons and Alpha Particles, ICRU Report No. 49 (1993); tables and graphs of these data are available at http://physics.nist.gov/PhysRefData/. Accessed 5 Dec 2016
5. J. Beringer et al., (Particle Data Group). Phys. Rev. D **86**, 010001 (2012)
6. The ALICE Collaboration, Int. J. Mod. Phys. A **29**, 1430044 (2014)
7. L.D. Landau, J. Exp. Phys. (USSR) **8**, 201 (1944)
8. P.V. Vavilov, Sov. Phys. JETP **5**, 749 (1957)
9. E. Fokas, G. Kraft, H. An, R. Engenhart-Cabillic, Ion beam radiobiology and cancer: Time to update ourselves. Biochimica et Biophysica Acta (BBA) - Rev. Cancer **1796**(2), 216–229 (2009)
10. Fondazione CNAO, Centro Nazionale di Adroterapia Oncologica per il trattamento dei tumori con protoni e ioni carbonio. Press Office
11. Particle Therapy Cooperative Group, Jermann M. Particle Therapy Statistics in 2014. Int. J. Particle Ther. **2**(1), 50–54 (2015)
12. G.R. Lynch, O.I. Dahl, Nucl. Instrum. Methods **B58**, 6 (1991)
13. C.D.F. Collaboration, Measurement of the $t\bar{t}$ production cross section in pp collisions at $\sqrt{s}$=1.96 TeV using lepton plus jets events with semileptonic B decays to muons. Phys. Rev. D **72**, 032002 (2005)

Chapter 7
Gaseous and Liquid Ionisation Detectors

Abstract Ionisation is one of the most common processes used today for radiation detection and measurement. Ionisation of gases and liquids forms the basis for a large number of devices that have been developed over a span of several decades. This Chapter presents the physics principles of ionisation, the properties of transport of electrons and ions in gases, and how these properties are exploited in common devices from the simplest, the Geiger counter, to the Time Projection Chamber.

7.1 Principles of Ionisation Detectors

In a gaseous ionisation detector, a charged particle passing through a certain volume filled with gas will create free charges from primary and secondary ionisation (see the sketch in Fig. 7.1), leaving a trail of charges.

Primary ionisation refers to electrons freed by the direct interaction with the incident particle. Secondary ionisation refers to charges freed as the result of ionisation of nearby atoms by the liberated (primary) electrons or to ionisation following atomic de-excitation. If an electric field is applied to the volume containing the ionised gas, charges will start to drift along the electric field lines. Due to the mobility of electrons and ions in gases being greater than in liquids, it is more common for these detectors to employ gases rather than liquids as an active medium. However, liquids have also found widespread applications, particularly detectors with liquid argon and xenon.

Multiple types of gaseous ionisation detectors have been developed over several decades from the mid 1900's, as summarised in the diagram in Fig. 7.2. Earlier configurations include the Ionisation chamber, the Proportional counter and the Geiger-Muller counter. The most advanced setup is that of the Time Projection chamber (TPC).

The basic configuration of a gaseous ionisation detector is shown in Fig. 7.3. A cylindrical volume is filled with a noble gas, chosen for its low ionisation potential and electron affinity. A potential difference is applied between the central wire and the casing creating a radial electric field. Different processes will happen depending on the magnitude of the potential difference V_0 applied between the anode and cathode. The diagram in Fig. 7.4 shows the general gain-voltage characteristics for a

L. Cerrito, *Radiation and Detectors*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-53181-6_7

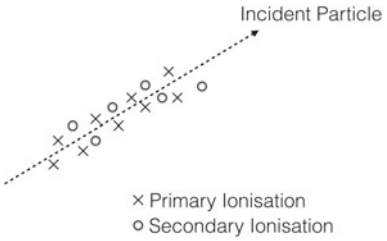


Fig. 7.1 Schematics of a radiation particle going through a gas and leaving a trail of primary and secondary ionisation charges

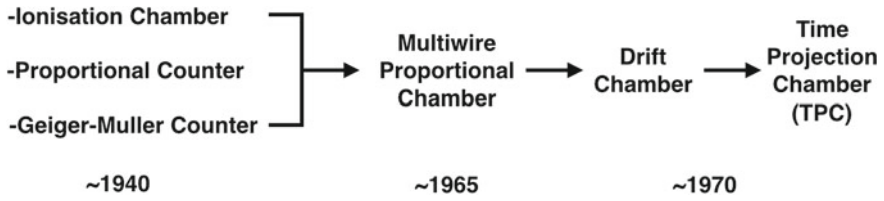


Fig. 7.2 Timeline of the development of gaseous ionisation detectors, from the first counters to the sophisticated Time Projection Chambers

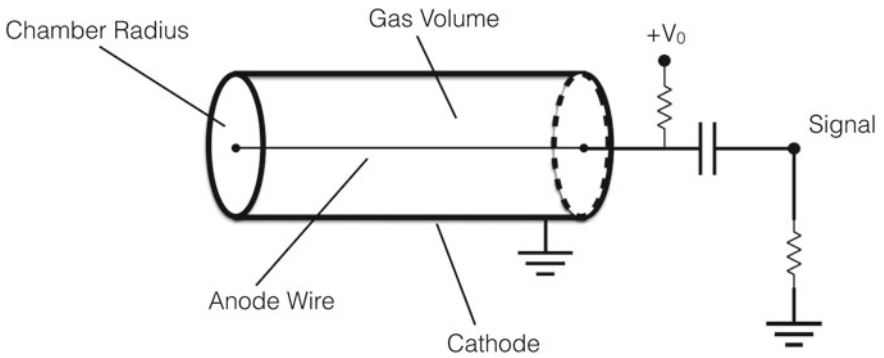


Fig. 7.3 Schematics of a ionisation chamber

gaseous ionisation detector. In region I, the electric field is not sufficiently high and all charges recombine before collection. The first plateaux on the left hand side of the plot, labelled region II and just beyond the recombination region (I), the voltage is sufficient to allow collection of all liberated charges, which drift towards the anode and cathode producing a small signal proportional to the liberated charge itself. This is the working regime of the ionisation chamber. In region III, the middle portion of the plot, the voltage is high enough to cause further ionisation by the electrons accelerated near the wire. In the space of a few millimeters from the wire, an avalanche of electron-ion pairs develops, with the number of e-ion pairs so created being proportional to the original number of pairs. This is effectively a charge amplification and the working

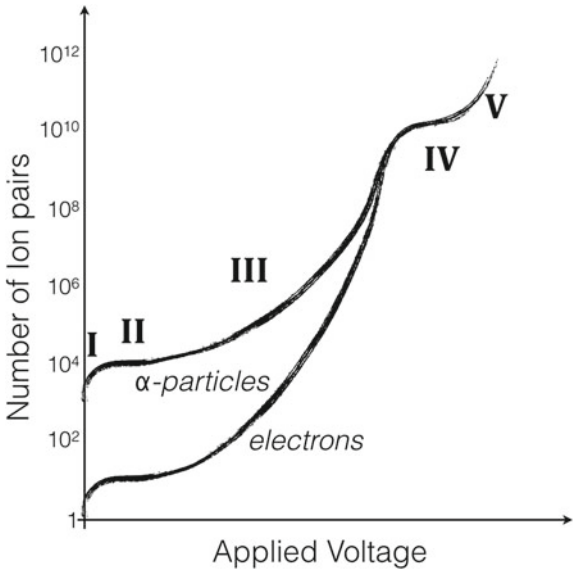


Fig. 7.4 Gain versus applied voltage in a gaseous chamber for incident electrons and α. Sketched after [1]

regime of the proportional counter, with amplification factors that can reach values of the order of 10^6.

In region IV, discharge occurs in the gas, i.e. a chain reaction of many avalanches around the entire anode. In this regime, the output current is completely saturated giving the same amplitude regardless of the initial deposited energy. The discharge is caused mainly by photons emitted by the de-excitation of the atoms or molecules and can be suppressed somewhat with quenching additives (photon absorbers). Detectors operating in this region are called Geiger counters. For voltage higher than region IV, the discharge happens continuously with or without incident radiation.

Ionisation chambers and Geiger counters are widely used today for radiation detection. Some smoke alarms (see an example in the photograph of Fig. 7.5) also use the ionisation chambers technology by employing a small amount of α emitting material and collecting the ionisation charges within a small chamber, which are monitored against a disrupted flux from smoke particles.

General considerations that guide the choice for the appropriate ionisation detector gases include low ionisation potential and the minimisation of the electron recombination process (therefore the choice of a gas with low electron affinity such as noble gases) and the minimisation of discharge under high voltage. The discharge is suppressed by adding a hydrocarbon (such as ethane, methane or CO_2). On the other hand, the drift speed, which affects the dead time and the occupancy of the detector and is crucial in some applications, may be optimised with the choice of special additive gases. Finally, the addition of an alcohol in small percentages reduces the accumulation of residuals onto the anode wire. The actual gases used for ionisation detectors are therefore an appropriate mixture of the above components, depending on the configuration of the detector and its measurement objectives.

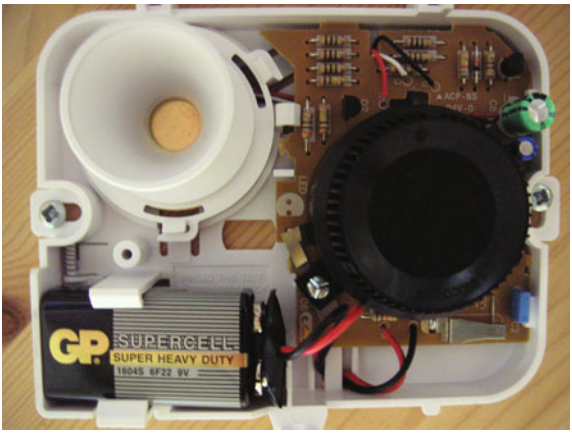


Fig. 7.5 The inside of an ionisation smoke alarm, with clear view of the horn (*top left*), battery (*bottom left*) and the ionisation chamber, on the *right hand side* [2]

7.2 General Characteristics of Ionisation Detectors

A detector is an instrument capable of measuring the presence and possibly some of the properties of radiation. The *sensitivity* is the capability of detection, and it is worth stressing from the onset that no detector is sensitive to all types of radiation and at all energy ranges. Detectors are optimised for one or more types of radiation and normally in a specific energy range.

The sensitivity of a ionisation detector depends generally on:

- the cross section for the ionising reactions
- the detector mass
- the detector noise
- the type and thickness of the protective material surrounding the sensitive volume of the detector

In addition to detection, most instruments give a measurement of energy as one of the basic properties, and this is achieved in ionisation detectors by relating the amount of ionisation with the deposited energy.

For detectors designed to measure energy, one of the most important parameters is therefore the *resolution*:

$$\text{Resolution} = \frac{\Delta E}{E}, \tag{7.1}$$

where ΔE is the full width at half maximum of the response curve to mono-energetic (E) incident radiation. Typical energy resolution values might be anywhere between a fraction of a percent, for a detector with good resolution, to values as high as tens of a percent. In general, because of the statistical nature of the energy deposition in ionisation, the resolution improves with higher energies, i.e. it is a function of the energy itself. This can be easily seen with Poisson statistics: if the average energy required to produced one ionisation process is w, the total number of ionisation processes can

be written as $n = \hat{E}/w$, where $\hat{E}$ is the total deposited energy (as opposed to the measured energy E, which is lower than $\hat{E}$). Since the measured energy E is proportional to the number of ionisation processes n, then according to Poisson statistics the uncertainty on E will be $\Delta E \propto \sqrt{n} \sim \sqrt{E}$, yielding

$$\text{Resolution} = \frac{\Delta E}{E} \propto \frac{\sqrt{E}}{E} \sim \frac{1}{\sqrt{E}}. \tag{7.2}$$

Therefore, the resolution of a measurement of energy improves with higher values of the energy deposited.

The relationship between the amount of released ionisation charge and the incident energy is called the *response*. The response might be a linear function, but in general this is not the case. The *response distribution* is the distribution of the response when the detector is subject to monochromatic (i.e. single energy) radiation. Such distribution might be close to a Gaussian but would typically exhibit non-Gaussian tails.

The *response time* is the time interval that the detector requires to form the response signal. The duration of the response time is important because a second event, for example one closely trailing the first event, might not be accepted by the detector during the response time, leading to the so-called *dead time*.

7.3 Ionisation Processes and Transport

There are two main processes that can occur in an atomic gas X under the passage of a heavy charged particle, such as a proton:

- Excitation: $X + p \rightarrow X^* + p$,
- Ionisation: $X + p \rightarrow X^+ + p + e^-$.

Excitation by a proton is a resonant reaction with a cross section of approximately $10^{-17}\,\text{cm}^2$. There are no free electrons or ions which are created, but the subsequent de-excitation might lead to further reactions. The ionisation by a proton has a cross section of about $10^{-16}\,\text{cm}^2$ with no resonant energy but a threshold energy is required to activate the process. Further processes might occur and include atomic de-excitation (e.g. the so-called Penning effect as in $\text{Na}^* + \text{Ar} \rightarrow \text{Na} + \text{Ar}^+ + e^-$), or the formation of molecular ions (e.g. $\text{He}^+ + \text{He} \rightarrow \text{He}_2^+$). The number of electron-ions pairs produced through the ionisation process depends on the energy deposited, and on the average ionisation energy for the gas. In general, an average energy of about 30 eV is needed to form an electron-ion pair, and this depends weakly on the incident particle or the type of gas. The Table 7.1 shows the ionisation energy and the mean energy for the creation of an electron-ion pair for noble gases. Note that the average energy needed fro the creation of an electron-ion pair is higher than the ionisation energy because a portion of the deposited energy is typically absorbed through excitation.

Table 7.1 First ionisation energy and mean energy for the creation of an electron-ion pair in noble gases

Gas	Ionisation energy (eV)	Mean energy for e-ion creation (eV)
Helium	24.6	41
Neon	21.6	36
Argon	15.8	26
Krypton	14.0	24
Xenon	12.1	22

For a minimum ionising charged particle that travels a distance L within a gas with an average electron-ion creation potential W, the average number of electron-ions pairs liberated can be approximately estimated from:

$$\bar{n}_{\mathrm{e-ion}} = \frac{\langle \mathrm{d}E/\mathrm{d}x \rangle \; L}{W}, \tag{7.3}$$

where $\langle \mathrm{d}E/\mathrm{d}x \rangle$ is the average energy loss of the incident particle. For example, for a particle at minimum ionisation losing 2 MeV cm^2/g in a gas of density $\rho_g = 10^{-3}$ g/cm^3 the energy $\bar{E}$ deposited is:

$$\bar{E} = \rho_g \left\langle \frac{1}{\rho} \frac{\mathrm{d}E}{\mathrm{d}x} \right\rangle = 2 \frac{\mathrm{MeV cm}^2}{g} 10^{-3} \frac{\mathrm{g}}{\mathrm{cm}^3} = 2\,\mathrm{keV/cm}. \tag{7.4}$$

For an average electron-ion liberation energy of 30 eV/pair, the number of electron-ion pairs produced by the travelling incident particle is therefore given by:

$$\bar{n}_{\mathrm{e-ion}} = \frac{2\,\mathrm{keV cm}^{-1}}{30\,\mathrm{eV/pair}} = 60\ \mathrm{pairs/cm}. \tag{7.5}$$

Of interest is also the mean free path λ_I between subsequent ionisation processes, and that is given by:

$$\lambda_I = \frac{1}{n_e \sigma_I}, \tag{7.6}$$

where n_e is the density of electrons and σ_I is the ionisation cross section. The Fig. 7.6 shows the electron production cross section in a collision process of protons onto three targets (He, H_2 and O_2) as a function of the proton energy between 5 keV and 5 MeV. Cross sections are of the order of 10^{-20} m^2 and using (7.6) it is possible to estimate a typical mean free path λ_I of the order of 20–100 μm.

Once the electron-ion pair has been created, it is important that they remain separated until collection of the charge on the anode and cathode. Recombination (that is the recombination of the ion and the electron to form a neutral atom, with the

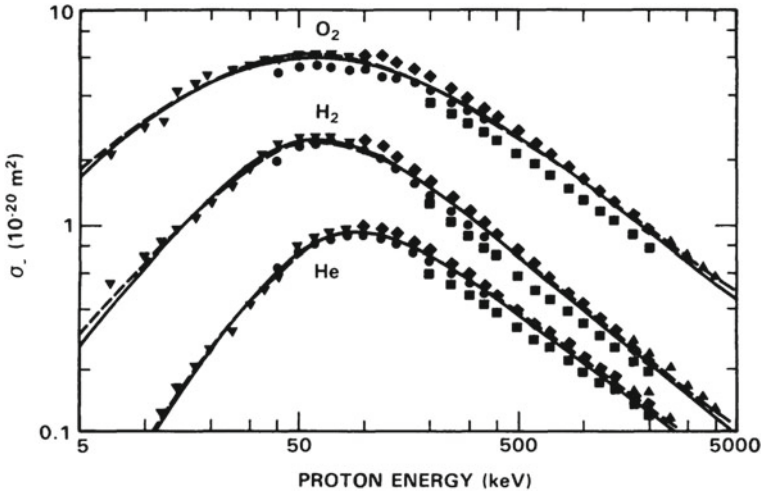


Fig. 7.6 Measured electron production cross section in a collision process of protons onto three targets (He, H_2 and O_2) compared to fits from various models. Reprinted from [3] with permission

possible emission of a photon) and electron attachment process $e^- + X \rightarrow X^- + \gamma$ are phenomena that might occur before collection and reduce the efficacy of the detection mechanism.

Once the electron-ion pairs are liberated, what happens next depends on whether there is an electric field in the active volume or not. If there is no electric field applied, then electrons and ions will move away from the point of liberation, colliding with nearby molecules and will eventually recombine. The electrons and ions move according to their thermal speeds $\bar{v}$, which can be estimated starting from the thermal kinetic energy:

$$\frac{1}{2}M\bar{v}^2 \sim \frac{3}{2}kT \tag{7.7}$$

$$\Rightarrow \bar{v} \sim \sqrt{\frac{3kT}{M}}, \tag{7.8}$$

where k is Boltzmann's constant, T is the temperature, and M is the mass of the gas molecule. Typical velocities are of order 10^4 cm/s for ions, and $\sim 10^6$ cm/s for electrons, which is faster due to their smaller mass.

The distribution of charges in space after a time t can be described by a Gaussian function centered about the point of initial ionisation:

$$\frac{\mathrm{d}N}{\mathrm{d}x} \sim e^{-x^2/f(t)}, \tag{7.9}$$

where $f(t)$ can be parametrised as a linear function of time with a *diffusion* coefficient D, i.e. $f(t) \sim Dt$.

Under the effect of an applied electric field, the ions and the electrons will accelerate along the electric field lines. The collective motion in this case corresponds to the superposition of drift and a thermal random motion and the average velocity in these conditions is called *drift velocity* (v_d). A simplified formula for the behaviour of the drift velocity can be derived considering the average acceleration $\bar{a}$ and time τ between collisions of the charge carriers

$$\bar{v}_d \sim \bar{a}\tau = \frac{eE}{M}\tau, \tag{7.10}$$

where M is the mass of the ion or the electron. The average time between collisions depends on the mean free path $\bar{L}$ and the thermal speeds v_T through $\tau = \bar{L}/v_T$ so that:

$$\bar{v}_d \sim \frac{eE}{M}\frac{1/(n\sigma)}{v_T} = \frac{eE}{M}\frac{A}{N_A\rho\sigma\sqrt{3kT/M}}, \tag{7.11}$$

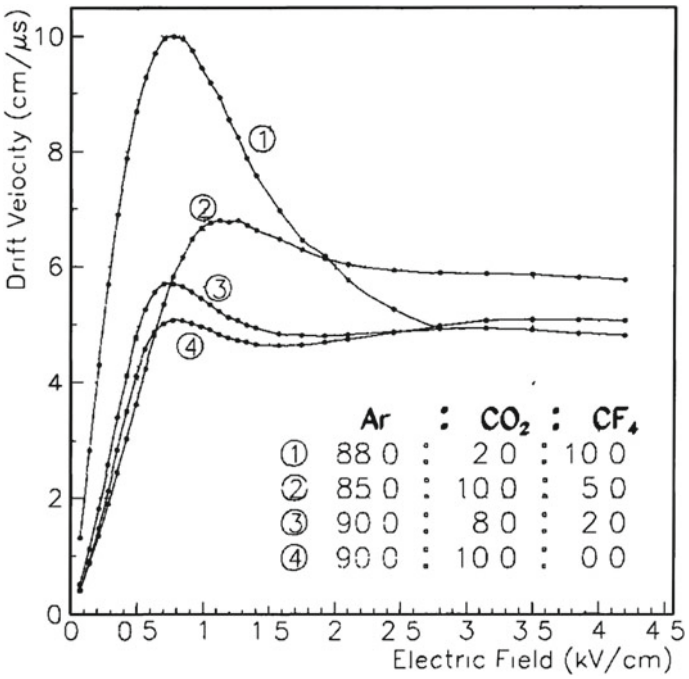


Fig. 7.7 Drift velocity of electrons as a function of electric field in several Ar-CO_2-CF_4 mixtures. Printed from [4] with permission

where n is the density of atoms, σ the cross section for elastic collisions, A is the atomic weight, N_A is Avogadros number and ρ is the mass density. While σ is approximately the same for electrons and ions, the mass M in the denominator determines that drift speeds are much slower, some two order of magnitudes, for ions than for electrons.

The electric field in the numerator also describes a linearly increasing speed with E; for electrons this is true up to a certain saturation threshold, after which the drift speed is approximately equal to the thermal speed. The plot in Fig. 7.7 shows the drift velocity of electrons as a function of electric field in several Ar-CO_2-CF_4 mixtures [4], showing the saturation regime beyond 2 kV/cm. The Ar-CF_4 combination is known to yield a fast gas, while CO_2 helps reducing the electrical breakdown of the gas. Finally, the *mobility* of electrons and ions in the gas is defined as

$$\mu = \frac{v_d}{E} \quad \text{(mobility)}, \tag{7.12}$$

where v_d is the drift velocity appropriate for the electrons or ions, and E is the strength of the electric field.

7.4 Ionisation Chamber

When a gas is subject to an electric field just sufficient to collect the electrons and ions on the electrodes, away from the point of separation, the chamber is said to operate in unit-gain mode. This means that most of the electron-ions pairs produced by the ionisation process are collected at the anode and cathode of the chamber. This is the working region for *ionisation detectors*. Contrary to what one might think, the pulse signals are not due to the charge collected, but to the current induced from the movement of ions and electrons as they drift towards the cathode and anode. The equivalent circuit is shown in Fig. 7.8 for a simplified ionisation chamber with parallel plate geometry. In the absence of any ionisation, the voltage across the resistor R, is $V_R = 0$ and the voltage V_0 appears across the plates.

When ionisation is created, the ion pairs drift between the plates and give rise to induced charges on the electrodes that reduce the chamber voltage from its equilib-

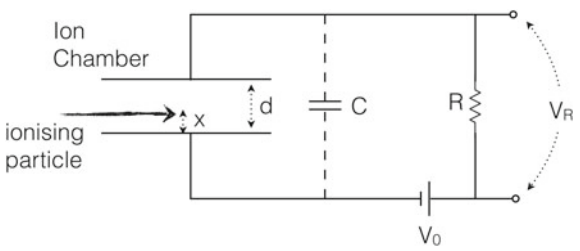


Fig. 7.8 Equivalent circuit of a basic ionisation chamber with parallel plate geometry

rium value of V_0. A voltage V_R, equivalent to the drop from V_0, will then appear across R. After the charge has all been collected, the circuit will start a return to equilibrium with $V_R \rightarrow 0$, the speed of which is given by the time constant of the circuit, RC.

The pulse shape can be calculated ignoring edge effects and recombination, from the following considerations [5]: a particle ionises the gas at a distance x from one of the plates (the distance between plates being d, as shown in the sketch of Fig. 7.8). The electric field between the plates is $E = V_0/d$. Suppose also the time constant of the circuit is longer than the drift time of both the electrons and the ions. If V_0 is the voltage applied to the chamber, the energy stored in the equivalent capacitor C in equilibrium is equal to:

$$\varepsilon = \frac{1}{2} C V_0^2. \tag{7.13}$$

Following an ionisation event and after a time t has elapsed, both the electrons and ions, moving with drift speeds v^- and v^+ respectively, will have travelled a distance $v^{\pm}t$ and moved to regions of lower potential, so that the energy balance is:

$$\frac{1}{2} C V_0^2 = n_0 e E (v^+ t) + n_0 e E (v^- t) + \frac{1}{2} C V_C^2, \tag{7.14}$$

where n_0 is the number of liberated electron-ion pairs. Rearranging the (7.14) yields:

$$\frac{1}{2} C \left(V_0^2 - V_C^2\right) = n_0 e E (v^+ + v^-) t, \tag{7.15}$$

and also

$$\frac{1}{2} C \left(V_0 + V_C\right) \left(V_0 - V_C\right) = n_0 e \frac{V_C}{d} (v^+ + v^-) t. \tag{7.16}$$

The signal voltage is measured across R and is $V_R = V_0 - V_C$, and is also much smaller than the input voltage, $V_R << V_0$ therefore by approximating:

$$V_0 + V_C = V_0 + V_0 - V_R \simeq 2V_0 \;\Rightarrow V_C \approx V_0. \tag{7.17}$$

Inserting (7.17) into (7.16) gives:

$$\frac{1}{2} C \left(2V_0\right) \left(V_R\right) = n_0 e \frac{V_-}{d} (v^+ + v^-) t, \tag{7.18}$$

and finally

$$V_R = \frac{n_0 e}{d\, C} (v^+ + v^-) t. \tag{7.19}$$

The relationship of (7.19), giving the voltage amplitude of the signal output, is valid only while both the electrons and the ions are drifting. After a time $t^- = x/v^-$ has elapsed, all the electrons reach the anode, so the second term in (7.19) becomes

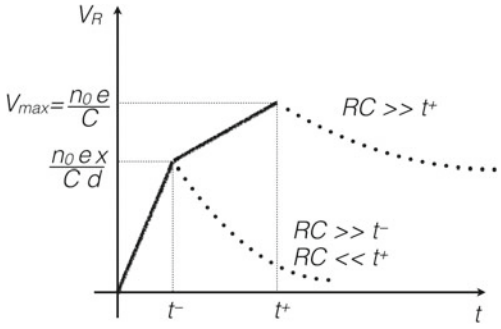


Fig. 7.9 Sketch of the puse shape for a basic parallel plate ionisation chamber configuration

a constant equal to the time interval $t^- = x/v^-$, and so the voltage V_R becomes

$$V_R = \frac{n_0 e}{d\,C}(v^+ t + x). \tag{7.20}$$

When all ions have arrived, after a time t^+

$$t^+ = \frac{d - x}{v^+}, \tag{7.21}$$

the voltage reaches:

$$V_R = \frac{n_0 e}{d\,C}(d - x + x) = \frac{n_0 e}{C}, \tag{7.22}$$

following which the system returns to equilibrium within a time characterised by a time constant RC. The value of (7.22) is also the maximum amplitude of the pulse when the time constant of the circuit $RC >> t^+$. The pulse derived in the (7.19), (7.20) and (7.22) is sketched in Fig. 7.9. Note that when the time constant of the circuit is chosen so as to use both the electrons and ions drifts, the maximum of the pulse is independent on the position of ionisation x, but the time of the pulse is very long therefore limiting the chamber occupancy. For faster operation only the electron drift can be used, but the pulse height for this part is proportional to x.

7.5 Proportional Chamber

In a proportional chamber, the operating voltage is high enough that the electrons from the primary ionisation are accelerated sufficiently to produce further ionisation of the gas; in fact, the number of electron and ions pairs produced from secondary ionisation might be several orders of magnitude larger, although still proportional to the original number of electrons and ions liberated by the ionising radiation. The process takes the form of an *avalanche*, with the liberated electrons causing further ionisations. As electron-ion pairs are being separated and move along the electric

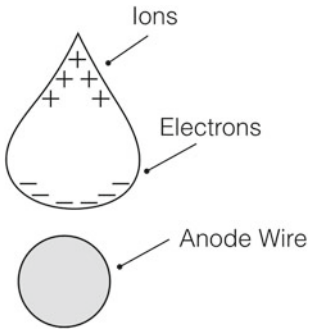


Fig. 7.10 Sketch of charge distribution of electrons and ions formed in an avalanche and moving towards or away from the wire, respectively, under the influence of the electric field

field lines, ions moving away from the positive terminal are slower and lag behind creating a *drop-like* charge distribution shape, as shown in Fig. 7.10.

Starting from the number of liberated electrons n_0, the total number of electrons, n, created in a path of length x can be expressed by:

$$n(x) \sim n_0 e^{\alpha x}, \tag{7.23}$$

where α is related to the mean free path of secondary ionisation by $\alpha = 1/\lambda_{S.I.}$. The ratio $n(x)/n_0$ is called the *multiplication factor*, where common proportional chambers operate with factors of the order of 10^6–10^7. Because the amplification and avalanche formation takes place only in regions where the electric field is high enough, the geometry is an important aspect of proportional chambers. The simplest configuration is cylindrical (as shown already in Fig. 7.3), with an anode wire running through the middle of a cylindrical cathode. The polarity is also important here and must be correct, electrons must be attracted towards the wire not move away from it, since with the radial electric field at distance r given by

$$E = \frac{1}{r} \frac{V_0}{\ln(b/a)}, \tag{7.24}$$

(where b and a are the radii of the chamber and the wire, respectively) only near the wire the field is strong enough to form the avalanche. A small diameter for the wire is preferred both to achieve high fields (of the order of MV/m) and to reduce the volume of multiplication.

An analysis of the signal pulse similar to what done for the parallel plate ionisation chamber can be applied here, and two major differences arise. The first is that about all the charge in the proportional chamber originates within the avalanche region, regardless of where the original ionisation was, and in a cylindrical configuration that means near the wire. Since almost all the electrons are created near the wire, the bulk of the induced signal is actually not from the electrons but from the ions which move *away* from the wire and towards the cathode. The time development of the signal pulse is also important; it can be divided into the drift time and the multiplication

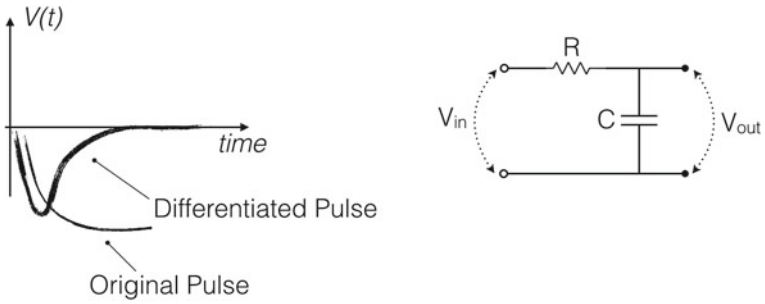


Fig. 7.11 Differentiation of a slow pulse by an RC circuit

time. The drift time is the time it takes the electrons to drift towards the wire and start the avalanche. This depends on the type of gas and the drift distance, with typical drift times of the order of micro-seconds. Once the avalanche formation starts, there is a multiplication time that affects the rising edge of the pulse. This time might be of the order of nano-seconds. Finally, because the electric field decreases with distance from the wire, the ions slow down as they move away, so that the pulse becomes very long and is typically shortened by using a signal differentiating circuit (RC), such as that in Fig. 7.11.

The basic fill gas requirements are similar to those for an ionisation chamber; noble gases such as argon are used for low ionisation potential and low working voltage for avalanche formation. To operate the chamber at the highest voltage, and therefore the highest gain, a photon absorber is usually required to suppress the process of discharge between anode and cathode. Methane (CH_4) or Ethane (C_2H_6) are often mixed to the noble gas for this purpose, in concentrations that might vary between 20 and 50%. Finally, as prolonged irradiation can cause the formation of polymers (that is strings made of groups of CH_2) coating the anode wires and reducing the chamber's gain and efficiency, an alcohol (or also O_2 but its electron affinity might not make it an optimal choice) is often added to the gas mixture to reduce such effect and to prolong the chambers life. The picture in Fig. 7.12 shows deposits on anode wires after prolonged irradiation.

Finally, charge division methods can be used to locate the position of ionisation along the direction parallel to the wire. This is done by collecting the charge at both ends of the anode wire and analysing the pulses with an Analog to Digital Converter (ADC). If the wire has length L (see Fig. 7.13), the position x is related to the collected charges Q_A and Q_B by:

$$\frac{x}{L} = \frac{Q_A}{Q_A + Q_B}. \tag{7.25}$$

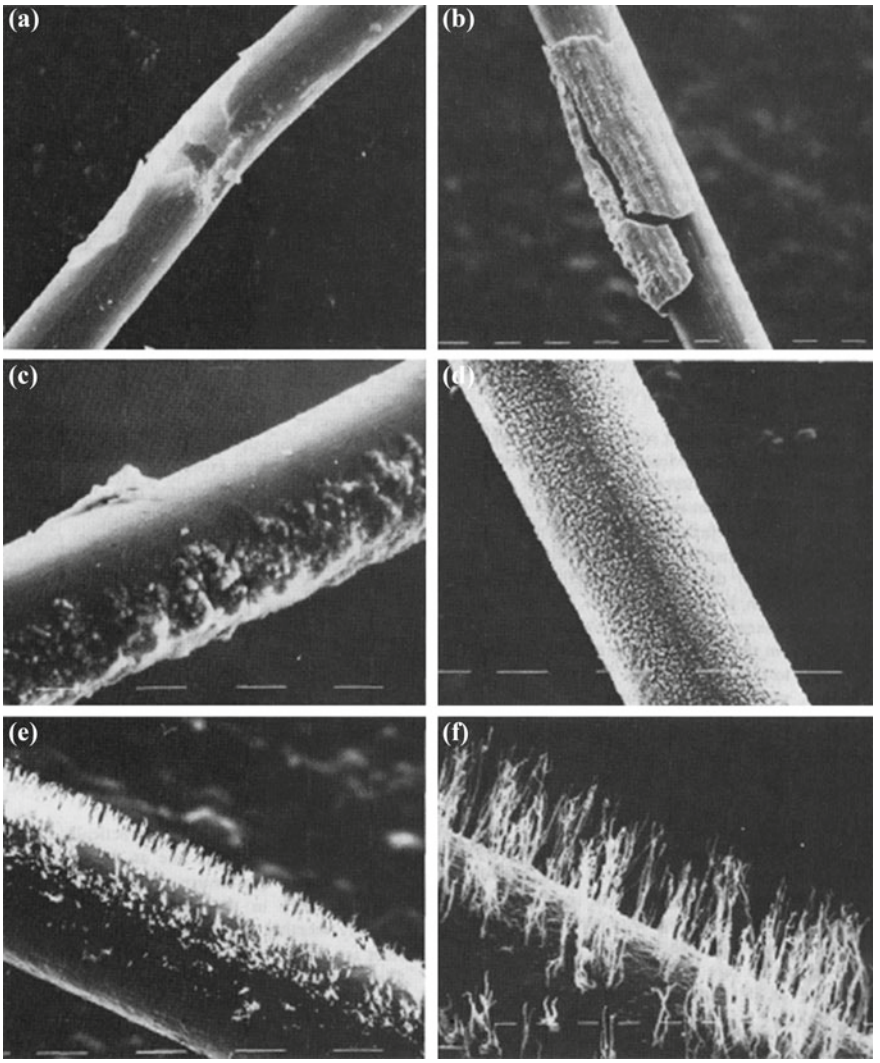


Fig. 7.12 Deposits on several test anode wires of 30 μm diameter, after prolonged irradiation. Reprinted from [6] with permission

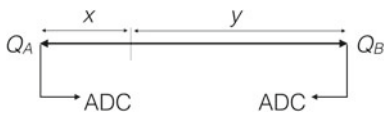


Fig. 7.13 Principle of charge division used to locate an ionisation event along the direction parallel to the anode wire in a proportional chamber

7.6 Multi Wire Proportional Chamber

A Multi Wire Proportional Chamber (MWPC) uses the proportional chamber principle to achieve particle trajectory tracking. The principle consists of arranging an array of many closely spaced wires which act each as an independent proportional chamber. By placing three or more panels of MWPC one can achieve a multiple point particle trajectory reconstruction. The single point resolution depends on the wire spacing, namely about 1/2 the wire spacing. The position can also be reconstructed using the center of gravity of amplitude pulses. Figure 7.14 shows an array of anode sense wires sandwiched by two cathode planes. The orientation is such that an incident particle will cross the planes and travel through the wires. The electric field lines run perpendicular to the cathode planes except for the region in the vicinity of the wires. Note that the mechanical precision of the wires is essential in order to achieve a uniform field configuration. Wire spacings are typically 2–5 mm and are limited to no less than 1 mm due to mechanical and electrostatic forces acting on the wires. Cathode planes are placed at a distance of about 10 mm. A possible choice of gas mixture for these chambers could be Ar (75%, appropriate for its low field for avalanche formation, low electron affinity and relatively low cost), Isobutane (24.5%, a photon quencher) and Freon (0.5%, to suppress the formation of polymer) which allows gains of about seven orders of magnitude. The timing resolution of the chamber will depend on the drift time of the electrons from the original ionisation position to reach the point of starting of the avalanche (a path that takes about 25–30 ns for fast gases).

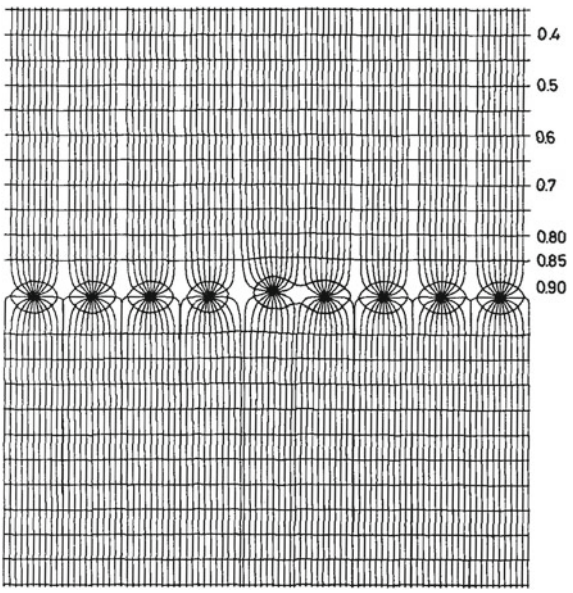


Fig. 7.14 Electric field and equipotential lines in a multiwire proportional chamber. The numbers indicate the relative potential and one wire has been displaced by 10%. Reprinted from [7] with permission

7.7 Multi Strip Gas Chamber and Resistive Plate Chamber

Since there are limitations in the MWPC on the wire spacing due to mechanical and electrostatic forces, devices called Multi Strip Gas Chamber (MSGC) have been developed to remove the need for wires. The anodes and cathodes are not wires but strips glued to an insulating material, and with this technique the spacing can be significantly reduced to about 100 μm (see Fig. 7.15). Also, by placing the cathode strips next to the anodes, the drift distance is substantially reduced, minimising the drift time of the ions.

A parallel technology to the MWPC and MSGC is the resistive plate chamber (RPC), initially developed in the 1980s [8]. While the working principle is the same of a gaseous detector, the electrodes are planes made of a mixture of phenolic resins and kept separated by spacers. The Fig. 7.16 shows the structure of a RPC in the ATLAS experiment. The gap between the resistive plates is of the order of 2 mm but can be reduced to 200–300 μm. High voltage in the gas volume causes drift and avalanche formation under the effect of a traversing charged particle. The signal is induced on pick-up copper strips on the outside of the resistive plates. Readout strips can be arranged in orthogonal configurations on back to back chambers for bi-dimensional position reconstruction. The timing resolution that can be achieved by these chambers is of the order of nano seconds, but RPC can reach 100 ps. Recent successful technologies include also Micro-MEsh Gaseous Structure (MicroMegas) [9] and gas electron multiplier (GEM) [10] detectors.

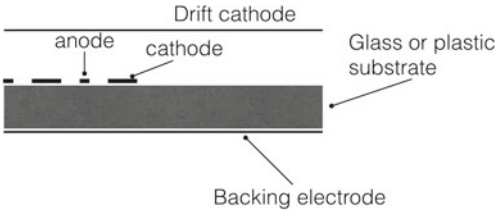


Fig. 7.15 Skematics of a multi strip gas chamber. Charges are collected on the strip cathodes; the spacing between anodes can be of 100–200 μm and the spacing that forms the drift volume is of a few mm

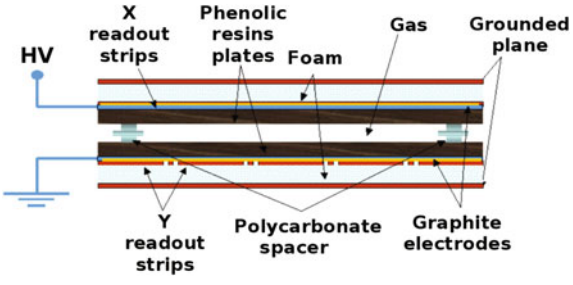


Fig. 7.16 The structure of a Resistive Plate Chamber in the ATLAS experiment at the Large Hadron Collider. Reprinted from [11] with permission

7.8 Drift Chamber

Early in the development of the MWPC it was realised that spatial information can be also obtained by measuring the drift time. If the drift velocity in the gas is known and constant, then the distance travelled by the electrons or ions in a time interval Δt is simply:

$$x = v_d \Delta t, \tag{7.26}$$

where Δt can be set as the interval between a trigger signalling the incoming radiation particle and the time of appearance of the electric pulse from the chamber. The (7.26) is known as the time to distance relationship and is linear as long as v_d is constant across the drift region. The (7.26) also implies that this detector requires an external trigger, however there are clear advantages in terms of the position resolution. Typical drift regions span a few centimeters, and with drift speeds of about 5 cm/μs the drift time is of the order of 1–2 μs. The structure of a drift chamber is similar to that already described of MWPC, but particular attention must be paid to the electric

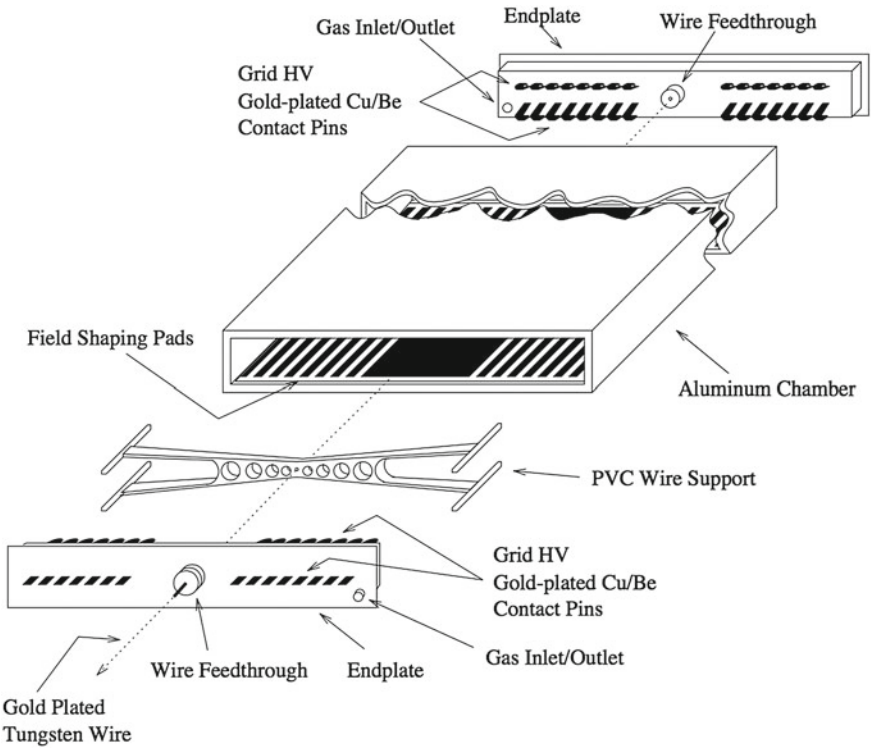


Fig. 7.17 Schematic view of a drift tube of the CDF experiment, used for the outer part of the muon system. Credit CDF Collaboration. Reprinted from [12] with permission

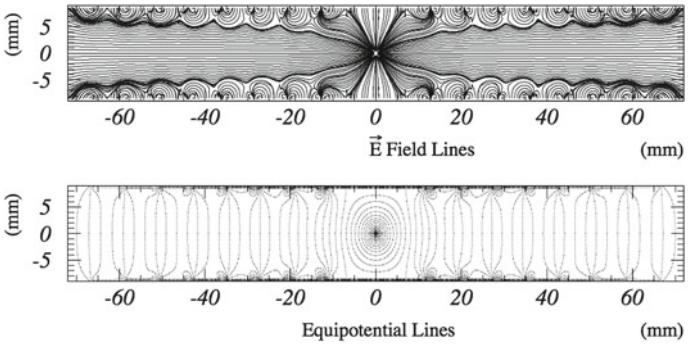


Fig. 7.18 Equipotential and electric field lines of the tube shown in Fig. 7.17. Credit CDF Collaboration. Reprinted from [12] with permission

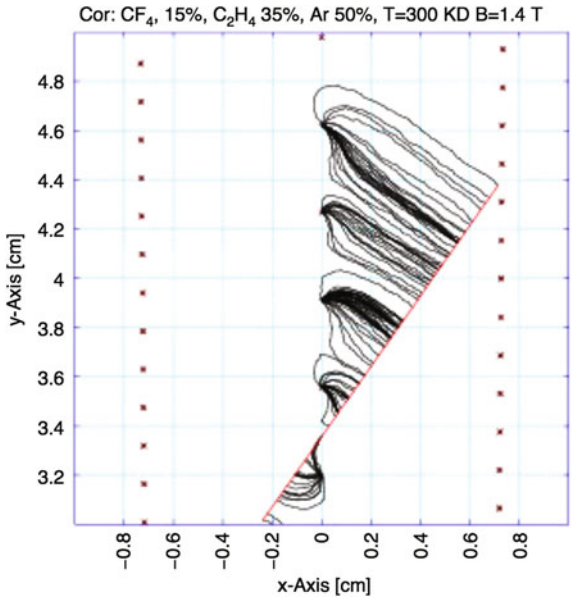


Fig. 7.19 Simulated drift trajectories for a radial (high momentum) track of the central tracker for the CDF experiment. Reprinted from [13] with permission

field distribution in the chamber. Low field regions need to be removed, and that is achieved with the addition of special extra wires, called *field wires*, whose goal is to adjust the electric field across the drift region. The choice of fill gas is largely similar to that used in MWPC, but the drift properties of the gas are very important. For instance, carbon tetrafluoride, CF_4, added to the gas mixture is known to allow higher drift speeds, but its environmental impact must also be considered. Spatial resolution of the order of tens of microns is routinely obtained in these chambers. The Fig. 7.17 shows a single-wire drift tube used for the CDF experiment at the Fermi National Accelerator Laboratory (1985–2011), and Fig. 7.18 shows the field

Fig. 7.20 The Central Outer Tracked of the CDF experiment. Reprinted from [13] with permission

and equipotential lines in the same tube. The chamber achieved a position resolution of 300 μm.

For operation within a magnetic field, which is often the case for particle detectors which measure the momentum of charged particles, the drift chamber must be designed taking into account that the drift of the ionisation charges will not follow the lines of the electric field but form an angle to it (the *Lorentz angle*), under the influence of the magnetic field. The design of detectors operating in such conditions accounts for the direction of the drifting charges. The Fig. 7.19 shows an element of the CDF Central Outer Tracker drift chamber. The sense wires run vertically in the center of the picture, and the drift trajectories are simulated from a high momentum track traversing the sensitive volume. The drift trajectories are set to be perpendicular to a track by tilting the drift chamber element, in this case with a Lorentz angle of 35°. The picture in Fig. 7.20 is the CDF Central Outer Tracker during assembly of the wires. Wire planes are visible as the tilted segments arranged as subsequent circular barrels. The chamber, equipped with 30,240 sense wires, achieved a position resolution of 150 μm with 96 radial point measurements, using a gas mixture of argon (50%), ethane (50%) and a small amount of isopropyl.

7.9 Time Projection Chamber

The Time Projection Chamber (TPC) uses the principles of both a MWPC and a drift chamber to achieve three-dimensional track reconstruction. An example is shown in Fig. 7.21 with the Time Projection Chamber of the ALICE experiment [14]. The chamber is arranged over two halves. In the middle of the chamber, a high voltage plane creates electric field lines that run from both endcaps to the central plane. Drifting electrons move towards the endcaps, where cathode pads and sense wires allow the reconstruction of a bidimensional projection of the ionising particle trajectory. As the drift time is proportional to the position of ionisation along the vertical axis, this allows for a three dimensional trajectory reconstruction. The presence of a magnetic field provides bending of the charged particles, and therefore momentum measurement, and reduces diffusion. Furthermore as the amount of charge collected is proportional to the energy deposited, these detectors can also be used for particle identification. This particular chamber for ALICE has a length of 5 m with a total drift time of 92 μs and is operated with a low diffusion (cold) gas mixture of Ne-CO_2-N_2 achieving a position resolution of 0.2 mm.

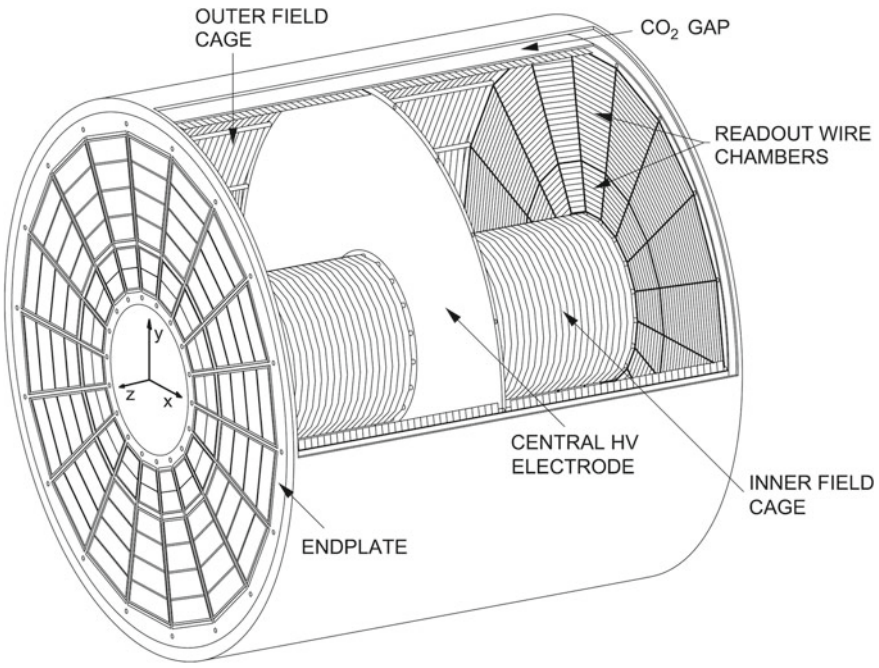


Fig. 7.21 The field cage of the Time Projection Chamber of the ALICE experiment. Reprinted from [14] with permission

7.10 Liquid Ionisation Detectors

The ionisation of liquid noble gases for detector applications has been exploited in recent years for many applications, for precision particle tracking and calorimetry, in particle physics, dark matter searches, γ-ray astronomy, to name just a few. Liquid noble gases possess a high density, and therefore good energy absorption, high spatial resolution due to low diffusion and the possibility of large volumes at relatively low cost. Furthermore, they combine ionisation and scintillation features, the latter of which will be the subject of Chap. 9. The most widely used active medium in this class today are liquid Xenon and liquid argon and there are some very good reviews of low energy applications of liquid ionisation detectors [15, 16]. An example of detectors currently or recently in operation is the liquid argon (LAr) calorimeter [17] of the ATLAS detector and several dark matter experiments (e.g. the Large Underground Xenon, and the ZoneED Proportional scintillation in Liquid Noble

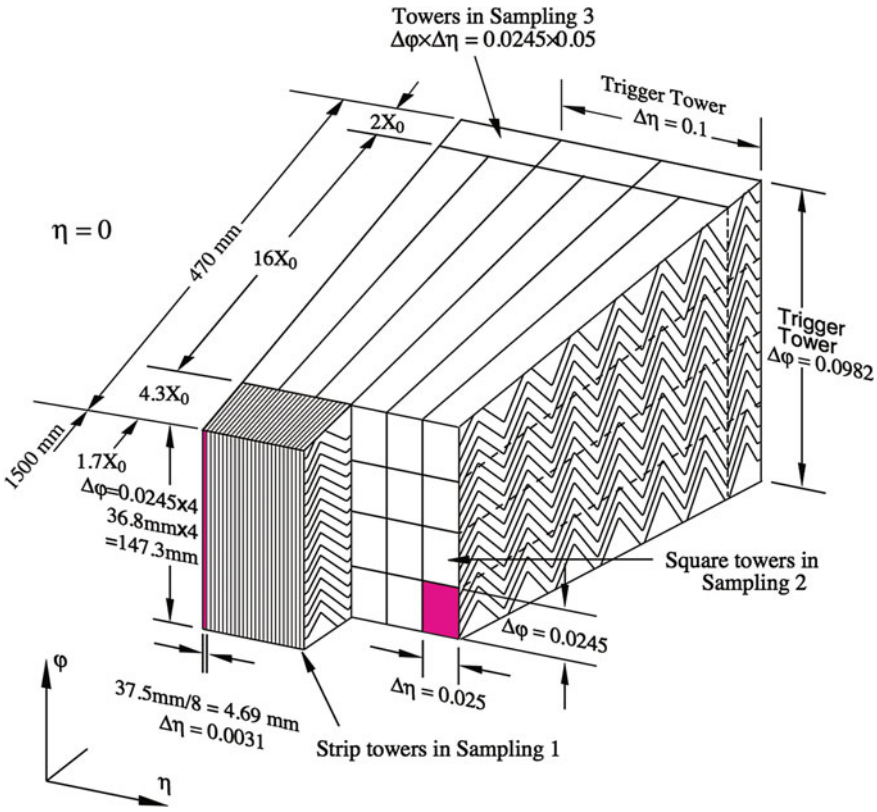


Fig. 7.22 Sketch of the accordion structure of the ATLAS liquid argon calorimeter. Reprinted from [17] with permission

gases experiment). The ATLAS LAr is a sampling of accordion-shaped electrodes filled with liquid argon as active material. The sketch in Fig. 7.22 shows the structure of the detector. The charges liberated by ionisation in the argon will drift within a time of 450 ns towards the electrodes. The use of liquid argon requires a cryogenic system operating at a temperature of 87 K, with contamination kept below 250 parts per billion. The function and design of calorimeters will be presented in detail in Chap. 10.

Glossary

Charge division Operating mode of a proportional chamber, e.g. in a single wire device, that allows the measure of the longitudinal coordinate via charge collection at both ends of the anode

Dead time Time interval over which the detector is not capable of forming a response to an event

Drift chamber Device based on the proportional chamber principle plus the use of a time to distance relationship to locate the distance of ionisation from the anode

Drift velocity Average velocity of electrons and ions after ionisation under the effect of an external electric field

Ionisation chamber Ionisation-based detector operating with unit-gain

Ionisation detector An instrument, based on the ionisation of the gas or the liquid forming the sensitive volume, capable of measuring the presence and possibly some of the properties of incident radiation

Mean free path of ionisation Average length between subsequent ionisation processes:

$$\lambda_I = \frac{1}{n_e \sigma_I},$$

where n_e is the density of electrons and σ_I is the ionisation cross section

MSGC Multi Strip Gas Chamber, a device based on the proportional chamber principle and constructed using anode strips rather than wires to reduce the pitch of the anodes sequence

MWPC Multi Wire Proportional Chamber, a device based on the proportional chamber principle and constructed placing a sequence of anode wires to achieve particle trajectory tracking

Primary ionisation Electrons freed by the direct interaction of an incident radiation particle with an atom or molecule

Proportional chamber Ionisation-based detector operating with high gain, typically with multiplicative factors of order 10^6–10^7

Recombination Process that leads to the attachment of an electron and an ion to form a neutral atom

Resolution Ratio between the full width at half maximum of the response to mono-energetic (E) incident radiation and the value of E:

$$\text{Resolution} = \frac{\Delta E}{E},$$

Response Relationship between the amount of released ionisation charge and the incident energy

Response distribution Distribution of the response when the detector is subject to monochromatic (i.e. single energy) radiation

Response time Time interval that the detector requires to form the response signal

RPC Resistive Plate Chamber, a device based on the proportional chamber principle and constructed with two planes of a mixture of phenolic resins kept at high potential voltage difference The signal is induced on pick-up copper strips on the outside of the resistive plates

Secondary ionisation Electrons freed as a result of atomic de-excitation or ionisation of nearby atoms by electrons from primary ionisation

References

1. A.C. Melissinos, *Experiments in Modern Physics* (Academic Press, New York, 1955)
2. An ionisation smoke alarmo from Ei Electronics. Credit Hiya111 at English Wikipedia, https://en.wikipedia.org/wiki/Smoke_detector#/media/File:Smokealarm.JPG. Accessed 20 Dec 2016, Public Domain
3. M.E. Rudd, R.D. DuBois, L.H. Toburen, C.A. Ratcliffe, T.V. Goffe, Phys. Rev. A **28**, 3244 (1983)
4. T. Zhao, Y. Chen, S. Han, J. Hersch, Nucl. Instrum. Methods Phys. Res. Sect. A **340**(3), 485–490 (1994)
5. G.F. Knoll, *Radiation Detection and Measurement* (Wiley, New York, 2010)
6. J. Adam et al., Nucl. Instrum. Methods Phys. Res. **217**(1–2), 291–297 (1983)
7. G. Charpak, D. Rahm, H. Steiner, Nucl. Instrum. Methods, **80**(1), 1334, 1 April 1970
8. R. Santonico, R. Cardarelli, Development of resistive plate counters. Nucl. Instrum. Methods A **187**(2–3), 377–380 (1981)
9. Y. Giomataris, Ph Rebourgeard, J.P. Robert, G. Charpak, Nucl. Instrum. Methods Phys. Res. Sect. A **376**(1), 29–35 (1996)
10. F. Sauli, Nucl. Instrum. Methods Phys. Res. Sect. A **386**(23), 531–534 (1997)
11. G. Cattani, A.T.L.A.S. Muon Collaboration, Nucl. Instrum. Methods Phys. Res. Sect. A **661**(Supplement 1), S6–S9 (2012)
12. S. Bertolucci et al., CDF Collaboration Public Note CDF/DOC/MUON/PUBLIC 6362 (2002)

13. T. Affolder et al., CDF Central Outer Tracker. Nucl. Instrum. Methods Sect. A **526**(3), 249 (2004). doi:10.1016/S0168-9002(01)
14. J. Alme et al., ALICE TPC collaboration. Nucl. Instrum. Methods Sect. A **622**(1), 316–367 (2010)
15. E. Aprile, A. E. Bolotnikov, A. I. Bolozdynya, and T. Doke, *Noble Gas Detectors* (Wiley-VCH Verlag, Weinheim, 2006)
16. E. Aprile, T. Doke, Rev. Mod. Phys. **82**, 2053 (2010)
17. ATLAS Collaboration, ATLAS liquid-argon calorimeter: Technical Design Report, ATLAS-TDR-2 ; CERN-LHCC-96-041 (1996) — Fig 1.2

Chapter 8
Semiconductor Detectors

Abstract Semiconductor detectors work, broadly speaking, using the same principles of gaseous and liquid ionisation detectors, but the active medium is denser and in solid state and the ionisation potential of semiconductors is about 10 times smaller than in gases. Semiconductor-based devices are the detectors of choice for precision vertex and tracking detection with high spatial resolution, and for energy measurements with high resolution in high energy physics.

8.1 Basic Semiconductor Properties

In solid state matter, the atomic electrons can populate a number of allowed energy levels. The energy values which are closely spaced can be visualised as energy bands and in materials with insulator or semiconductor electrical properties, electrons which are bound to lattice sites are said to belong to the *valence band*, whereas electrons that can migrate through lattice atoms are said to belong to the *conduction band*. A sketch of energy bands is shown in Fig. 8.1. In the absence of thermal excitation ($T = 0\,\mathrm{K}$), the valence band is completely filled and the conduction band is completely empty (i.e. there is no electrical conduction). Electrons in the conduction band can move freely under the influence of an external electric field, whereas they cannot do so when in the valence band. However, insulators and semiconductors differ by how large the energy gap is. Semiconductors are materials for which the energy gap between the valence and the conduction bands is sufficiently small (up to a few electron volts) to allow some electrons to reach the conduction band under normal thermal excitation. Therefore, the size of the energy gap is crucial to determine the electrical properties of the material.

The most commonly used elemental semiconductor materials for particle detectors are silicon (Si) and germanium (Ge). The energy gap at a temperature of 0 K is 1.16 eV for Si and 0.75 eV for Ge. Diamond would also be suitable, and has a larger band gap than either Si or Ge, but it is also very expensive. Silicon is instead relatively inexpensive and abundant, and at room temperature has a sufficiently large energy gap compared to thermal energy to only require cooling for very low-noise applications,

L. Cerrito, *Radiation and Detectors*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-53181-6_8

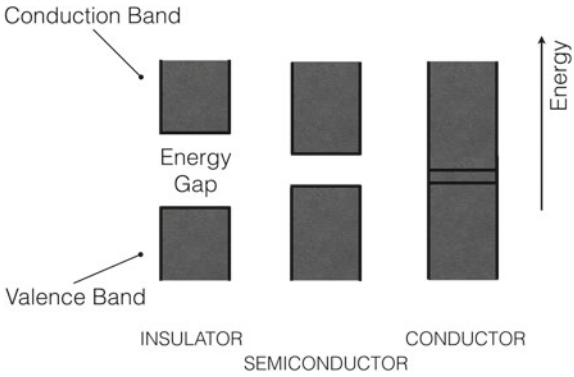


Fig. 8.1 Sketch of the arrangement of electron energy levels in cristalline materials

Fig. 8.2 Photograph of a piece of purified silicon [1]

and to mitigate radiation damage. The photograph in Fig. 8.2 shows a piece of purified silicon.

The atomic configuration of silicon has four electrons in the outermost shell. In a silicon crystal, which is a cubic crystal structure, those electrons form covalent bonds with four neighbouring Si atoms. Covalent bonds are chemical bonds in which pairs of electrons are shared between atoms. At non-zero temperature, electrons can be excited into the conduction band, leaving their valence energy levels vacant (called *holes*). A pair of one electron and one hole can be thought of as the analogous of the electron-ion pair in gases; for instance the electron and hole can be made to drift by the application of an electric field. The sketch in Fig. 8.3 shows a planar representation of a pure silicon lattice with temperature $T > 0$.

At any given temperature there is an equilibrium in the number of free electrons and holes; the concentration n_i of electrons in the conduction band is given by the expression

$$n_i = AT^{3/2} e^{-E_g/2kT}, \tag{8.1}$$

where A is a proportionality constant, E_g is the gap energy at the temperature of $T = 0\,\mathrm{K}$, k is Boltzmann's constant, and T is the semiconductor temperature expressed in Kelvins. Typical concentrations of electrons, or equivalently of holes,

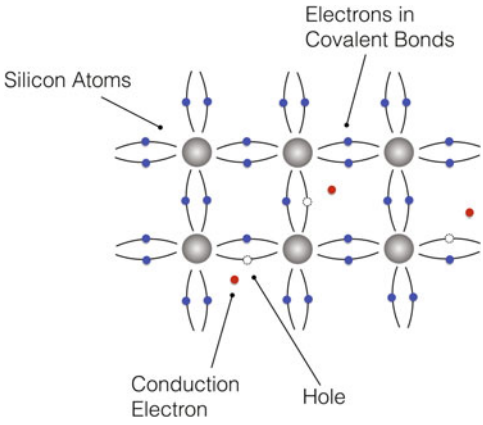


Fig. 8.3 Planar representation of the cubic lattice structure of a pure silicon crystal, indicating the covalent bonds of the four outermost-shell electrons in each silicon atom and the formation of electro-hole pairs when one of the electrons is liberated from its bond

at 300 K are 10^{10} to 10^{13} per cubic centimeter. Note that the typical atomic concentrations in these materials are 10^{22} atoms per cubic centimeter, so the electron-hole concentrations are relatively low even without cooling.

If the electron-hole pairs were all created in one point, they would spread with the passing of time according to a Gaussian distribution whose width σ can be parametrised as

$$\sigma \sim \sqrt{2\,D\,t}, \tag{8.2}$$

where D is the diffusion coefficient that depends on the mobility of the electrons and holes and on the temperature. Under the action of an externally applied electric field, the electrons and holes start *drifting* with an average velocity

$$v_e = \mu_e(E, T) \times E, \tag{8.3}$$

$$v_h = \mu_h(E, T) \times E, \tag{8.4}$$

where μ_e, μ_h are mobility functions (which are specific for electrons and holes respectively) dependent on both the electric field E and the temperature T. The mobility functions are approximately constant for electric fields up to 1 kV/cm and then decreasing for higher fields with an inverse proportionality to E, $\mu \sim 1/E$. This means that the drift velocities grow linearly with E and then tend to saturate to a constant value. The graphs in Fig. 8.4 show the drift speeds for electrons and holes in silicon as a function of the applied electric field for two different temperatures. The drift velocity is of the order of 1–10 cm/μ s, with holes about a factor of two slower for the same electric field intensity. The saturation occurs as higher fractions of energy are taken away by the collisions with the lattice atoms.

The functional dependence of the mobility of electron and holes with temperature is approximately

$$\mu \sim T^m, \tag{8.5}$$

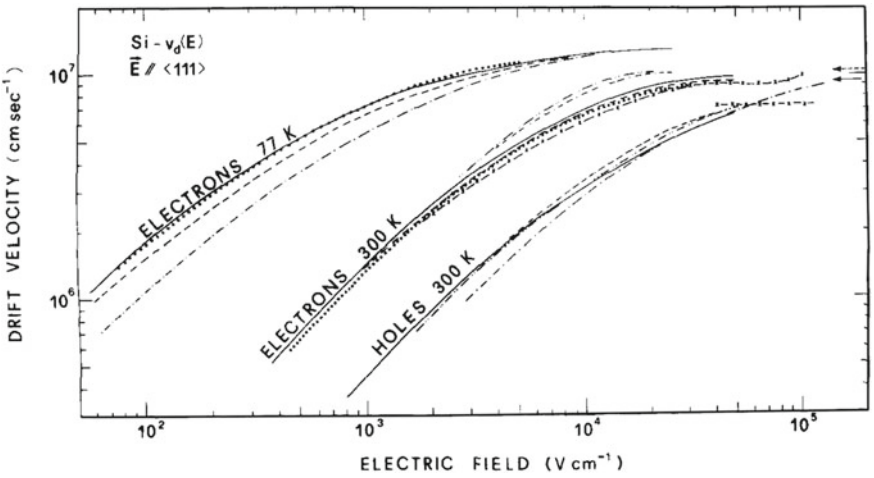


Fig. 8.4 Experimental results of electron- and hole-drift velocities as functions of electric field applied to a ⟨111⟩ crystallographic direction. Reprinted from [2] with permission

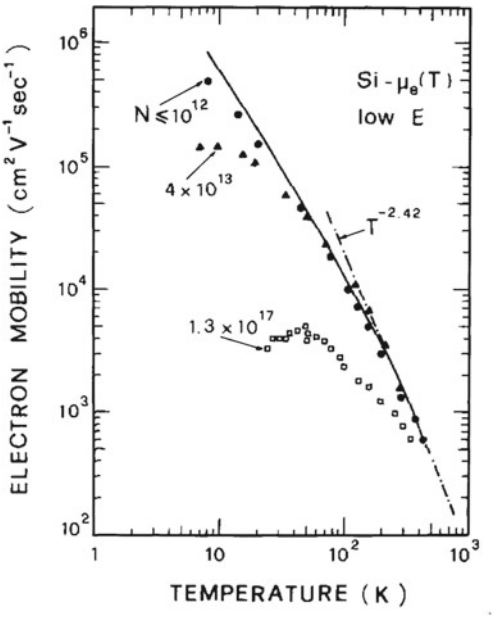


Fig. 8.5 Ohmic mobility of electrons in silicon as a function of temperature. Reprinted from [2] with permission

with $m \approx -2$. The plot in Fig. 8.5 shows the ohmic mobility of electrons as a function of temperature, with a fit to data yielding $m = -2.42$ at around room temperature. A similar curve can be obtained for holes [2].

If an electric field is applied to a silicon crystal, the current density (J, measured in A/m^2) can be written as:

$$J = \rho v, \tag{8.6}$$

where ρ is the charge density and v the average speed of charges, and therefore combining (8.4) and (8.6) gives

$$J = \rho v = e n_i \left(\mu_e + \mu_h\right) E. \tag{8.7}$$

Since the current density J can also be expressed as a function of the conductivity σ (or also the resistivity $\mathrm{R} = 1/\sigma$) and of the electric field E, as

$$J = \sigma E, \tag{8.8}$$

then comparing (8.7) and (8.8) yields the conductivity of silicon expressed in terms of the charge carriers and their mobility,

$$\sigma = e n_i (\mu_e + \mu_h). \tag{8.9}$$

Intrinsic resistivity of silicon at 300 K is of 2300 Ω m, and will depend strongly on the presence of *impurities* in the silicon lattice. Impurities refers to different atoms positioned in the silicon lattice. The presence of impurities adds energy levels within the band gap and these additional energy levels can trap electrons, they can speed up the recombination of electrons and holes or even block it, with a typical associated reduction in the number of free charges.

Example: Can pure silicon be used as a detector medium?

Pure silicon will not work as a detector medium because the number of charges liberated by radiation is too low compared to the intrinsic charge carriers. This can be estimated rather simply through the following considerations. Silicon has a mean ionisation energy (I_0) at $T = 300$ K of approximately 3 eV and suppose it is traversed by a heavy charged particle releasing energy at the rate of the minimum ionisation point, $\langle (1/\rho) dE/dx \rangle = 3\,\mathrm{MeV/cm^2 g}$. For a silicon wafer of surface area $A = 1\,\mathrm{cm}^2$, thickness $d = 300\,\mu\mathrm{m}$ and density $\rho = 2\,\mathrm{g/cm^3}$ the number of electron-hole pairs liberated by the ionising radiation can be estimated as:

$$N_{\mathrm{e-h\,pairs}}^{\mathrm{ionisation}} = \frac{\langle (1/\rho) dE/dx \rangle \rho_s\, d}{I_0} = \frac{3\,\mathrm{MeV} \times 2 \times 300 \times 10^{-4} \mathrm{cm\,cm^{-1}}}{3\,\mathrm{eV}} = 6 \times 10^4. \tag{8.10}$$

The number of pairs obtained from (8.10) is to be compared to the number of intrinsic carriers (with a density n_i of about 10^{10} per cubic centimeter, where the i index stands for intrinsic) in the same detector volume, assuming a temperature of $T = 300$ K:

$$N_{\mathrm{e-h\,pairs}}^{\mathrm{intrinsic}} = n_i\, d\, A = 3 \times 10^8. \tag{8.11}$$

Notice therefore that the number of thermally created electron-holes is four orders of magnitude larger than the charge created by the ionising radiation. It is clear from

these estimates that in order to make a viable semiconductor detector, the number of freely moving charge carriers must be reduced significantly.

8.2 Doped Semiconductors

Silicon is tetravalent, and in the normal crystalline structure it forms covalent bonds with the four nearest silicon atoms. A silicon crystal can be doped with pentavalent atoms, such as phosphorus to form what is called a *n-type* semiconductor. The addition of phosphorus in the crystal structure results in an orphan electron located at an energy level just below the conduction band (see Fig. 8.6), which under thermal excitation can easily reach the conduction band. The energy difference between these impurity levels and the conduction band is as small as 0.01 eV (in silicon) or 0.05 eV (in germanium). Concentrations of dopant atoms of only one per million will result in a donor density of about $10^{17}/\text{cm}^3$ which in turns means there are now potentially several orders of magnitude more free electrons than the regular thermal concentration of the pure semiconductor. When all of the weakly bound donor electrons are excited, the concentration of free electrons (n) at moderate temperature is dominated by the concentration of donor electrons (N_D), therefore $n \simeq N_D$. This is called the *extrinsic region*.

Furthermore, the number of free holes (p) diminishes due to the abundance of electrons in such a way that, in thermal equilibrium

$$n \times p = n_i \times p_i, \tag{8.12}$$

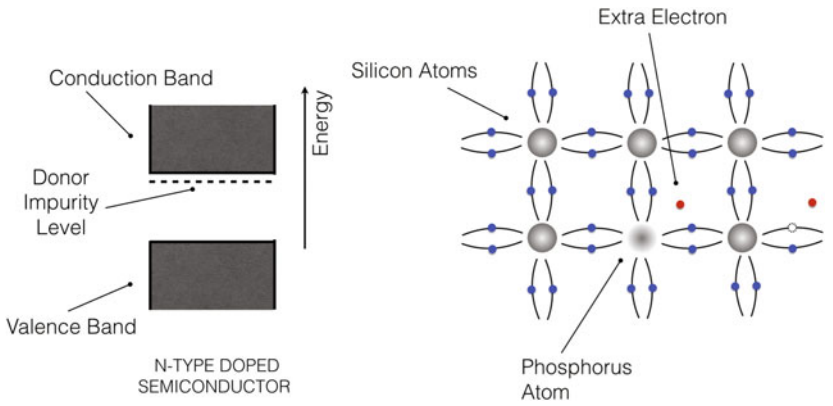


Fig. 8.6 Planar representation of the cubic lattice structure of a pure silicon crystal with a phosphorus atom, indicating the weakly bound orphan electron resulting from the pentavalent configuration of phosphorus

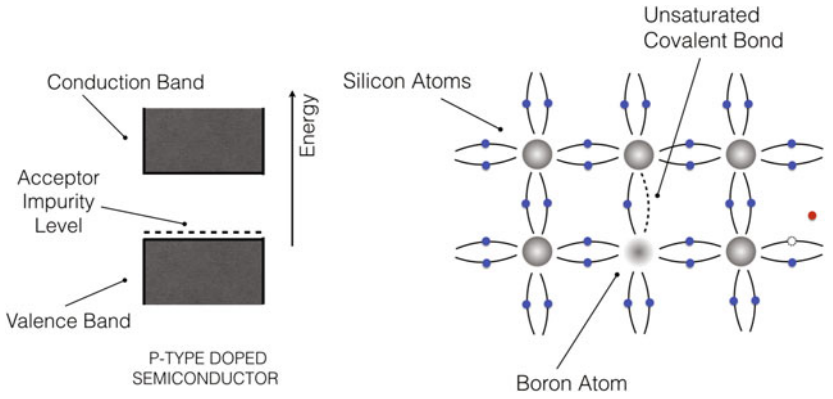


Fig. 8.7 Planar representation of the cubic lattice structure of a pure silicon crystal with a boron atom, indicating the unsaturated covalent bond resulting from the trivalent configuration of boron

where n_i and p_i are the intrinsic electron and holes concentration (without dopant). This is called the law of Mass Action. For a concentration of free electrons (n) of $10^{17}/\text{cm}^3$ and intrinsic concentrations $n_i = p_i = 10^{10}/\text{cm}^3$ it follows that $p \sim 10^3/\text{cm}^3$. Note that the total number of charge carriers is much greater than the intrinsic number, and that in a n-type doped semiconductor the electrical conductivity is due almost exclusively to electrons. Other suitable n-type dopants are the elements in the 15th group of the periodic table, such as arsenic and antimony.

If the doping of the silicon lattice is performed with a trivalent impurity, such as boron, for each boron atom there will be one covalent bond that remains unsaturated. Such configurations determine a vacant energy level (acceptor) located just above the valence band, as shown in Fig. 8.7 and the semiconductor is called *p-type* doped. Electrons in the crystal will fill the acceptor vacant energy level, leaving holes in the crystal's electronic structure. To a good approximation, the concentration of holes (p) is equal to the concentration of acceptors (N_A). In this case also, (8.12) implies that the product of the concentrations of free holes and electrons equals the product of intrinsic concentrations and therefore the majority carriers in a p-type doped semiconductor are holes. Other suitable p-type dopants are the elements in the 13th group of the periodic table, such as aluminum and gallium.

8.3 p-n Semiconductor Junction

A p-n junction consists of positioning in contact of a p-type and a n-type semiconductor, as shown in Fig. 8.8. In practice, a single semiconductor is doped differently in two parts. The n-type side has an excess of electrons which migrate to the p-type side to fill the excess of holes. The result is the build up of net electric charge in the vicinity of the junction which creates an electric field opposing further diffusion of

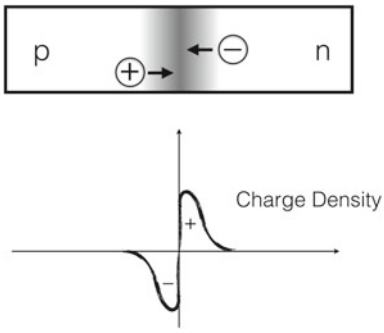


Fig. 8.8 Schematics of a p-n junction in a semiconductor. Negative charges will move across to fill vacancies and will leave behind an excess of holes. The charge density acquires a non neutral configuration on each side across the junction, in the depletion zone (indicated in *shaded grey*)

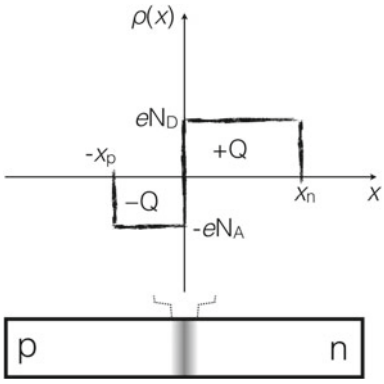


Fig. 8.9 Depth of the depletion zone in a p-n semiconductor junction

charges from one side to the other. The physical zone under these conditions is called a *depletion* zone, because of the relative absence of free charges. In the depletion zone any free electron or hole that might appear is swept aside by the electric field.

The characteristics of a depletion zone are particularly attractive for radiation detectors: ionising radiation will liberate electron-hole pairs in the semiconductor which are then drifting under the influence of an electric field. However, the width of the depletion zone is generally small. Such width depends on the dopant concentrations and can be estimated as follows; let x indicate the distance from the contact point between the n-type and p-type semiconductors, and suppose the charge density $\rho(x)$ can be approximated by a constant within the depletion zone (as shown in Fig. 8.9), hence:

$$\rho(x) = \begin{cases} +eN_D, & \text{for } 0 < x < x_n \\ -eN_A, & \text{for } -x_p < x < 0 \end{cases} \tag{8.13}$$

where N_D and N_A are the donor and acceptor impurity concentrations, respectively, and x_n (x_p) is the extension of the depletion zone on the n-side (p-side). Since the total charge is conserved, it follows that

$$N_A\, x_p = N_D\, x_n. \tag{8.14}$$

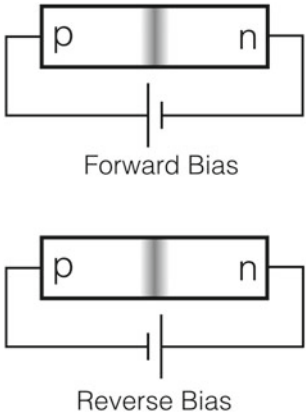


Fig. 8.10 Forward and reverse bias in a p-n semiconductor junction

Notice therefore that the depletion zone is more extended in the area with lower dopant concentration. If the charge density is known, Poisson's equation

$$\frac{d^2V}{dx^2} = -\frac{\rho(x)}{\varepsilon}, \tag{8.15}$$

can be used to relate the potential (V) to the charge density, with ε the dielectric constant. With two integrations of (8.15), the depletion depth (d) is expressed by

$$d = x_n + x_p = \left(\frac{2\varepsilon V_0}{e}\frac{N_A + N_D}{N_A N_D}\right)^{1/2}, \tag{8.16}$$

where V_0 indicates the contact potential, approximately 1 V. Typical dimensions of the depletion zone are of the order of 10^{-6} meters, and therefore despite this configuration being intuitively attractive as a sensitive detector volume, the performance of such detector would be quite low.

The depth of the depletion zone is rather small, but application of an appropriate external potential across the p-n junction makes this device a diode and one of the most useful detectors. In forward bias, that is with the positive terminal connected to the p-side as shown in Fig. 8.10, current will flow across the device. In reverse bias, with the positive terminal connected to the n-side, the electrons will be pulled away from the junction and similarly for the holes on the p-side from the negative terminal, hence *widening* the depletion zone. Since the resistivity of the depletion region is much higher than for the p-side and n-side alone, the applied voltage (V_{bias}) will appear across the depletion region. The same (8.16) applies to a reverse bias junction, but V_0 is replaced by V_{bias}. In practice, a small *leakage* current will flow across the junction when in reverse bias and care must be taken in the manufacture of the device to minimise such effect. Depletion depths of the order of millimeters can be obtained by applying a reverse bias of the order of a few hundred volts.

8.4 Silicon Detector Configurations

The depletion zone in a semiconductor p-n junction represents a volume in which new electron-hole pairs, liberated by radiation interacting with the semiconductor material, can be readily collected at the terminals, much like a gaseous ionisation detector. The average ionisation energy for an electron-hole pair is around 3 eV, and therefore about ten times lower than for gases, yielding two advantages: lower energy uncertainty due to a smaller fraction of energy per carrier, and better signal to noise ratio in low signals. The induced charge accompanying the drifting of electrons and holes is similar to the mechanism shown for the ionisation chamber, however the mobility of electrons and holes in semiconductors is within a factor of two to three of each other rather than orders of magnitude different as in the case of electrons versus ions in gases.

The simplest semiconductor detector configuration is a Pad, that is a large surface diode. The Fig. 8.11 shows the silicon pad [3] of the PHOBOS experiment at the Relativistic Heavy Ion Collider (RHIC). The detector consists of a p-n junction formed of a p+ doped pad and a n-type substrate. An oxide layer covers the pad and aluminum contacts are made both on the pad and on the back side of the plate. The n+ layer improves the ohmic contact. A protection ring around the pad is designed

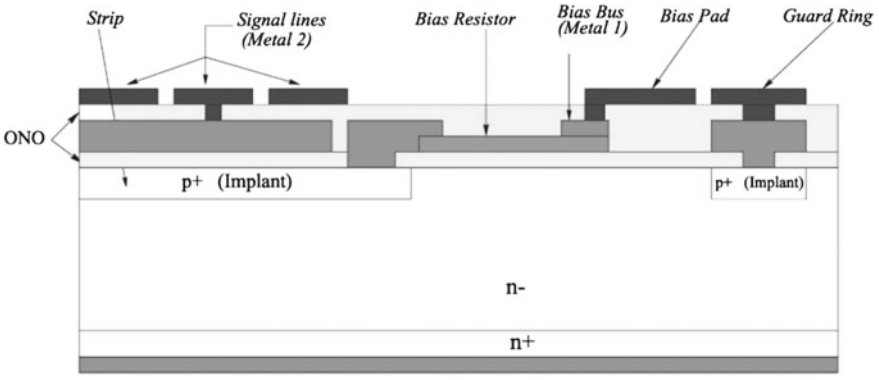


Fig. 8.11 Schematic diagram of the cross section of the silicon pad for the PHOBOS experiment at RHIC. Reprinted from [3] with permission

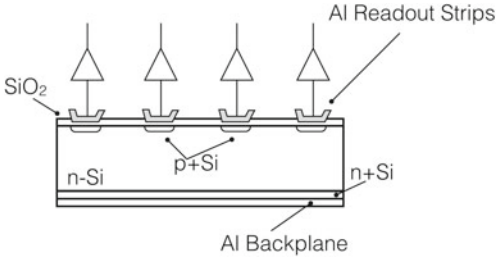


Fig. 8.12 Schematic working principle of a silicon microstrip detector. The segmentation in the p-n junctions allows for signal collection on a small number of strips

to build up a depletion zone and make a sharper boundary of the active area, leading to a decrease in leak currents and noise. No position resolution is possible within a single pad but the use of multiple pads allows particle tracking.

Alternatively, silicon microstrip detectors are designed for good position resolution. The schematic in Fig. 8.12 shows the working principle of an AC-coupled silicon microstrip detector. The holes liberated from radiation crossing the sensitive area drift towards the p+ strips, while electrons move towards the n+ backplane. Charges on the doped strips induce a signal via capacitive coupling on the aluminum readout directly connected to a charge amplifier. A thin layer (about 100–200 nm) of SiO_2 is placed between the aluminum contact and the p+ strips to avoid leakage current from the amplifier. Typical values for sensor thickness are around 300 μm, with n-bulk resistivity of 2 kΩ and a bias voltage (applied through a separate circuit, not shown) of up to approximately 200 V, while the strip pitch is optimised for position resolution and signal collection. An extensive overview of silicon microstrip detectors can be found in [4].

If the measurement's objective is a two dimensional position hit reconstruction, then a second coordinate can be obtained by mounting a second microstrip detector under the first with the strips non aligned with the first detector and forming an opening angle between the two, called a *tilt*. A tilt of a few degrees is typically sufficient but this also means doubling the material, which is a factor in precision particle detection. A more elegant solution, which also reduces material, is to arrange the differently oriented strips on the back of the same bulk with both a n+ and p+ segmentation, to form a double sided microstrip detector as the one shown in Fig. 8.13

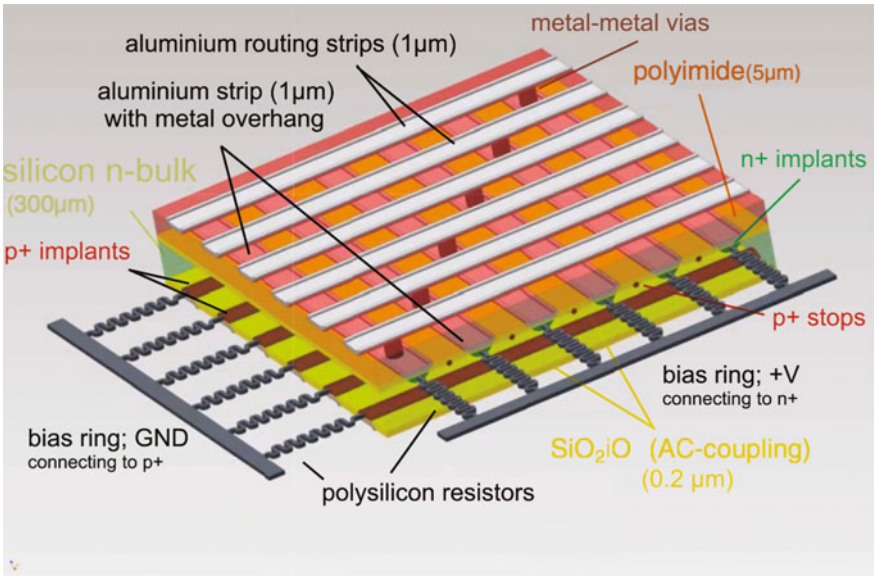


Fig. 8.13 The double-sided silicon sensor scheme of the DELPHI experiment. Reprinted from [6] with permission

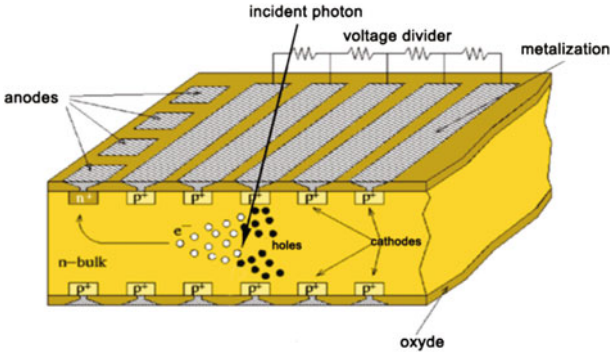


Fig. 8.14 Schematic working principle of a large area silicon drift detector. Reprinted from [10] with permission

from the DELPHI experiment [5] at the Large Electron and Positron Collider (1989–2000). Special insulation is required on the n-side to avoid the buildup of static charge between the strips, and consequently double sided devices are more expensive and complicated in the handling and manufacturing.

Alternative configurations include silicon drift detectors, first proposed in the 1980s [7] and then used for example in the STAR experiment at Brookhaven and ALICE experiment at the LHC [8], as well as for the proposed Large Observatory For X-ray Timing (LOFT) mission of the European Space Agency [9]. As shown in the schematics of Fig. 8.14, holes are collected on the cathode strips and electrons drift towards the anodes through an electric field with a decreasing gradient towards the n+ anode.

Complex semiconductor tracking systems are widely used worldwide. In modern particle physics experiments they are used for precision tracking and vertex reconstruction, including the measurement of momentum in a magnetic field. The cartoon in Fig. 8.15 shows the tracking system of the ATLAS experiment [11]. Silicon semiconductor detectors form the eight concentric layers from the beam pipe, at radii between 50.5 and 514 mm.

In the position closer to the interaction points not strips but hybrid *pixel* detectors [14, 15] are often employed. These consist of pixelated particle sensors that are essential for resolving tracking with very high occupancy (when multiple particles are likely to traverse the same strip sensor). The pixel detectors are made from a matrix of reverse bias silicon diodes and offer excellent disambiguation, large signal to noise ratio and low leakage current, at the expense of a large number of connections and electronics and higher costs. The ATLAS pixel detector uses 80 million pixels covering an area of 1.7 m^2, but the total surface covered by both pixel and strip modules accounts for 61 m^2. This is more than a tenfold increase over the previous generation of detectors at the Tevatron collider. To date, the largest detector area covered by silicon detector technology for particle tracking has been developed

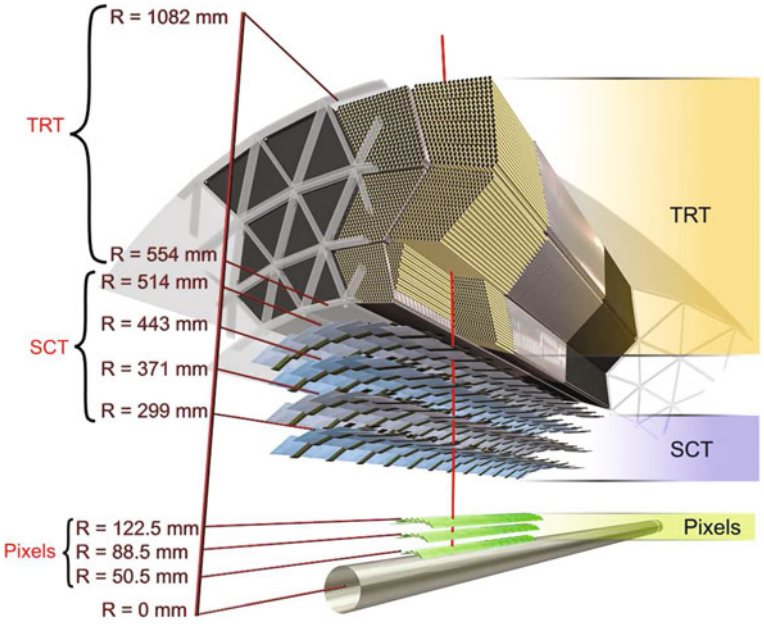


Fig. 8.15 Schematic view of a slice of the tracking detectors of the ATLAS experiment. Silicon semiconductor detectors form the eight concentric layers from the beam pipe. A charged track is seen traversing the structure. Reprinted from [12, 13] with permission

by the CMS experiment, reaching some 200 m^2 split into 15,148 submodules and instrumented with 10 million channels.

Today, hybrid pixel sensors are also widely used for X-ray photons detection, with interesting characteristics in terms of high signal to noise ratio, wide dynamic range and short readout time. Other notable developments include three dimensional silicon detectors [16].

8.5 Particle Tracking and Momentum Measurement

Semiconductor detectors and gaseous ionisation detectors are widely used in particle physics for particle detection and momentum measurement, by operating in a magnetic field that deflects the particle trajectory. Considering a relativistic particle of speed v, electric charge e and mass m moving onto a plane and being tracked through a segment L of radius R, as shown in Fig. 8.16, a simple application of Lorentz's force gives:

$$m\gamma \frac{v^2}{R} = evB, \tag{8.17}$$

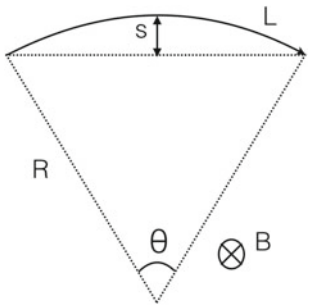


Fig. 8.16 Two-dimensional sketch of the curved trajectory of a charged particle in a magnetic field

and therefore the particle momentum is

$$p = e\ B\ R, \tag{8.18}$$

hence the momentum can be derived from the magnetic field and the radius of the bending. The (8.18) can be expressed directly in the convenient units of GeV/c by substituting the electric charge in coulombs, and expressing energy in GeV, the magnetic field in tesla and the radius in meters, so that momentum as a function of the bending radius is approximately equal to

$$p\,[\mathrm{GeV/c}] \simeq 0.3\,B[\mathrm{T}]\,R[\mathrm{m}]. \tag{8.19}$$

The resolution of the momentum measurement is directly proportional to the accuracy of the measurement of the bending radius (as well as the magnetic field), hence it is strictly dependent on the capabilities of the tracking detector

$$\frac{\Delta p}{p} = \frac{\Delta R}{R}. \tag{8.20}$$

However, it is not the radius that is measured in a tracking detector, but the sagitta, $s \pm \Delta s$, shown in Fig. 8.16, and s is related to the radius R by

$$s = R - R\cos\frac{\theta}{2} \simeq R\frac{\theta^2}{8} \simeq \frac{L^2}{8R}. \tag{8.21}$$

By substituting R of (8.21) into (8.19) the resulting expression for momentum is:

$$p\,[\mathrm{GeV/c}] \simeq 0.3\,B[\mathrm{T}]\frac{L^2}{8s}[\mathrm{m}]. \tag{8.22}$$

The error on the sagitta is related to the error on the radius by:

$$\Delta s \simeq \frac{L^2}{8R^2}\Delta R. \tag{8.23}$$

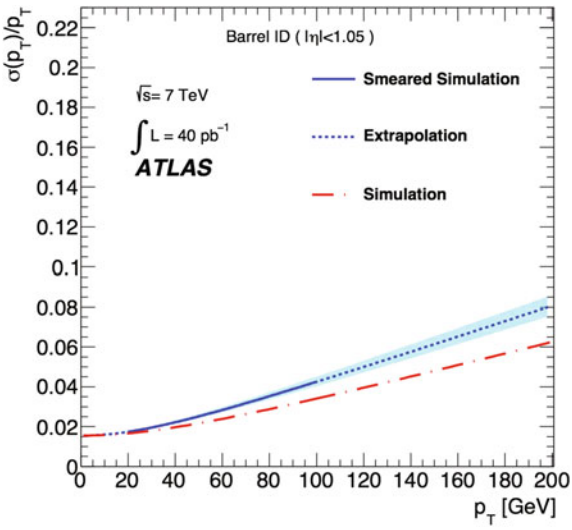


Fig. 8.17 Momentum resolution of muons in the central part of the ATLAS detector as a function of the muon transverse momentum. Printed from [17] with permission

The relative uncertainty on p is thus:

$$\frac{\Delta p}{p} = \frac{\Delta R}{R} \simeq \frac{8R}{L^2}\Delta s \simeq \left[\frac{8}{L^2 0.3\, B}\right] p\, \Delta s, \tag{8.24}$$

with lengths expressed in meters, the magnetic field in tesla and momenta in GeV/c. The sagitta resolution depends on the resolution of the individual hits position. The relative error on the momentum is therefore expected to be a linear function of momentum itself, with the error inversely proportional to the intensity of the magnetic field and to the square of the arc length. The plot in Fig. 8.17 shows the simulated muon transverse momentum resolution as a function of transverse momentum in the central portion of the ATLAS detector [17], showing the characteristic raise with p_T from resolutions of just below 2% to about 8%. The momentum resolution is limited by the multiple scattering of the particle with the detector material (discussed in Chap. 6), which is responsible for the leveling off of the resolution at low momentum values, and is dependent also on the accuracy of the knowledge of the magnetic field.

Glossary

Conduction band Set of electron energy levels for which the electrons can migrate through the sites of the lattice
Covalent bond Chemical bond in which pairs of electrons are shared between atoms
Depletion zone Volume of semiconductor depleted of free electrons and holes, as is typical in the proximity of a p-n semiconductor junction

Doped semiconductor Addition of pentavalent or trivalent atoms to a tetravalent semiconductor lattice
Energy gap The energy difference between the highest energy level in the valence band and the lowest energy level in the conduction band
Forward bias Application of an external potential to a p-n semiconductor junction, with the positive terminal connected to the p-side
Hole Vacant energy level in the valence band
Impurities Atoms of different nature positioned within the semiconductor lattice
Leakage current Electric current flowing across a p-n junction subject to reverse bias
n-type semiconductor A tetravalent crystal, such as silicon, doped with pentavalent atoms, such as phosphorus
p-n junction Positioning in contact of a p-type and a n-type semiconductor, most commonly achieved with the different doping of a single semiconductor
p-type semiconductor A tetravalent crystal, such as silicon, doped with trivalent atoms, such as boron
Reverse bias Application of an external potential to a p-n semiconductor junction, with the positive terminal connected to the n-side
Valence band Set of electron energy levels for which the electrons are bound to the sites of the lattice

References

1. Credit Enricoros at English Wikipedia, https://en.wikipedia.org/wiki/File:SiliconCroda.jpg. Accessed 20 Dec 2016. Public Domain
2. C. Jacoboni, C. Canali, G. Ottaviani, A. Alberigi Quaranta, Solid-State Electron. **20**, 77–89 (1977)
3. R. Nouicer et al., Nucl. Instrum. Methods Phys. Res. A **461**(1–3), 143–149 (2001)
4. A. Peisert, *Silicon Microstrip Detectors, Instrumentation in High Energy Physics* (World Scientific, Singapore, 1992)
5. N. Bingefors et al., The DELPHI microvertex detector, NIM **A328**, 447471 103 (1993)
6. F. Hartmann, *Evolution of Silicon Sensor Technology in Particle Physics*, vol. 231, Springer Tracts in Modern Physics (Springer, Berlin, 2009)
7. E. Gatti, P. Rehak, Semiconductor drift chamber an application of a novel charge transport scheme. Nucl. Instrum. Methods Phys. Res. **225**, 608614 (1984)
8. A. Rashevsky, V. Bonvicini, P. Burger, S. Piano, C. Piemonte, A. Vacchi, Large area silicon drift detector for the ALICE experiment. Nucl. Instrum. Methods Phys. Res. A **485**, 5460 (2002)
9. M. Feroci, L. Stella, M. van der Klis, T.J.-L. Courvoisier, M. Hernanz, R. Hudec, A. Santangelo, D. Walton, A. Zdziarski, D. Barret, T. Belloni, J. Braga, S. Brandt, C. Budtz-Jrgensen, S. Campana, J.-W. den Herder, J. Huovelin, G.L. Israel, M. Pohl, P. Ray, A. Vacchi, S. Zane et al., The Large Observatory for X-ray Timing (LOFT), Experimental Astronomy, 100 (2011)
10. Y. Evangelista, R. Campana, E. Del Monte, I. Donnarumma, M. Feroci et al., Simulations of the X-ray imaging capabilities of the silicon drift detectors (SDD) for the LOFT wide-field monitor, in *Proceedings SPIE 8443, Space Telescopes and Instrumentation 2012: Ultraviolet to Gamma Ray*, 84435P, 7 Sep 2012. doi:10.1117/12.926000

11. ATLAS Collaboration, ATLAS inner detector: Technical Design Report, 1, ATLAS-TDR-4; CERN-LHCC-97-016 (1997)
12. The ATLAS Collaboration, G. Aad et al., The ATLAS Experiment at the CERN Large Hadron Collider, JINST **3**, S08003 (2008)
13. J. Schieck, ATLAS Collaboration, Track-based alignment for the ATLAS inner detector tracking system. J. Instrum. **7** (2012). doi:10.1088/1748-0221/7/01/C01012
14. E.H.M. Heijne, P. Jarron, A. Olsen, N. Redaelli, The silicon micropattern detector: a dream? Nucl. Instrum. Methods Phys. Res. A **273**, 615 (1988)
15. CERN RD19 collaboration, Development of silicon micropattern pixel detectors, Nucl. Instrum. Meth. A **348**, 399 (1994)
16. G-F. Dalla Betta et al., Recent developments and future prospectives in 3 D radiation sensors. J. Instrum. **7**, C10006 (2012)
17. The ATLAS Collaboration, G. Aad, T. Abajyan et al., Muon reconstruction efficiency and momentum resolution of the ATLAS experiment in proton-proton collisions at $\sqrt{s} = 7$ TeV in 2010. Eur. Phys. J. C **74**, 3034 (2014). doi:10.1140/epjc/s10052-014-3034-9

Chapter 9
Scintillation Process and Light Detectors

Abstract Scintillation detectors convert energy from the incident radiation into visible or near visible light. The light is guided out of the sensitive volume and appropriately detected with photosensors. This chapter describes first the physics mechanism of luminescence and the materials most commonly used before dealing with the transport of light, the wavelength shifting and photoelectric light detection. A section on bolometers, which are sensitive thermal light detectors, ends the chapter.

9.1 Scintillation for Radiation Detection

Scintillation is the process in which the energy from a certain radiation interacting with a volume of sensitive material (called a *scintillator*) is converted into electromagnetic waves. The frequency of the emitted electromagnetic waves is within or near the visible spectrum. The sensitive material is what produces the scintillation and the idea behind its use is that the light can then be detected with appropriate devices, or *photosensors*. The human eye, for example, is a type of photosensor.

The phenomenon underlying the scintillating materials is called *fluorescence*, which refers to the fast emission of light, with a delay of the order of 10 ns from the time of deposition of the energy, following the interaction between the radiation and the material. The emission of multiple photons, i.e. the light output, follows an exponential decay and in first approximation, the time evolution of the number of emitted photons is described by a simple form

$$N_{\text{photons}}(t) = \frac{N_{\text{tot}}}{\tau_d} e^{-t/\tau_d}, \tag{9.1}$$

where N_{tot} is the total number of emitted photons and τ_d is a characteristic decay constant.

In most cases however, the time evolution of the light output does not strictly follow a single exponential decay, but is better described by a double exponential with one slow component and one fast component

$$N_{\text{photons}}(t) = Ae^{-t/\tau_A} + Be^{-t/\tau_B}, \tag{9.2}$$

L. Cerrito, *Radiation and Detectors*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-53181-6_9

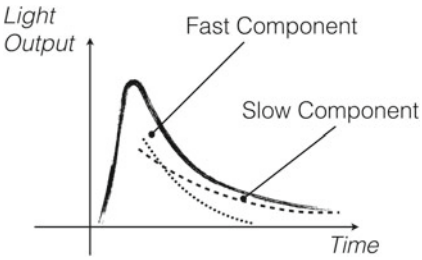


Fig. 9.1 Sketch of the time evolution of the fluorescence light output

where A and B are normalisation constants, and the τ_A and τ_B terms are the decay constants that govern the duration of the fast and slow components, as shown in the sketch of Fig. 9.1.

The amount of light emitted is called the light *response*, and it is strictly related to the efficiency for converting ionisation energy to photons in the sensitive material. In general, the main features of scintillating materials and related devices are:

- Sensitivity to energy deposition: above a certain threshold, most scintillators emit an amount of light proportional to the energy deposited
- Fast response: scintillators are fast devices compared to other detectors
- Pulse shape: the shape of the emitted light pulse can depend on the type of particle, and can therefore be used for extracting additional information on the incident radiation

Materials which are used for detection purposes as scintillators include inorganic crystals and organic plastics and liquids, gaseous scintillators and some glasses. Gaseous scintillators are mainly noble gases, whose atoms are individually excited and return to their ground state with the emission of light mainly in the UV region. Their light output response is fast, of the order of 1 ns. Glass scintillators are used primarily for neutron detection, and their structural resistance makes them ideal for extreme conditions. Inorganic crystals and organic plastics will be described in detail in the next Section.

General aspects in the characterisation and choice of a particular scintillation material for radiation detection include:

- High efficiency, for the conversion of energy into light
- Transparency, for the transmission of the scintillator own fluorescence that needs to reach the light sensors
- Emission of light in a detectable frequency, meaning matching between the emitted light frequency and the sensitive frequency of the photosensors
- Short decay time, in order to facilitate a fast response.

9.2 Inorganic Scintillators

Inorganic scintillators are crystals that produce luminescence when struck by radiation. The photon emission is typically facilitated by the presence of small amount of impurities in the crystal lattice. For instance, materials such as sodium iodide (NaI), cesium iodide (CsI), barium fluoride (BaF_2) are normally used as scintillators with predominantly tallium (Tl) being added as activator impurity. The scintillation mechanism in inorganic scintillators is characteristic of the electronic band structure of crystals. Electrons are excited into the conduction band through the energy released by the incident radiation. The impurity has the role of adding energy levels in the band gap, facilitating a sequence of de-excitations accompanied by the emission of photons with frequencies around the visible spectrum. Some *phosphorescence* may also be present, i.e. the emission of photons sometime after the primary luminescence component. This is due to electrons being trapped into energy levels which forbid direct de-excitation, and which the electrons can only leave through additional excitation transitions.

The response time of inorganic scintillators is of the order of 500 ns, with an additional light component due to phosphorescence, and the time evolution is well described by one or two exponential decays. The response frequency is in the 300–700 nm range. The graphs in Fig. 9.2 show the spectra of emission, here called intensity, of NaI(Tl), CsI(Tl) and CsI(Na). The distributions are somewhat bell-shaped with a clearly visible peak. One of the disadvantages of inorganic scintillators is the strong temperature dependence of the light output. The Fig. 9.3 shows the temperature dependence for the light output and the decay time in NaI(Tl). The material has multiple decay-time components but the primary decay constant is of about 250 ns at room temperature, with a significant temperature dependence. Furthermore, inorganic scintillators are highly hygroscopic (with some exceptions as for instance the

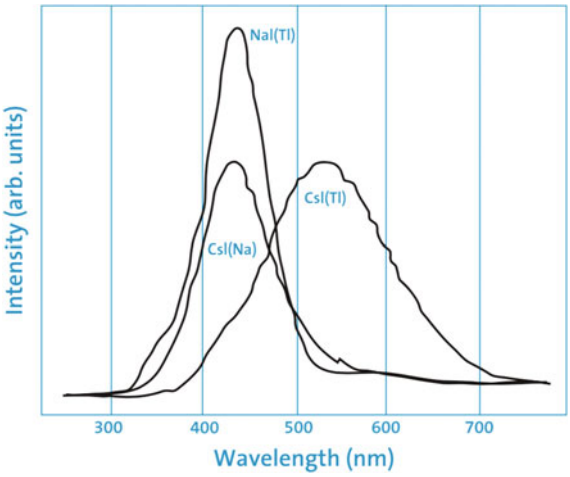


Fig. 9.2 Scintillation emission spectrum of NaI(Tl), CsI(Tl) and CsI(Na). Courtesy of Saint-Gobain Crystals, a division of Saint-Gobain Ceramics & Plastics, Inc. Printed from [1] with permission

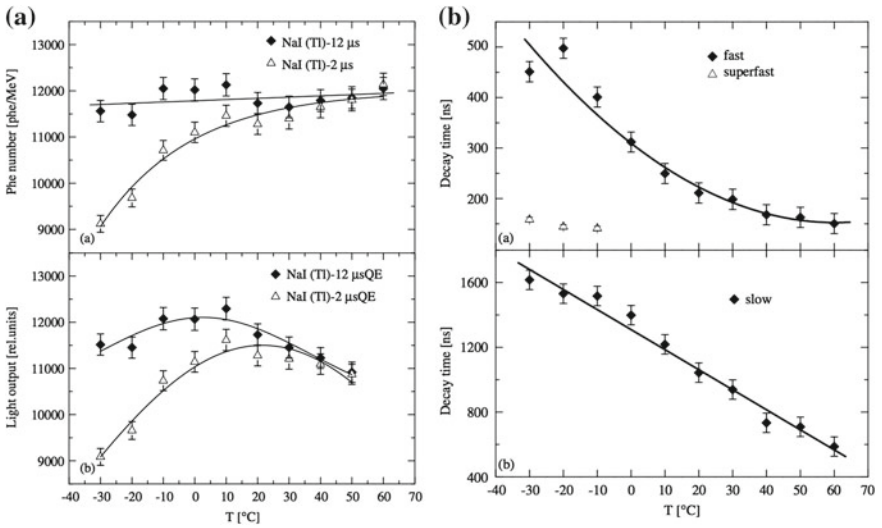


Fig. 9.3 Dependence of **a** the response and **b** the decay time on temperature for a NaI(Tl) scintillator. Printed from [2] with permission

bismuth germanate oxide, $B_4Ge_3O_{12}$ or BGO), so they must be protected into air tight enclosures. Advantages include a high radiation stopping power and high light output, which ensure a good energy resolution.

Sample parameters that characterise NaI(Tl) crystals include the wavelength of maximum emission, $\lambda_{max} = 415\,\text{ns}$ ($h\nu_{max} = 3.0\,\text{eV}$), and the number of photons (about 38,000) emitted per MeV of energy deposited by γ radiation, which corresponds to 26 eV per emitted photon. The light yield ε, that is the ratio between the energy absorbed from the incident radiation and the energy of the luminescence produced, can hence be estimated as

$$\varepsilon = 38,000 \times \frac{3.0\,\text{eV}}{10^6\,\text{eV}} = 11.4\%. \tag{9.3}$$

Note that the response is in general different for different types of incident particle at the same energy, and it might not vary linearly with energy.

Inorganic crystals for radiation detection are widely employed worldwide. The picture in Fig. 9.4 shows a lead-tungstate ($PbWO_4$, shortened as PWO) crystal [3] used for the electromagnetic calorimeter [4, 5] of the CMS experiment at the Large Hadron Collider.

Individual crystals are then arranged in complex arrays to form wide area detectors, from relatively small, such as the one pictured in Fig. 9.5 used for the HERMES experiment at the DESY laboratory, to tens of square meters.

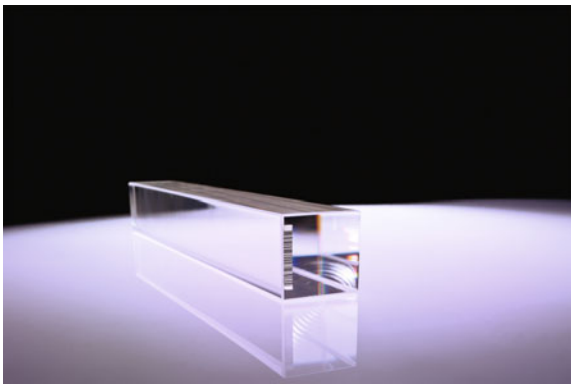


Fig. 9.4 A lead-tungstate crystal of the type used in the CMS electromagnetic calorimeter. Credit © CERN, for the benefit of the CMS Collaboration, printed with permission [3, 18]

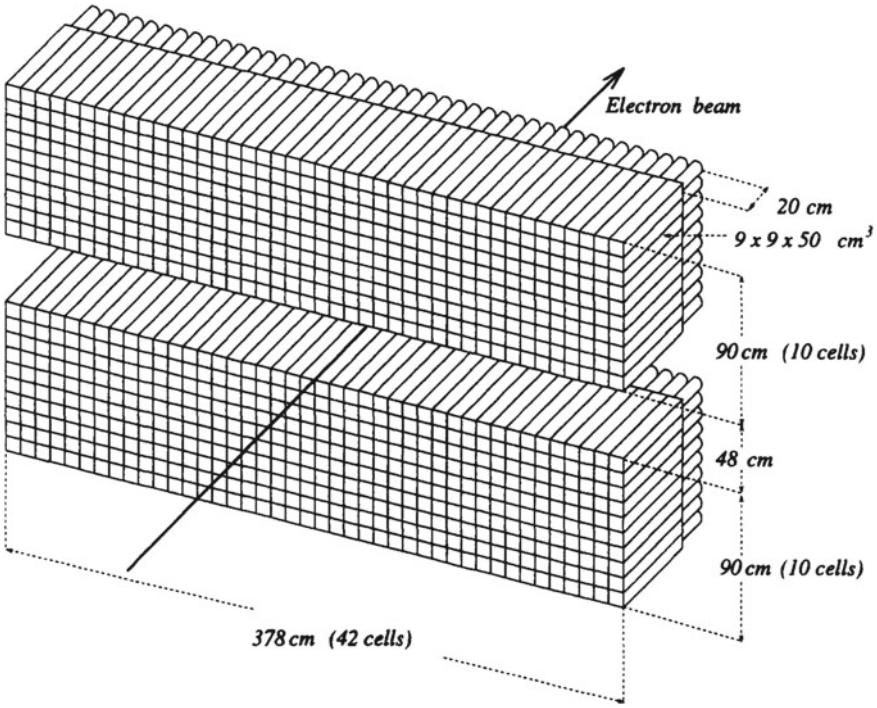


Fig. 9.5 Isometric view of the calorimeter of the HERMES experiment. Printed from [6] with permission

9.3 Organic Scintillators

Organic scintillators are aromatic hydrocarbon compounds, such as Naphthalene ($C_{10}H_8$), Antracene ($C_{14}H_{10}$) or Stilbene ($C_{14}H_{12}$). The luminescence, unlike for inorganic crystals, is of molecular origin, i.e. transition of electrons between energy levels within the molecules. Fluorescence is produced with decay times of order 10^{-8}–10^{-9} s, which is significantly faster than for inorganic scintillators, while phosphorescence is a slow process with times greater than 10^{-4} s. The frequency output for the fluorescent component is in the UV region (around 300 nm). Organic crystals are difficult to cut and the response might sometimes depend on the orientation of the crystal, therefore they are most commonly used as solutions in plastic or liquid materials. This has the additional benefit that the plastic or liquid compounds can be *loaded* with additives with specific properties, e.g. wavelength shifters to modify the UV-based light response. Normally, the absorption of energy is due mainly to the solvent and the energy is passed on to the scintillation solute. The most common plastic solvent is polyvinyl toluene (PVT), which is doped with one of the organic scintillators to produce a plastic scintillator.

Since the decay constant of the signal in plastic scintillators is very fast, the raise of the curve cannot be ignored, such that the best description of the light output (expressed e.g. as the number of photons, $N(t)$) is the convolution of a Gaussian and an exponential function

$$N(t) = N_0 G(\sigma, t) e^{-t/\tau}, \tag{9.4}$$

where N_0 is a normalisation constant and $G(\sigma, t)$ is a Gaussian function of width σ that fits the raising profile of the curve. While relatively easy to machine, plastic scintillators can be damaged by organic solvents and body acids, causing *cracking* of the plastic (crease). Therefore, great care must be taken to handle the materials and avoid contaminating the surface.

Compared to inorganic scintillators, plastic scintillators have a small temperature dependence and are easily shaped and cheaper to produce. However, they emit a lower light yield and can be damaged by radiation breaking the C-H bonds, reducing the light output significantly. The plot on the left in Fig. 9.6 shows as an example the light output as a function of photon wavelength for a typical premium plastic scintillator (the Saint-Gobain BC-408). The plot on the right in Fig. 9.6 shows the light output as a function of the incident particle type and energy. Good linearity is shown in this case.

In general the relationship between the light output (L.O.) and the energy deposited is not necessarily linear, and a good approximation for the material and particle dependence is of the following form (Birk's law):

$$\frac{\mathrm{d\,L.O.}}{\mathrm{d}x} = \frac{A\,\mathrm{d}E/\mathrm{d}x}{1 + C\,\mathrm{d}E/\mathrm{d}x}, \tag{9.5}$$

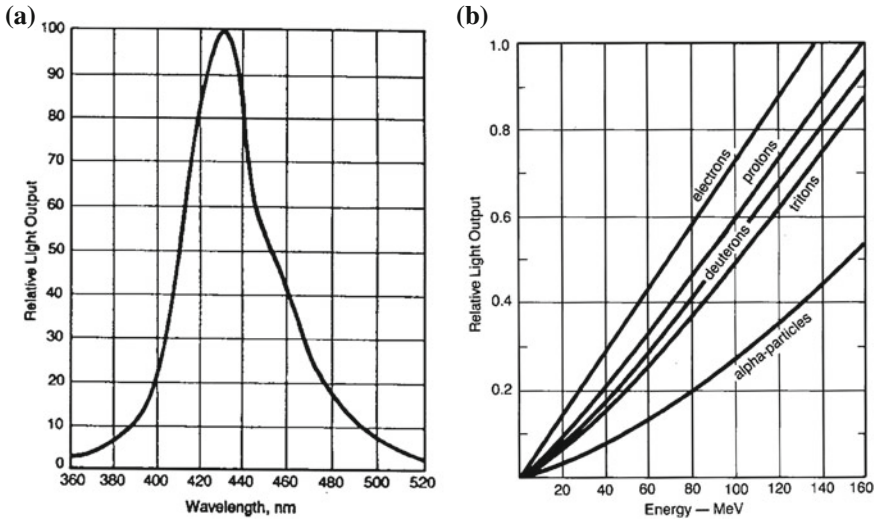


Fig. 9.6 **a** Dependence of the response versus wavelength for a premium plastic scintillator (the Saint-Gobain BC-408). Courtesy of Saint-Gobain Crystals, a division of Saint-Gobain Ceramics & Plastics, Inc. **b** Dependence of the response of the plastic scintillator NE 102 versus incident radiation. Adapted from [7] with permission

where x is the path length, A is the scintillation efficiency and C is an experimental constant that depends on the scintillating material.

The energy needed for photon production ranges from 60 eV/photon (Antracene) to 100–300 eV/photon (for most plastic scintillators). Antracene is the most efficient of the organic scintillators and used as a reference. Notice how these energy thresholds are generally higher than for inorganic scintillators. Also, unlike inorganic crystals, the temperature dependence of organic scintillators is moderate, in particular the light output is mostly independent of temperature at room temperatures, but it becomes a factor at extreme values (below $-60\,^{\circ}\mathrm{C}$ or over $100\,^{\circ}\mathrm{C}$). The shape of the light pulse is also dependent on the incident particle, as a consequence of the excitation of different states by different particles. This in turn allows the exploitation of the light output for particle identification purposes, as shown in Fig. 9.7 for an example of neutron-gamma discrimination.

In summary, the choice between organic or inorganic type of scintillators is down to the required detector performance, working environment, and costs. Inorganic scintillators have a high light yield and high density, and therefore large energy deposition and good energy resolution. However, the crystal growth is relatively complex, they tend to have a large temperature dependence and are relatively expensive. On the other hand, organic scintillators are relatively inexpensive, are characterised by a very fast response, they are easily shaped and have a small temperature dependence. However, their light response is lower than for inorganic scintillators and they are more susceptible to radiation damage.

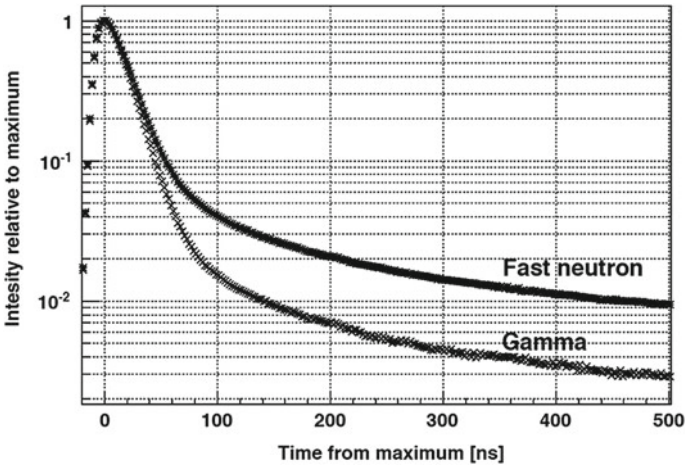


Fig. 9.7 Average waveforms from ^{252}Cf for neutrons and gammas in a hydrocarbon-based plastic, normalised to the same electron-equivalent energy range. Printed from [8] with permission

9.4 Transport and Detection of Light

Not always it is possible to couple a reading device directly onto the scintillation material, therefore the light often needs to be transported along a guide. Light guides are made of optical quality plexiglass, and internal reflection carries the light away from the scintillator. Note that a light guide must be adiabatic, that is if the cross sectional area is A_{in} at the input and A_{out} at the exit, then at best only a fraction A_{in}/A_{out} can be transported. Common configurations include *fish tail* and *twisted* light guides (an array of strips adapting the edge of the scintillator to a phototube). The photograph in Fig. 9.8 shows a few types commonly in use. To avoid light escaping off the sides, a reflective material such as aluminium is wrapped around both the scintillator and the light guide. Black covering is typically placed outside the aluminium to avoid contamination of light from the outside.

The actual detection of light is performed through devices called *photomultipliers* (PMTs). Photomultipliers are electron tubes which convert light into a measurable current. The sketch in Fig. 9.9 shows an illustration of a photomultiplier. It is composed of a photosensitive cathode followed by an electron amplification system. The photocathode emits electrons via the photoelectric effect, the anode then collects the cascade electrons and the signal is amplified through a number of stages to the final output. The electric current will be directly proportional to the number of incident photons.

The photocathode is constructed as a thin layer of photosensitive material laid onto a glass. The kinetic energy E of the emitted electrons follows the relationship

$$E = h\nu - \phi, \tag{9.6}$$

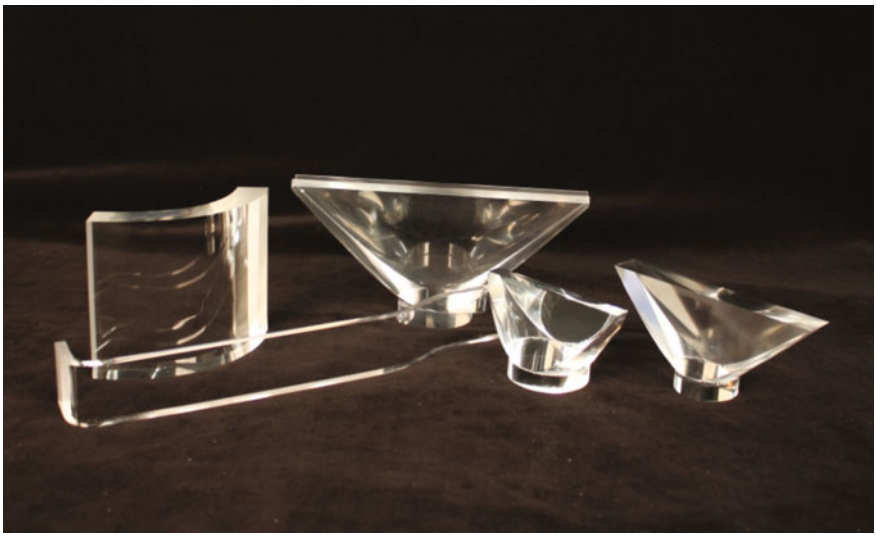


Fig. 9.8 Example of various shapes of light guides. Courtesy of Eljen Technology (US). Printed from [9] with permission

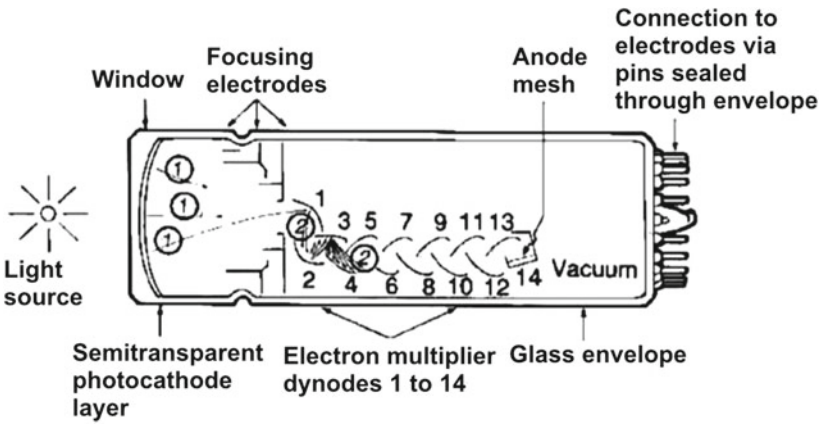


Fig. 9.9 Schematic illustration of a photomultiplier, showing the photoelectrons (①) and the electrons from secondary emission (②). Reprinted from [10] with permission

where ν is the frequency of the incident photon and ϕ is the work function. In general, the efficiency of photon to electron conversion (called *quantum efficiency*) is not 100% and is non linear. Quantum efficiency is defined as

$$\eta(\lambda) = \frac{\text{number of photoelectrons}}{\text{number of incident photons}}. \tag{9.7}$$

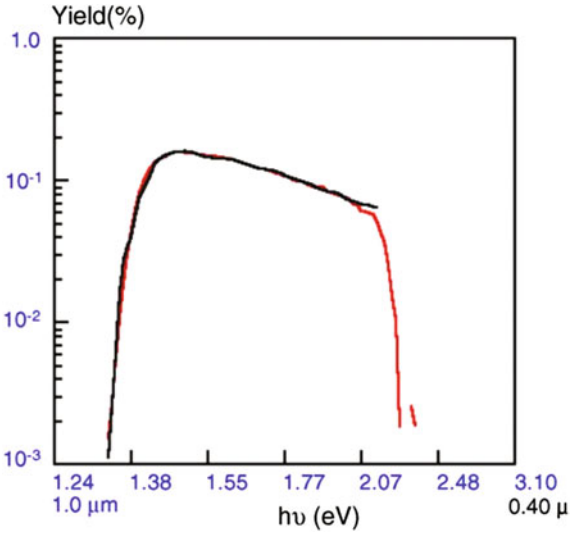


Fig. 9.10 Quantum yield curve for a GaAs photocathode. Reprinted from [12] with permission

Typical efficiencies are of the order of 10–25%, depending on the photon frequency (see Fig. 9.10 for an example of a quantum yield spectrum). The GaAsP and UBA photocathodes have currently some of the highest quantum efficiencies (reaching 40–50%), but a narrow band of sensitivity, while InP/InGaAs are sensitive from 300 nm to 1600 nm (a full comparison of various photocathodes is given in [11]). Additionally, there are substances which can be used to optimise the quantum efficiency for specific incident wavelengths, from the infrared to the ultraviolet. Most photocathodes are made of semiconductor materials.

Once photoelectrons have been emitted, the remainder of the photomultiplier is an electron amplification device. The first step of amplification consists in collecting the emitted electrons using an electric field. General requirements involve the collection efficiency and the time, which must be independent on the point of emission from the photocathode. After collection, an electron amplifier multiplies the weak photocurrent by using a series of secondary emission electrodes. The gain of each electrode (called a dynode) is known as *secondary emission* factor, δ, defined as the ratio between the number of electrons produced over the number of electrons incident on the electrode. Typical gain for individual stages is between 2 and 10. Semiconductor materials are also used for this purpose and the theory behind the process is similar to the description of the photoelectric effect. Because an electric field is required to accelerate and collect electrons at each subsequent stage, and the gain of each dynode is related to the energy of the electrons and therefore to the voltage between dynodes, the semiconductor is placed onto a conducting material. General requirements call for high and stable secondary emission and low thermic emission (in order to minimise the thermal noise). Most PMTs contain between 10 and 14 stages of secondary emission and reach gains of 10^6–10^7. The Fig. 9.11 shows a photomultiplier with a 10-stage venetian-blind dynode structure. Various configurations

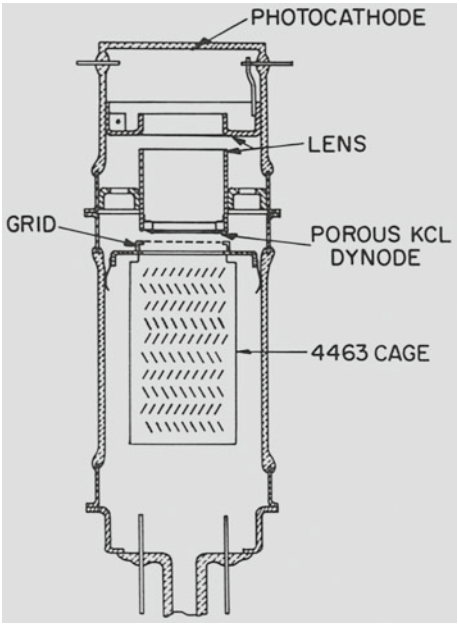


Fig. 9.11 Experimental photomultiplier with a porous transmission dynode and a 10-stage venetial-blind dynode structure. Reprinted from [14] with permission

Fig. 9.12 The 9357FLA Electron Tubes photomultiplier. The glass window has been coated with a wavelength shifter in this particular device. Printed from [15] with permission

for the amplification stages, including circular-cage and box-grid structures can be seen in [13]. The picture in Fig. 9.12 shows a photomultiplier with hemispherical glass window, one of the many PMT shapes and sizes that be constructed [13].

The key to PMT operations is the stability of the voltage between dynodes, crucial to minimise gain fluctuations, and the inter-dynode voltage must be sufficient to achieve full charge collection at each stage. Voltage supply must be regulated to within 0.1%, and PMTs are often left to run for at least a few hours in order to stabilise. The pulse shape depends on the time constant of the circuit. The PMT is said

to operate in current mode if the time constant of the photomultiplier circuit is much smaller than the time constant of the scintillator. Conversely, it is said to operate in voltage mode if the opposite is true for the two time constants. Voltage mode is normally preferred since it provides increased signal stability, if pile up is not an issue for the specific application. The time response of the PMT depends on the initial electron collection and the subsequent steps. The time spread introduced by the first dynode is already about 0.5 ns and thermal noise contributes additionally to the time resolution. Finally, PMTs are sensitive to magnetic fields due to the extended paths of the electron collection lines, so a metallic screen is typically required if the PMT is to operate in a strong magnetic field.

9.5 Bolometers

Bolometers are thermal detectors that absorb photons and transform their energy into heat, which is then detected by a sensitive thermometer. While thermal detectors have a non-specific spectral response, they are particularly suited for submillimeter and infrared wavelengths and have been extensively developed to be today the detector of choice for low light levels in this spectral range. The basic working principle of a bolometer is shown in Fig. 9.13. A detector element is connected via a thermal link to a heat sink, kept at temperature T_0. Under an incident radiation, the detector absorbs a constant power P_0 raising its temperature by T_1. The value of T_1 depends on the thermal conductance G (measured in watts per kelvin) [16]:

$$T_1 = \frac{P_0}{G}, \tag{9.8}$$

therefore a measurement of T_1 corresponds to a measurement of the absorbed power. In practice the input power has a time dependence, so by introducing a variable power component $P_\nu(t)$ deposited by the radiation of frequency ν, the temperature will raise with a time gradient:

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{\eta}{C} P_\nu(t), \tag{9.9}$$

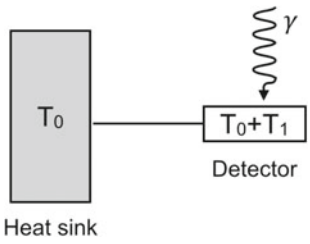


Fig. 9.13 Schematic diagram of a bolometer

where C is the heat capacity and η is the fraction of incident power absorbed by the detector. The total power $P(t)$ absorbed by the detector is therefore:

$$P(t) = P_0 + \eta P_\nu(t) = GT_1 + C\frac{\mathrm{d}T}{\mathrm{d}t}. \tag{9.10}$$

If $P_\nu(t)$ changes from zero at time $t < 0$ to P_1 at time $t \geq 0$ with a step function, the equation (9.10) becomes

$$P(t) = P_0 + \eta P_1 = GT_1 + C\frac{\mathrm{d}T}{\mathrm{d}t}, \tag{9.11}$$

and its solution is

$$T_1(t) = \frac{P_0}{G} + \frac{\eta P_1}{G}\left(1 - e^{-t/\tau}\right), \tag{9.12}$$

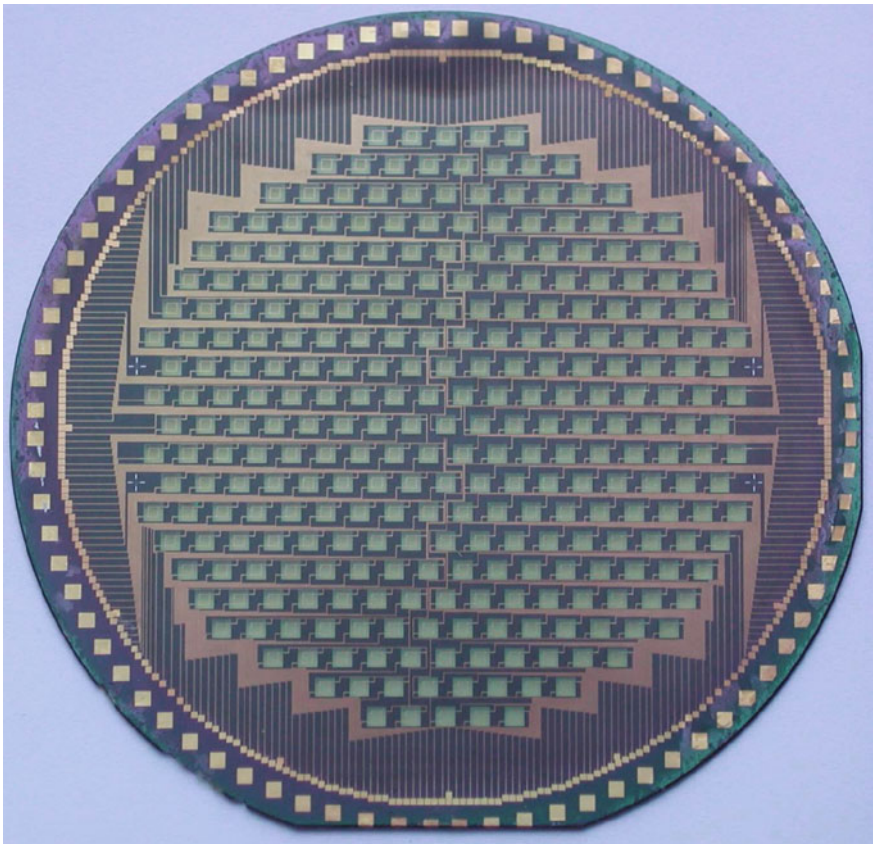


Fig. 9.14 Large APEX BOlometer CAmera (LABOCA) bolometric array. Each square in the central pattern is a bolometer. Credit [17], reproduced with permission © ESO

with $\tau = C/G$. The exponential part describes the situation soon after $t = 0$, while for times long compared to τ the temperature stabilises at a value directly proportional to the total absorbed power.

In modern bolometers, the temperature is measured by monitoring the voltage across the detector component, made of semiconductor material, as it varies in response to a change in its electrical resistance which depends on the temperature. Because of the sensitivity of these devices to electronic noise, good performance can be obtained typically by using very low temperatures. Today bolometers are also used in detector arrays reaching hundred to a few thousand elements. The picture in Fig. 9.14 shows a bolometric array of the Large APEX Bolometer CAmera (LABOCA) [17], a receiver for the Atacama Pathfinder Experiment 12 m submillimeter telescope (APEX) located in Northern Chile. The LABOCA detector consists of 295 semiconducting composite bolometers used for the detection of astronomical objects emitting in the frequency range of 300–400 GHz.

Glossary

Birk's law Empirical relationship between the light output (L.O.) and the energy deposited in a scintillator

$$\frac{\mathrm{d\,L.O.}}{\mathrm{d}x} = \frac{A\mathrm{d}E/\mathrm{d}x}{1 + C\mathrm{d}E/\mathrm{d}x}, \tag{9.13}$$

where x is the path length, A is the scintillation efficiency and C is an experimental constant that depends on the scintillating material

Bolometers Thermal detectors that absorb photons transforming their energy into heat

Dynode Part of a photomultiplier acting as a secondary emission electrode, i.e. increasing the number of photoelectrons moving through the device

Fluorescence Fast emission of light by a substance following the deposition of the energy from an incident radiation. It is a type of luminescence

Inorganic scintillator Crystals that produce luminescence when struck by radiation, such as sodium iodide (NaI), cesium iodide (CsI), barium fluoride (BaF_2), often with a small amount of activator impurities

Light guide Plexiglass used to transport the light emitted from a scintillator onto a photosensor

Light response The amount of light emitted from a scintillator following the deposition of the energy from an incident radiation

Light yield Ratio between the energy absorbed from the incident radiation and the energy of the luminescence produced

Organic scintillator Aromatic hydrocarbon compounds, such as Naphthalene ($C_{10}H_8$), Antracene ($C_{14}H_{10}$), Stilbene ($C_{14}H_{12}$), producing luminescence when struck by radiation

Phosphorescence Emission of photons sometime after the primary luminescence component, due to electrons being trapped into energy levels which forbid direct de-excitation

Photocathode Part of a photomultiplier which emits electrons via the photoelectric effect

Photomultiplier (PMT) Electron tubes converting light into a measurable current

Plastic scintillator A substance composed of a plastic solvent, most commonly polyvinyl toluene (PVT), doped with one of the organic scintillators

Quantum efficiency Efficiency of photon to electron conversion of a photocathode, i.e. the ratio between the number of electrons produced and the number of photons incident onto a photosensitive material

Scintillator Material emitting light, at visible or near visible frequencies, when struck by radiation

Time response Time interval between the deposition of energy and the emission of luminescence by a scintillator substance

Wavelength shifter Fluorescent material absorbing light at a given frequency and re-emitting at a lower frequency.

References

1. Saint-Gobain Ceramics & Plastics, Inc, http://www.crystals.saint-gobain.com. Data from IEEE NSS NS-30, 380 (1983), http://www.crystals.saint-gobain.com/sites/imdf.crystals.com/files/documents/sodium-iodide-material-data-sheet.pdf. Accessed 7 Dec 2016
2. M. Moszyński et al., Nucl. Instrum. Methods Phys. Res. Sect. A **568**(2) 739–751 (2006). doi:10.1016/j.nima.2006.06.039
3. A.A. Annenkov, M.V. Korzhik, P. Lecoq, Lead tungstate scintillation material. Nucl. Instrum. Methods Phys. Res. Sect. A **490**(12), 30–50 (2002). http://dx.doi.org/10.1016/S0168-9002(02)00916-6
4. C.M.S. Collaboration, The CMS experiment at the CERN LHC 2008. JINST **3**, S08004 (2008)
5. CMS Collaboration, CMS Physics Technical Design Report, Volume 1: Detector Performance and Software 2006, CERN/LHCC-2006-001; CMS-TDR-008-1 (2006)
6. H. Avakian et al., Performance of F101 radiation resistant lead glass shower counters. Nucl. Instrum. Methods Phys. Res. Sect. A **378**(12), 155–161 (1996)
7. T.J. Gooding, H.G. Pugh, Nucl. Instrum. Methods **7** 189–192 (1960)
8. N. Zaitseva, Pulse shape discrimination with lithium-containing organic scintillators. Nucl. Instrum. Methods Phys. Res. A **729**, 747–754 (2013)
9. Eljen Technology, 1300 W. Broadway, Sweetwater, TX 79556, www.eljentechnology.com. Accessed 7 Dec 2016
10. D. Renker, Nucl. Instrum. Methods Phys. Res. Sect. A **527**(12), 15–20 (2004)

11. Hamamatsu Photonics K.K., http://www.hamamatsu.com/us/en/technology/innovation/photocathode/index.html. Accessed 7 Dec 2016
12. L. Lei et al., The variation of spectral response of transmission-type GaAs photocathode in the seal process. Appl. Surface Sci. **251**, 273–277 (2005)
13. Hamamatsu Photonics K.K., Photomultiplier Tubes, Basics and Applications, 3rd edn. (2007), https://www.hamamatsu.com/resources/pdf/etd/PMT_handbook_v3aE.pdf
14. Smith et al., Performance of a photomultiplier with a porous transmission dynode. IEEE Trans. Nucl. Sci. **13**(3) (1966)
15. P. Benetti et al., Nucl. Instrum. Methods Phys. Res. Sect. A **505**(12), 89–92 (2003)
16. G. Rieke, *Detection of Light, From the Ultraviolet to the Submillimeter* (Cambridge University Press, Cambridge, 2003)
17. G. Siringo et al., Astron. Astrophys. **497**(3), 945–962 (2009)
18. Image Number CERN-EX-0803027-02, CERN Document Server, https://cds.cern.ch/record/1100385#02. Accessed 21 Dec 2016

Chapter 10
Electromagnetic and Hadronic Showers: Calorimeters

Abstract The ionisation mechanism presented in Chap. 6 is only one of the processes by which charged particles interact with matter, and is the dominant for heavy particles in thin media. Electrons and photons are subject to a number of other processes which cannot be neglected and are exploited in the development of calorimeters. Hadrons too loose energy by more mechanisms than just ionisation. The additional processes are the topic of this chapter. Following a review of the characteristics of electromagnetic and hadronic showers, the later sections present the design of calorimeters, both electromagnetic and hadronic, which exploit this type of phenomenology.

10.1 Interaction of Electrons with Matter

The total energy loss of electrons interacting with matter is the result of a number of processes, some with larger impact than others and active in different ranges of the electron energy. The plot in Fig. 10.1 shows the fractional energy loss of electrons and positrons in lead, as a function of the electron or positron energy. The graph shows two dominant processes: ionisation from the collision with atomic electrons, particularly for low energies, and *bremsstrahlung*, also referred to as *radiative* energy loss, for moderate and high-energy electrons and positrons. The rate of energy loss can therefore be approximated as

$$\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right) \simeq \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{radiative}} + \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{ionisation}} \tag{10.1}$$

Bremsstrahlung arises in this context when charged particles are accelerated in the Coulomb field of the nucleus, and its magnitude is proportional to the inverse of the square of the mass of the particle. Because of this dependence, it is only really relevant for electrons and positrons, which are relatively light particles, and ultra relativistic muons while it can be neglected for heavier particles and ions. The energy loss by bremsstrahlung alone can be approximated and parametrised as:

L. Cerrito, *Radiation and Detectors*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-53181-6_10

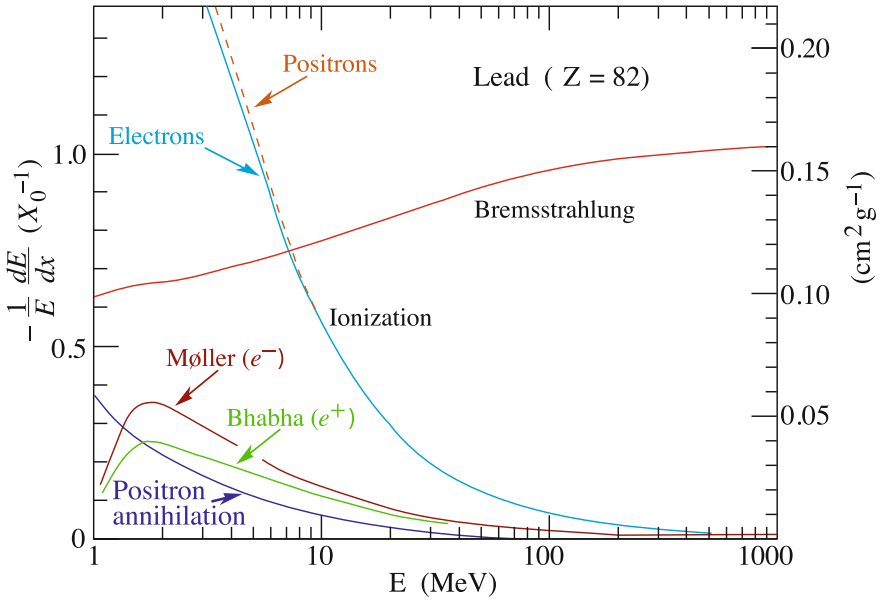


Fig. 10.1 Fractional energy loss in lead as a function of the electron or positron energy. Originally adapted from [1]. Reprinted from [2] with permission

$$-\frac{1}{\rho}\frac{\mathrm{d}E}{\mathrm{d}x} \simeq \frac{E}{X_0}, \tag{10.2}$$

where ρ is the density of the traversed material, measured in g/cm^3, x is the thickness in cm and X_0 is called the *radiation length*. Integrating (10.2) on both sides yields

$$E(x) = E_0 e^{-x\rho/X_0}, \tag{10.3}$$

which shows how the radiation length X_0 expresses the mean thickness (in g/cm^2) after which the electron or positron energy is reduced by a factor $1/e$ of its initial energy, E_0. Typical values of X_0 can range from 36.62 g/cm^2 for air at standard pressure to 6.37 g/cm^2 for lead (see Table 10.1). When converted into a mean distance these values correspond to about 300 m in air and 5.6 mm in lead.

Since energy loss by ionisation is dominant at low electron energies, while bremsstrahlung is leading at high energies, there is a value for which their contribution to the energy loss is equal. The *critical energy* is defined as the energy, E_c, of the electron for which

$$\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{radiative}} = \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{ionisation}}, \tag{10.4}$$

Table 10.1 Radiation length, X_0, and critical energy E_c of various materials. Values from [3]

Material	X_0 (g/cm^2)	X_0 (cm)	E_c (MeV, for e^-)
Air (dry, 1 atm)	36.62	30,390	87.92
Water	36.08	36.08	78.33
Scint. (Polyvinyl toluene)	43.90	42.62	94.11
Lead	6.37	0.561	7.43
Iron	13.84	1.76	21.68
Aluminum	24.01	8.89	42.70
Standard rock	26.54	10.01	49.13
Shielding concrete	26.57	11.55	49.90
Lead glass	7.87	1.26	10.41

and an approximate estimate of E_c for different materials may be obtained from the following empirical formula:

$$E_c \simeq \frac{800\,\text{MeV}}{Z + 1.2}, \tag{10.5}$$

where Z is the atomic number of the target material. For most materials, E_c is in the range of 10–100 MeV; for instance in water and air it is about 100 MeV, while in lead it is around 7 MeV. Precise values can be obtained from [3] and some examples are shown in Table 10.1. There are other processes which affect the energy loss of electrons and positrons as they pass through a given material, and the plot in Fig. 10.1 instructively shows their impact: Møller and Bhabha scattering refers to the scattering of electrons or positions, respectively, off the atomic electrons with energy loss per collision below 0.255 MeV. These processes are shown in Fig. 10.2a and b. These same processes are called ionisation for higher values of the energy loss per single collision. The positron annihilation is the reaction $e^+e^- \to \gamma\,\gamma$, in which matter and anti-matter particles are turned into photons, as shown in Fig. 10.2c. Notice also, from the behaviour of the curves in Fig. 10.1, that the rate of energy loss is approximately constant with E for all but bremsstrahlung, for which it grows linearly with E.

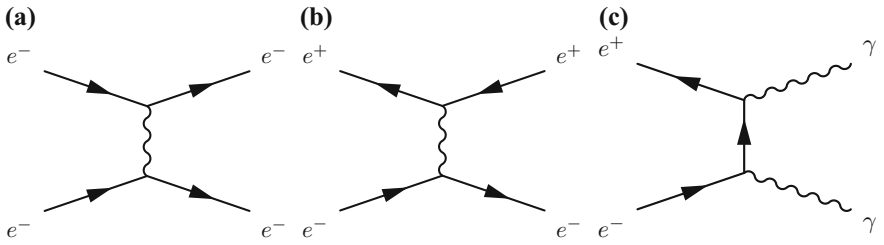


Fig. 10.2 Leading order Feynman diagrams for some of the processes of electron interaction in matter. **a** Møller scattering. **b** Bhabha scattering. **c** Electron–positron annihilation

10.2 Interaction of Photons with Matter

Photons interact with matter through the electromagnetic interactions via three main processes. The first is the photoelectric effect, in part already described in Chap. 9 in the context of detecting light from scintillation. The photoelectric effect refers to the absorption of a photon by an atom and the subsequent emission of an electron. The second is the Compton scattering, that is the scattering of photons off atomic electrons, which might also cause ionisation of the atom itself. The third mechanism is the pair production of one electron and one positron form the disappearance of a photon, in the field of an atomic nucleus.

The photoelectric effect is shown in Fig. 10.3, where a photon colliding with an atomic electron leads to the ejection of the electron and a recoiling ion. The electron energy is related to the photon frequency by

$$E_e = h\nu - \phi, \tag{10.6}$$

where ϕ is the binding energy and $h\nu$ is the energy of the photon.

The cross section for the photoelectric process, $\sigma_{p.e.}$, as a function of the photon energy, E_γ, drops rapidly for soft photons

$$\sigma_{p.e.} \sim Z^5 E_\gamma^{-3/2} \quad \text{for} \quad E_\gamma << m_e c^2, \tag{10.7}$$

and has a milder E_γ dependence for hard photons

$$\sigma_{p.e.} \sim Z^5 E_\gamma^{-1} \quad \text{for} \quad E_\gamma >> m_e c^2. \tag{10.8}$$

The plot in Fig. 10.4 shows the behaviour of the total photon cross section on lead. The kinks on the photoelectric curve are due to the ejection of the outer electrons in the atomic shell.

For photon energies of around 1 MeV, the Compton scattering is the dominant contributor to the total cross section. A sketch of the kinematics of the process is shown in Fig. 10.5, and the process was already discussed in Sect. 4.4. The energy of the outgoing photon, γ', can be easily calculated (see Sect. 4.4) and is:

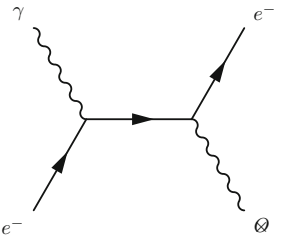


Fig. 10.3 Leading order Feynman diagram of the photon–electron scattering in a nuclear field (photoelectric effect)

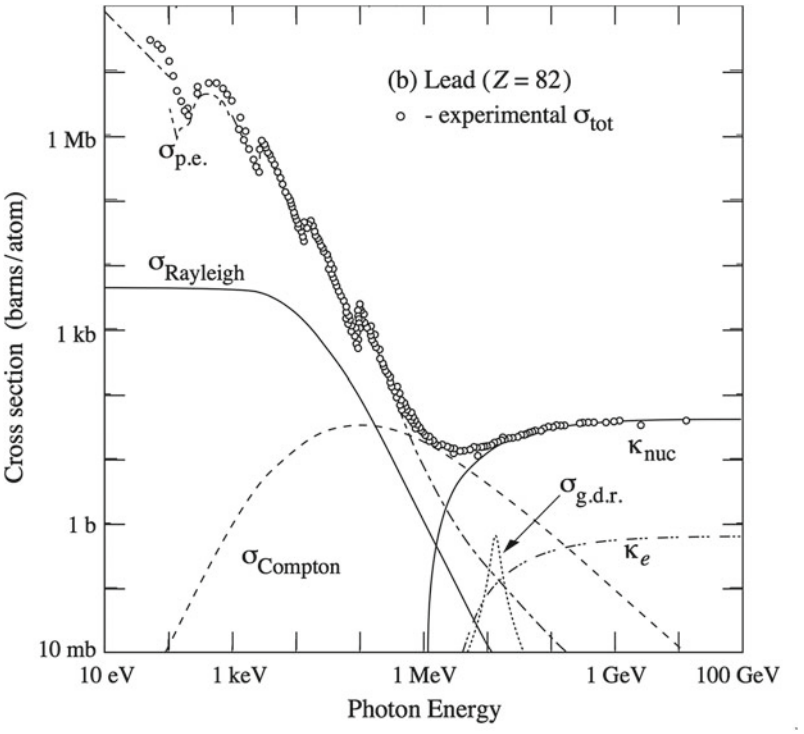
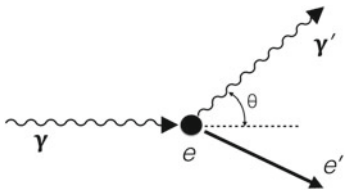


Fig. 10.4 Photon total cross sections as a function of the photon's energy in lead, showing contributions from several sub-processes. The $\sigma_{p.e.}$ is the photoelectric cross section, σ_{Compton} is the compton scattering off an electron, and κ_{nuc} is the pair production in the nuclear field. Reprinted from [4] with permission

Fig. 10.5 Diagram of compton scattering



$$E_{\gamma'} = \frac{E_\gamma m_e}{m_e + E_\gamma(1 - cos\theta)} = \frac{E_\gamma}{1 + \frac{E_\gamma}{m_e}(1 - \cos\theta)}. \tag{10.9}$$

The maximum energy transferred to the electron

$$E_{e,\,\max} = E_\gamma - E_{\gamma',\,\min}, \tag{10.10}$$

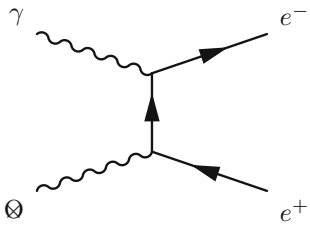


Fig. 10.6 Pair production in the nuclear field

corresponds to a maximum value called the Compton edge. Finally, the dominant mechanism to the total photon cross section off nuclei at high photon energy is pair production in the nuclear field. This process is the transformation of a photon into an e^+e^- pair and the reaction needs the presence of the nucleus in order to conserve energy and momentum. A diagram of the process is shown in Fig. 10.6. The reaction proceeds as long as E_γ is

$$E_\gamma > 1.022 \text{ MeV}, \tag{10.11}$$

which corresponds to twice the electron rest mass, so the process has a threshold that forbids it for softer photons.

10.3 Electromagnetic Showers

When an energetic photon enters a dense target material this triggers not just one but a chain of interaction processes. As described in Sects. 10.1 and 10.2, photons, electrons and positrons are all produced by mechanisms such as bremsstrahlung and pair production. One photon might produce a e^+–e^- pair which on it own account will interact with the material producing bremsstrahlung and therefore leading to more photons. A schematic of the chain reaction is shown in Fig. 10.7, indicating how the number of particles grows progressively to produce a so-called *electromagnetic shower*. A photon will go on to produce one e^+–e^- pair after travelling a distance of about 1 radiation length. The shower will stop when the energy of the secondary photons is no loger sufficient for further pair production, so the longitudinal depth will depend on the energy of the initial photon. The transverse size of the shower is given by the empirical formula [5, 6]:

$$R_\text{M} \simeq \frac{21 \text{ MeV}}{E_c} X_0, \tag{10.12}$$

where R_M is a parameter called the Molière radius, E_c is the critical energy in the material and X_0 is the radiation length. Typical values for R_M are of a few centimeters, for instance 1.6 cm in lead, 4.3 cm in lead glass and 9.1 cm in scintillator.

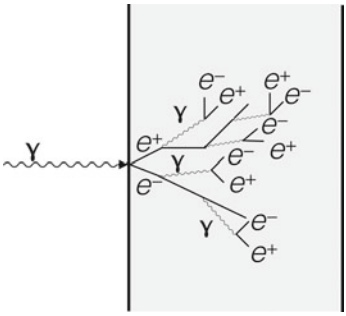


Fig. 10.7 Sketch of the development of an electromagnetic shower

The longitudinal size can be estimated through the following simplified model. Let t measure the depth in units of the radiation length, i.e. $t = x/X_0$, and assume that each electron loses half of its energy to a single photon after every traversed X_0, so that X_0 measures the steps of each pair production and bremsstrahlung event. The number of shower particles at a depth t will be:

$$N(t) = 2^t, \tag{10.13}$$

and

$$N_\gamma \simeq N_{e^+} \simeq N_{e^-}. \tag{10.14}$$

The average energy per particle at depth t is

$$E(t) = \frac{E_0}{N(t)} = E_0 2^{-t}, \tag{10.15}$$

and the shower proceeds until the energy of the electrons or positrons is less than the critical energy E_c for bremsstrahlung, i.e. when $E(t) \leq E_c$. At this point, the penetration depth is maximum, $t_{\max}$, and inverting (10.15) gives:

$$t_{\max} \simeq 1.4 \ln\left(\frac{E_0}{E_c}\right). \tag{10.16}$$

For instance, given a material with $E_c = 50\,\text{MeV}$ traversed by a photon with initial energy $E_0 = 100\,\text{GeV}$, the corresponding shower depth is $t_{\max} \simeq 11$. Notice that the depth increases with the logarithm of the initial energy. The maximum number of particles is

$$n_{\max} = 2^{t_{\max}} = E_0/E_c, \tag{10.17}$$

also the total integrated track length is proportional to the energy E_0. The Fig. 10.8 shows a simulation of an electromagnetic shower induced by a 30 GeV electron passing through iron. The plot reports the fractional energy deposition per radiation

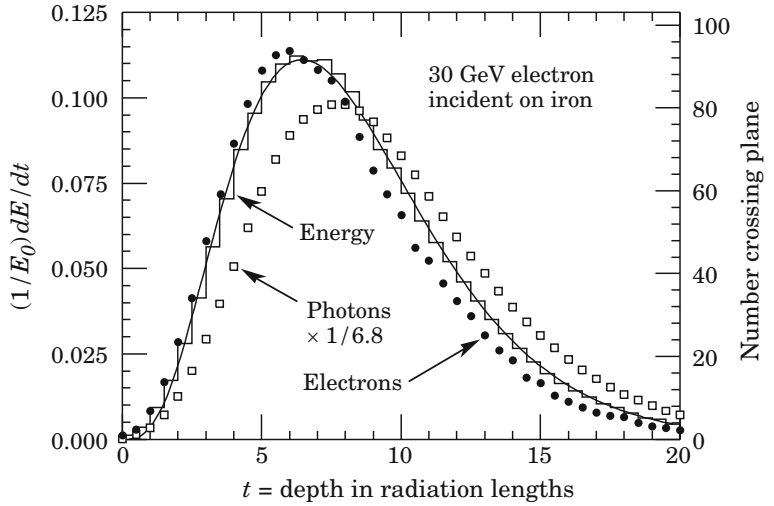


Fig. 10.8 Simulation of an electromagnetic shower induced by a 30 GeV electron in iron. The histogram is the fractional energy deposition per radiation length normalised by the initial energy as a function of the depth of the shower. The *curve* is a fit to the histogram with a Gamma function. The *filled circles* and *open squares* are the number of electrons and photons, respectively, crossing planes at $X_0/2$ intervals. Printed from [2] with permission

length, normalised by the initial energy, as a function of the depth of the shower. Also indicated are the number of electrons and photons crossing planes at $X_0/2$ intervals.

Example: Consider an electron with energy of 30 GeV incident on iron. Since the critical energy for iron is 21.68 MeV (for electrons), the simplified model of (10.16) would yield a maximum shower depth of:

$$t_{\max} \simeq 1.4 \ln \left(\frac{30\,\text{GeV}}{0.02168\,\text{GeV}} \right) = 10.1 \tag{10.18}$$

which is not too dissimilar form the simulation of Fig. 10.8.

10.4 Electromagnetic Calorimeters

The characteristics of the electromagnetic showers are exploited in *electromagnetic calorimeter* detectors, or simply *EM calorimeters*. The Fig. 10.9 shows the simulation of a 20 GeV electromagnetic shower developing within a lead glass crystal [7]. A similar crystal from the CMS experiment was shown in the picture of Fig. 9.4.

Calorimeters are detectors designed to measure energy through the total absorption of particles. The energy will be deposited in various forms and the signal produced in the detector volume, being light or electric charges, is proportional to the

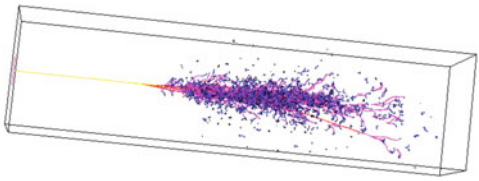


Fig. 10.9 Simulation of a 20 GeV electron shower developing within a lead glass crystal [7]

total deposited energy. Determining the proportionality constant or constants is part of the *calibration* of the detector. In general, the energy deposited E is proportional to the number of interactions occurring in the detector volume n

$$E \propto n, \tag{10.19}$$

and therefore the width of the energy deposited, σ_E, as in a Poisson process is

$$\sigma_E \propto \sqrt{n}. \tag{10.20}$$

Combining (10.19) and (10.20) gives:

$$\frac{\sigma_E}{E} \sim \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}} \sim \frac{1}{\sqrt{E}}. \tag{10.21}$$

Therefore, according to (10.21), the energy resolution of a calorimeter improves with higher energy, which is reasonable as more deposited energy implies more interactions. This behaviour of the resolution is opposite to what happens with momentum (p) measurements, for which the measurement degrades with higher momentum:

$$\frac{\sigma_p}{p} \sim p, \tag{10.22}$$

as discussed in Sect. 8.5 and (8.24).

Example: the energy resolution of the ATLAS electromagnetic calorimeter can be found expressed as

$$\frac{\sigma_E}{E} \simeq \frac{0.1}{\sqrt{E}}, \tag{10.23}$$

which corresponds to a resolution of 1% for energy $E = 100\,\text{GeV}$. The ATLAS momentum resolution is

$$\frac{\sigma_p}{p} \approx 5 \times 10^{-4}\, p \tag{10.24}$$

which corresponds to 5% at $p = 100\,\text{GeV}$.

In a large detector complex, calorimeters are set to work alongside tracking detectors. They have complementary resolution capabilities, and allow the detection of neutral particles and hadronic jets originated from quarks and gluons. Additionally, calorimeters can cover large areas, provide fast signals (from 1 to 10 ns), and their segmentation in depth allows the separation of hadrons, such as p, n, $\pi^{\pm}$, from electromagnetic particles (γ, $e^{\pm}$).

In terms of technology and design, EM calorimeters are designed according to two broad classes: homogeneous or sampling. Homogeneous calorimeters are constructed in such a way that the whole detector volume is filled by a high-density material working as both absorber and sensing medium. The signal might be detected through different physics processes. Examples include materials such as BGO and BaF_2 (detecting scintillation light), lead glass (detecting Cherenkov light) or Liquid noble gases (detecting ionisation charge). Such calorimeters have a very good energy resolution but are normally quite expensive. The CMS lead tungstate calorimeter [8] is an example of an homogeneous calorimeter, which detects light via avalanche photodiodes.

Sampling calorimeters are constructed alternating layers of absorber and active material (they are also called sandwich calorimeters, for a sketch see Fig. 10.10). Typical absorbers include iron, lead or uranium, while the sensing material is plastic scintillators, silicon, liquid or gaseous ionisation detectors. Their resolution is normally worse than homogeneous calorimeters, due to the fluctuations in the number of particles reaching the active regions, and owing to the fact that only part of the deposited energy is actually detected. However, they can be very compact and are normally less expensive than their homogeneous counterparts. A recent example, but their use has been widespread for some time, include the liquid argon calorimeter of the ATLAS experiment [9] (see Fig. 10.11).

Although the energy resolution of calorimeters is at first order described by (10.21), it is more accurately parametrised with three terms

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{10.25}$$

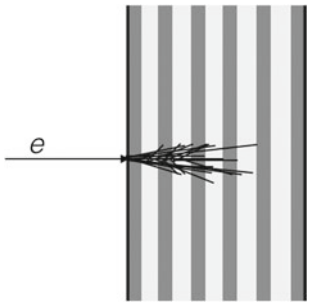


Fig. 10.10 Schematic diagram of a sampling calorimeter, with alternating layers of absorber and active material

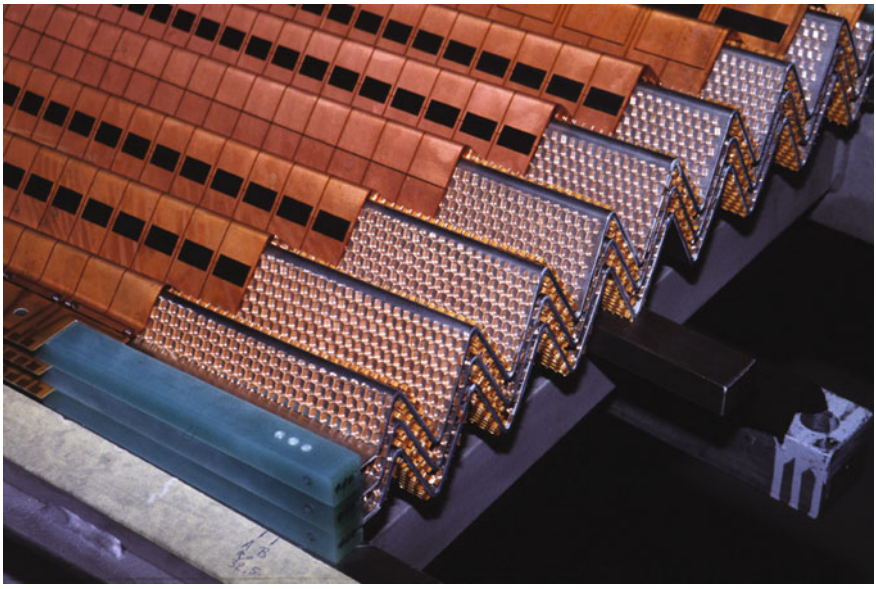


Fig. 10.11 A detail of the ATLAS liquid argon electromagnetic calorimeter, showing the honeycomb structure and the *accordion-shaped* electrodes. The gaps are filled with liquid argon which acts as sensing medium. Credit © CERN, printed from [10] with permission

Table 10.2 Electromagnetic calorimeters and their resolution in some recent collider experiments. The (-h) and (-s) in the type column indicate homogeneous or sampling design, respectively

Experiment	Design Year	Type	Resolution (σ_E/E)	Depth (X_0)
CDF	1988 [11, 12]	Scint./Pb-s	$13.5\%/\sqrt{E\,(\text{GeV})}$	18
BaBar	1995 [13]	CsI(Tl)-h	$2.3\%/E^{1/4}(\text{GeV}) \oplus$ 1.4% [14]	16.0–17.5
CMS	1997 [8]	$PbWO_4$-h	$2.8\%/\sqrt{E\,(\text{GeV})}$ $\oplus 12\%/E\,(\text{GeV}) \oplus$ 0.3% [15]	25
ATLAS	1996 [9]	LAr/Pb-s	$10\%/\sqrt{E\,(\text{GeV})}$ $\oplus 30\%/E\,(\text{GeV}) \oplus$ 0.4% [16]	25

where a, b and c are constants, the $\oplus$ symbol indicates a sum in quadrature, and the three terms are due to energy fluctuations, electronics noise and shower leakage (the shower not being fully contained in the detector volume), respectively. The Table 10.2 shows the resolution features of some recent calorimeters in operations at collider experiments.

10.5 Hadronic Showers and Calorimeters

Hadronic charged particles (protons, pions, kaons, etc.) are subject to energy loss via ionisation of atoms in the material they encounter, as discussed in Chap. 6. Furthermore, they undergo Compton electromagnetic scattering off nuclei, but such effect is relatively small. What is not negligible is that, from energies of the order of 10^9 eV, they undergo significant inelastic collisions with nuclei, and nuclear interactions, in which new particles (of which pions are the lightest and most probable) are formed. An example reaction is

$$p + \text{Nucleus} \rightarrow \pi^+ + \pi^- + \pi^0 + \cdots + \text{Nucleus}^*, \tag{10.26}$$

where $*$ indicates an excited nucleus and the de-excitation can lead to further particles, including also neutrons, α and nuclear fragments. The parameter used to describe nuclear collisions is the *Interaction length*, λ_{int},

$$\lambda_{\text{int}} = \frac{1}{\sigma\, n}, \tag{10.27}$$

where σ is the cross section for the hadronic interaction and n is the number of atoms per unit volume in the target material. The net result of this process is the development of an hadronic shower, in which π^+, π^-, π^0 are abundantly produced in roughly equal numbers, and until the energy is sufficient to produce new particles. The threshold energy for the development of the shower is given by about twice the pion mass

$$E_{\text{th}} \approx 2m_\pi = 0.28\,\text{GeV}. \tag{10.28}$$

The decay $\pi^0 \rightarrow \gamma\gamma$ also leads to the development of an electromagnetic shower component within the hadronic shower. Since λ_{int} is much larger than X_0, the longitudinal dimension of an hadronic shower is larger than for an electromagnetic shower of corresponding energy. The transverse size is also generally larger than electromagnetic showers.

Calorimeters for hadronic showers are designed with similar considerations to electromagnetic calorimeters. The larger shower size means however that more depth is necessary to contain all the energy. Moreover, nuclear binding and re-arrangement effects participate in the energy balance so that the fraction of energy detected is normally smaller than in electromagnetic calorimeters. Stochastic fluctuation of this energy component will limit the resolution, so even though the resolution function of hadronic calorimeters is similar to (10.25)

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b, \tag{10.29}$$

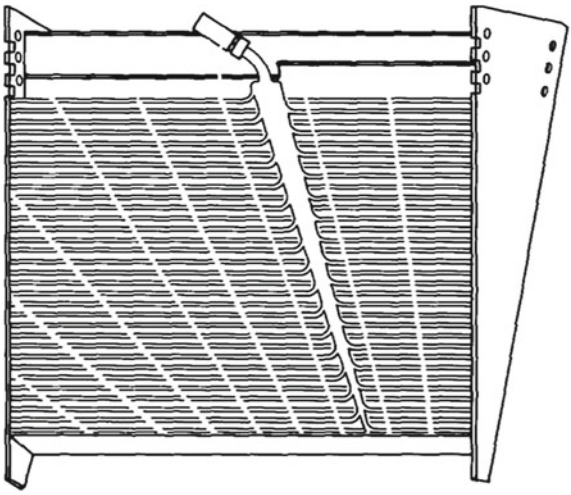


Fig. 10.12 Module of the central hadronic calorimeter of the CDF experiment, showing the alternating layers of steel and plastic scintillator and a sketched light guide. Reprinted from [17] with permission

with a stochastic coefficient a and a constant resolution term b, the coefficient a is in fact several times larger than for an electromagnetic calorimeter. In other words, hadronic calorimetry is less precise than electromagnetic because the medium, that is the nuclei, is an active part of the underlying processes that drive the fluctuations of energy. The Fig. 10.12 shows a diagram of the central hadronic calorimeter of the CDF experiment [17]. The calorimeter is a sampling of 48 steel and scintillator layers providing a depth of $\lambda_{\text{int}} \simeq 4.7$.

Glossary

Bhabha scattering Positron scattering off atomic electrons with energy loss per collision below 0.255 MeV

Bremsstrahlung Emission of photons by a charged particle consequent to the acceleration of an electric charge

Compton scattering Scattering of a photon off an atomic electron

Critical energy (electron) Energy of an electron for which the rate of energy loss by bremsstrahlung equals the rate of energy loss by ionisation

Critical energy (positron) Energy of a positron for which the rate of energy loss by bremsstrahlung equals the rate of energy loss by ionisation

Electron–positron annihilation The reaction $e^+e^- \rightarrow \gamma\,\gamma$, in which matter and anti-matter is turned into photons

Electromagnetic calorimeter Detector designed to measure the energy of electrons, positron or photons by total absorption

Electromagnetic shower Chain reaction initiated by an energetic electron, positron or photon into a substance, in which multiple electrons and photons are created or liberated

Energy resolution (calorimeter) Ratio between the Gaussian width of the energy response and the value of the input energy

Hadronic shower Chain reaction initiated by an energetic hadron through the inelastic collisions with nuclei, and nuclear interactions.

Homogeneous calorimeter Calorimeter designed such that the whole detector volume is filled by a high-density material working as both absorber and sensing medium

Interaction length Nuclear interaction length is the distance in a given substance after which an incident number of hadrons diminishes by a factor $1/e$ by nuclear interaction

Møller scattering Electron scattering off atomic electrons with energy loss per collision below 0.255 MeV

Pair production Production of one electron and one positron form the disappearance of one photon, in the field of an atomic nucleus

Photoelectric effect Absorption of a photon by an atom and the subsequent emission of an electron

Radiation length Mean thickness of a material (in g/cm^2) after which an electron or positron energy is reduced by a fraction of $1/e$ of its initial energy by bremsstrahlung only

Radiative energy loss see bremsstrahlung

Sampling calorimeter Calorimeter designed alternating layers of absorber (e.g. iron, lead or uranium) and active material plastic scintillator, silicon, liquid or gaseous ionisation medium)

Stochastic coefficient Contribution to the resolution of a calorimeter due to stochastic effects (typically the energy deposition and detection)

References

1. *Figure 3.2 from Messel and Crawford, Electron-Photon Shower Distribution Function Tables for Lead, Copper, and Air Absorbers* (Pergamon Press, Oxford, 1970)
2. J. Beringer et al., (Particle Data Group). Phys. Rev. D **86**, 010001 (2012)
3. K.A. Olive et al. (Particle data group). Chin. Phys. C, **38**, 090001 (2014); and 2015 update, Atomic and Nuclear Properties of Materials, http://pdg.lbl.gov/2015/AtomicNuclearProperties/
4. C. Amsler et al., Particle data group. Phys. Lett. B **667**, 1 (2008)
5. W.R. Nelson et al., Phys. Rev. **149**, 201 (1966)
6. G. Bathow et al., Nucl. Phys. B **20**, 592 (1970)

7. Electromagnetic Shower Simulator, https://www.mppmu.mpg.de/~menke/elss/home.shtml. Accessed 7 Dec 2016
8. CMS Collaboration, The CMS electromagnetic calorimeter project: Technical Design Report, CERN/LHCC/97-03 (1997)
9. ATLAS Collaboration, ATLAS liquid-argon calorimeter: Technical Design Report, CERN/LHCC/96-041 (1996)
10. Copyright CERN, http://cds.cern.ch/record/39737. Accessed 12 Dec 2016
11. F. Abe et al., Nucl. Instrum. Methods A **271**, 387 (1988)
12. L. Balka et al., The CDF central electromagnetic calorimeter. Nucl. Instrum. Methods Phys. Res. Sect. A **267**(23), 272–279 (1988)
13. B. Aubert et al., The BaBar detector, BABAR collaboration. Nucl. Instrum. Methods A **479**, 1 (2002)
14. M. Kocian, Performance and calibration of the crystal calorimeter of the BABAR detector, in *Proceedings of the 10th International Conference on Calorimetry in Particle Physics (CALOR 2002)* (Pasadena, CA, 2002), pp. 167–174
15. C.M.S. Collaboration, P. Adzic et al., Energy resolution of the barrel of the CMS electromagnetic calorimeter. JINST **2**, P04004 (2007). doi:10.1088/1748-0221/2/04/P04004
16. ATLAS Collaboration, The ATLAS liquid argon calorimeter: overview and performance. J. Phys. Conf. Ser. **293**, 012044 (2011). doi:10.1088/1742-6596/293/1/012044
17. S. Bertolucci et al., The CDF central and endwall Hadron calorimeter. Nucl. Instrum. Methods Phys. Res. Sect. A **267**(23), 301–314 (1988)

Chapter 11
Cherenkov and Transition Radiation: Detectors for PID and Neutrinos

Abstract Cherenkov and Transition Radiation are forms of interaction of radiation with matter that are actively exploited for their unique features. The uniqueness arises from the strong dependence of these mechanisms on the relativistic γ of the interacting particle. Hence particle identification (PID) is the main advantage in exploiting such fascinating effects. The capability of detecting elusive particles such as the neutrinos has also been significantly advanced in the past decades thanks to the detection of Cherenkov light from the neutrino interaction chain, and this is the topic of the last section.

11.1 Cherenkov Radiation

Cherenkov radiation is the result of a charged particle travelling in a medium at a speed that is faster than the speed of light in that medium. This is possible because the speed of light in a medium is generally lower than c, it is in fact c/n where n is the index of refraction. Therefore, in formulae, Cherenkov radiation is produced if the speed v of a charged particle is

$$v > \frac{c}{n}, \tag{11.1}$$

and since the relativistic β is defined as $\beta = v/c$ this condition can also be written as

$$\beta > \frac{1}{n}. \tag{11.2}$$

If the condition of (11.2) is satisfied, an electromagnetic shockwave builds up from the particle as in the analogue situation of an object travelling supersonically in the case of a sonic boom (see Fig. 11.1). Let x indicate the distance covered by the particle in a time interval t ($x = vt$), and x' the distance travelled by the electromagnetic wave in the same interval. The Cherenkov angle θ_C between the direction of propagation of the light and the direction of motion of the particle is given by:

L. Cerrito, *Radiation and Detectors*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-53181-6_11

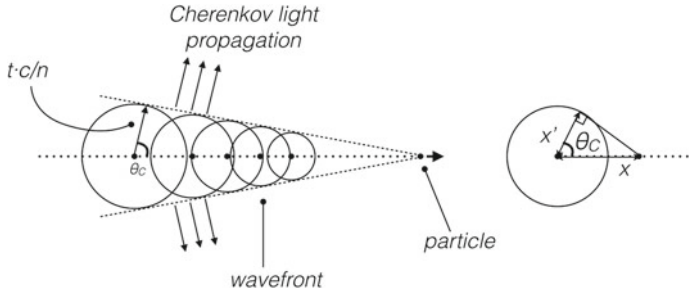


Fig. 11.1 Sketch of the wavefront of Cherenkov radiation produced by a moving charged particle

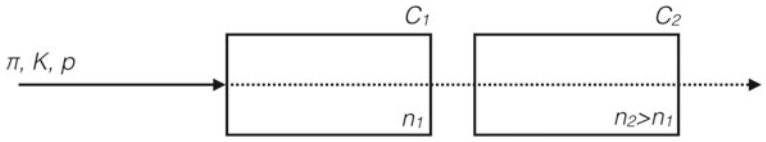


Fig. 11.2 Diagram of a sequence of Cherenkov detectors used for particle identification

$$\cos\theta_C = \frac{x'}{x} = \frac{ct}{n}\frac{1}{vt} = \frac{1}{\beta n(\nu)}, \tag{11.3}$$

where n has been written explicitly as a function of the emitted light frequency (ν), and the (11.3) is valid for an infinite radiating medium.

The energy loss through Cherenkov radiation is small compared to other mechanisms of interaction, it is of the order of 1/1000th of the energy loss by ionisation. Furthermore it is already included in the Bethe-Block formula ((6.12) in Chap. 6). The emitted light has wavelengths between the ultraviolet and the infrared, and assuming n is constant for a range of wavelengths, the number of photons produced per unit path is:

$$\frac{dN}{dx} = 2\pi\alpha\sin^2\theta_C\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right), \tag{11.4}$$

where α is the fine structure constant. For example, in water $n = 1.33$, $\theta_C \simeq 41°$, and considering only λ between 350 and 500 nm the light output is 100 photons/cm, with an energy release of about 100 eV/cm.

The features of the Cherenkov radiation have long been used in particle detectors to distinguish different charged particles, exploiting the β dependence of the Cherenkov angle in (11.3). One simple technique is the threshold detection shown in a simplified form in Fig. 11.2. The emission of Cherenkov radiation is subject to the condition of (11.2), that is the particle β has to pass a certain threshold value:

$$\beta > \beta_{\text{thresh.}} = \frac{1}{n}, \tag{11.5}$$

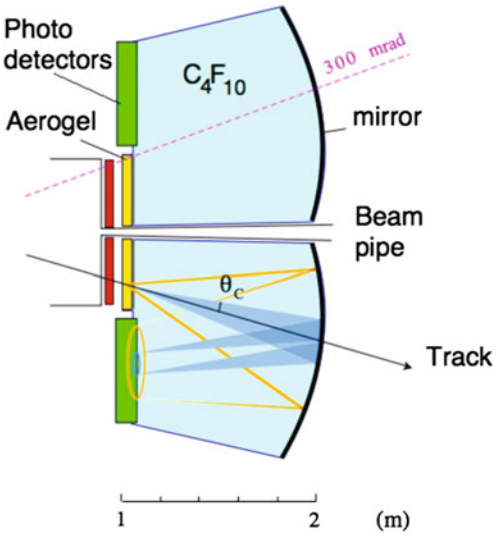


Fig. 11.3 Layout of the LHCb RICH1 detector, illustrating the focusing of Cherenkov light. Printed from [1] with permission

and recall that

$$\beta = \frac{v}{c} = \frac{p}{m_0 \gamma c}, \tag{11.6}$$

where m_0 is the rest mass of the particle moving with momentum p. Consider a beam of pions, kaons and protons all with momentum p entering a sequence of two Cherenkov detectors (C_1 and C_2) with index of refraction n_1 and n_2, respectively, and with $n_2 > n_1$. If n_1 and n_2 are chosen such that

$$n_2 : \qquad \beta_p < \frac{1}{n_2} < \beta_\pi, \beta_K \tag{11.7}$$

$$n_1 : \qquad \beta_K, \beta_p < \frac{1}{n_1} < \beta_\pi \tag{11.8}$$

then Cherenkov light detected in C_1 and in C_2 implies an identified pion. Conversely, light in C_2 but not in C_1 implies an identified kaon. Finally, from light neither in C_1 nor in C_2 follows the conclusion that the particle was a proton.

An example of a recent Cherenkov detector is the Ring Imaging CHerenkov (RICH) [1] at the LHCb experiment [2]. It consists of a silica aerogel transit medium (see Fig. 11.3), with refraction index $n = 1.03$, suitable for the lowest momentum tracks, and a gaseous C_4F_{10} for the intermediate momenta, both composing the RICH1 apparatus. The highest momentum particles, up to 150 GeV, radiate through gaseous CF_4 in an additional detector, the RICH2. Travelling particles emit Cherenkov radiation making a specific angle θ_C with the particle trajectory. The Cherenkov light is reflected by a set of mirrors back into photodetectors where it projects into circles which are reconstructed and measured. The plot in Fig. 11.4

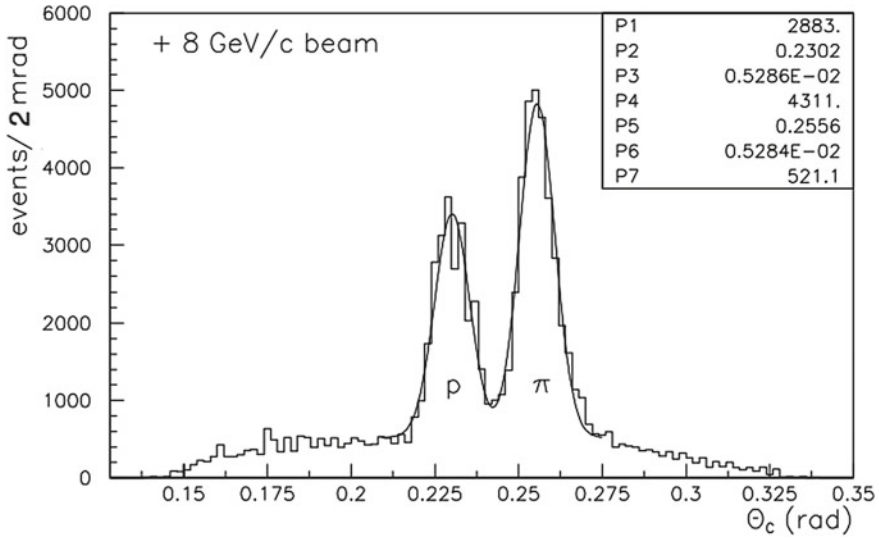


Fig. 11.4 Particle identification through Cherenkov angle measurements, allowing the separation of protons and pions at the LHCb RICH1-prototype detector. Reprinted from [1] with permission

shows the Cherenkov angles measured from the rings of 8 GeV pions and protons in the RICH1 detector, indicating a fairly clear separation between two types of particle.

11.2 Transition Radiation

Transition radiation is emitted by a charged particle when it crosses between two materials with different relative static permittivities. Relative permittivity is the ratio between the electric energy stored by, e.g. an applied voltage in a material and the electric energy stored by the same voltage in vacuum. During the transition between the materials, the electric field of the particle is rapidly rearranged and this is accompanied by the emission of electromagnetic (transition) radiation. For highly relativistic particles, the radiation is mostly emitted at frequencies of the X-ray region and it is sharply peaked in the direction of the particle motion. The intensity of the radiation is proportional to the relativistic γ of the particle [3, 4].

Since the total radiation yield of the transition radiation at a single interface is relatively small, the transition radiation is often exploited in detectors using a *radiator*, that is a sequence of several material interfaces. An example of a radiator is represented by of a stack of thin foils of low-Z material interleaved with air. With such an arrangement, a charged particle going through the radiator would traverse many interfaces, emitting coherent radiation proportional to the number of interfaces.

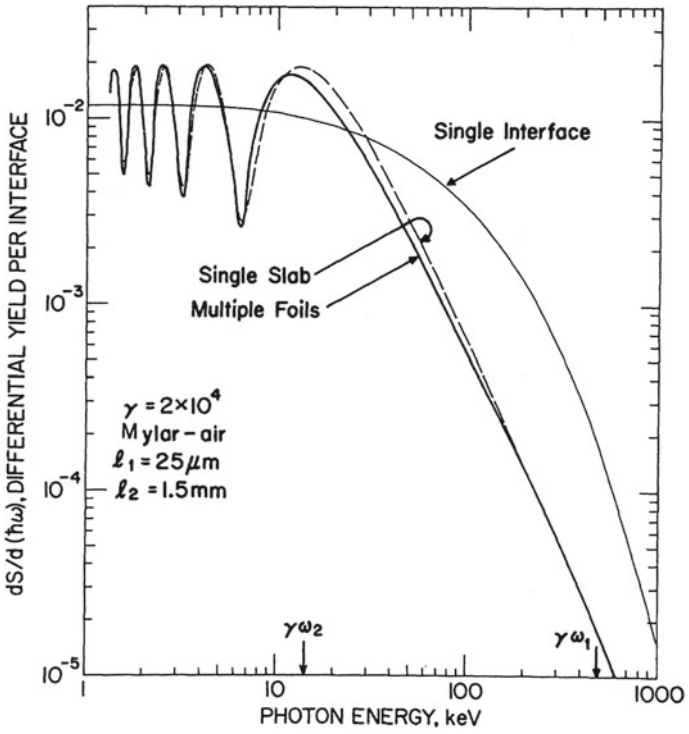


Fig. 11.5 Computed differential distribution of the transition radiation photon energy, per interface, for a radiator consisting of a single interface or multiple foils. Reprinted from [3] with permission

Moreover, interference effects will lead to a modulation of the frequency distribution, as shown in Fig. 11.5.

Due to the dependence of the transition radiation on the Lorentz factor $\gamma = E/m_0c^2$ (where m_0 is the rest mass of the particle), there is a wide momentum range, say between 1 and 100 GeV/c, where electrons are the only particles producing transition radiation. At higher energies, kaons can be separated from pions in a momentum range between about 200 and 700 GeV/c. An example of exploitation of transition radiation for detection and particle identification is the Transition Radiation Tracker (TRT) [5, 6] at the ATLAS experiment. In the ATLAS detector layout, the TRT surrounds the inner tracking system with a 52 cm-long sequence of individual tubular drift chambers, with 4 mm diameter, interleaved by fibers or foils creating a sequence of interfaces (see Fig. 11.6). Electrons can be identified by the transition radiation they produce, while all particles are tracked by the individual signals of the drift tubes.

One more example of a transition radiation detector is the TRansition Detector (TRD) at the ALICE experiment [7]. In the ALICE detector layout, the TRD surrounds the time projection chamber and consists of over 500 drift chambers inter-

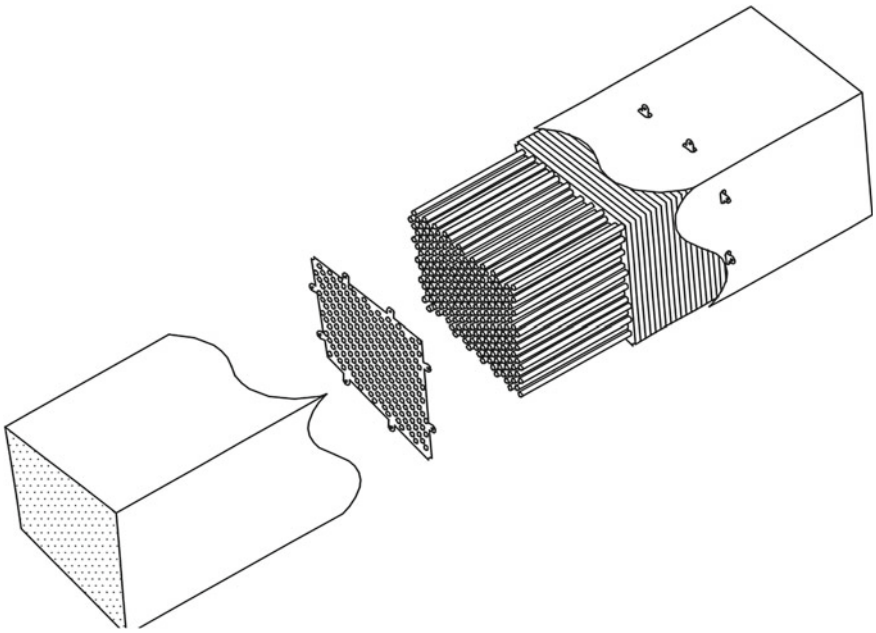


Fig. 11.6 Module of the ATLAS Transition Radiation Detector, showing the individual drift tubes and the stacks of polypropylene/polyethylene fibre radiators. Reprinted from [6] with permission

leaved with radiators made of foam and fiber materials. A schematic of one TRD chamber is shown in Fig. 11.7. Charged particles cross the chamber perpendicularly to the plane of the anode wires. An electric field allows for the multiplication and collection of the ionised gas charges (for ionisation detectors see Chap. 7). Transition radiation is produced by electrons crossing the radiator and will cause an afterpulse in the drift chamber at later drift times, hence allowing the separation of electrons from heavier charged particles.

11.3 Detecting Neutrinos

Neutrinos carry no electric charge and can be detected only via the weak interaction through the exchange of a W boson (charged current) or a Z boson (neutral current). Examples are shown in Fig. 11.8 and in the following list:

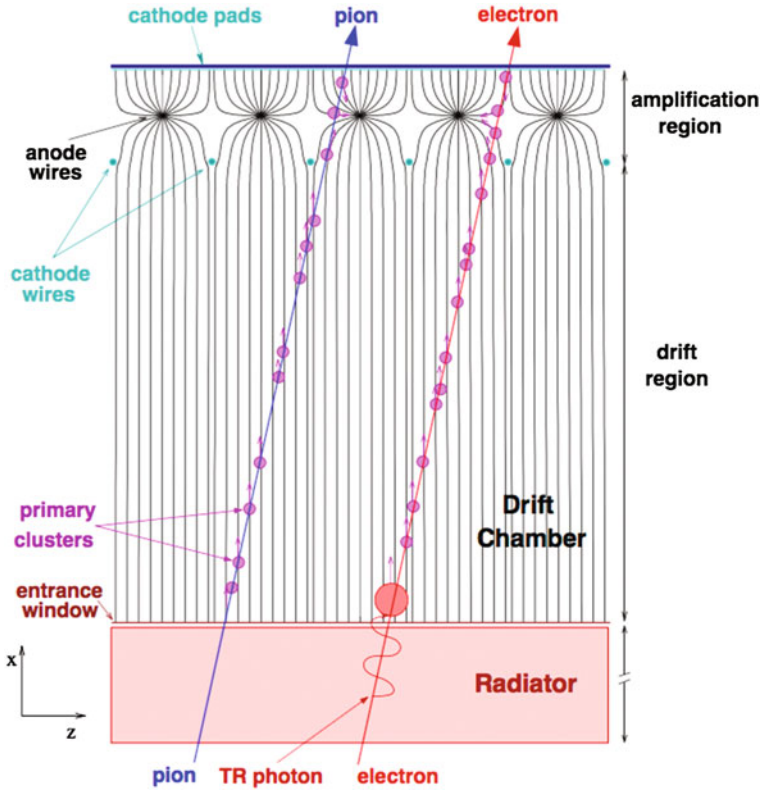


Fig. 11.7 Schematic illustration of a portion of the ALICE transition radiation detector. The electric charges produced by energy loss and by transition radiation absorption move along the drift region producing a detectable signal. Reprinted from [8] with permission

$$\nu_e + n \rightarrow e^- + p \quad \text{(charged current)}, \tag{11.9}$$
$$\bar{\nu}_e + p \rightarrow n + e^+ \quad \text{(charged current)}, \tag{11.10}$$
$$\bar{\nu}_e + e^- \rightarrow \mu^- + \bar{\nu}_\mu \quad \text{(charged current)}, \tag{11.11}$$
$$\nu_e + e^- \rightarrow \nu_e + e^- \quad \text{(neutral current)}, \tag{11.12}$$
$$\nu_\mu + e^- \rightarrow \nu_\mu + e^- \quad \text{(neutral current)}, \tag{11.13}$$
$$\nu_\tau + e^- \rightarrow \nu_\tau + e^- \quad \text{(neutral current)}. \tag{11.14}$$

Due to low interaction cross sections, neutrino detectors usually need large volumes and an high degree of shielding from cosmic and terrestrial backgrounds. The detection is based on the sensitivity to the charged particles or neutrons produced in the electroweak interactions. The radiochemical methods were highlighted in Chap. 2 with the description of the Homestake experiment, and are based on the measurement of the concentration of specific radioactive elements produced by a charged current

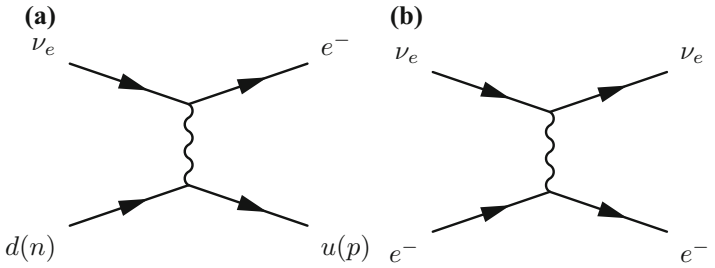


Fig. 11.8 Leading order Feynman diagrams of neutrino interactions. The charged current W exchange **a** between the neutrino and a quark in the nucleon, and the neutral current Z exchange **b** between the neutrino and an atomic electron

interaction with the detector material (e.g. chlorine, gallium). These detectors provide counting but no information on the kinematics or arrival time of the neutrinos. When positrons are liberated, as in the interaction of an electron anti-neutrino with a proton producing a neutron and a positron (the inverse beta decay), the resulting electron-positron annihilation process is easily detectable. The annihilation reaction generates two photons in coincidence, which can be detected with a scintillator medium surrounded by photomultipliers.

The choice of technology for neutrino detection is actually owed in part to the sensitivity requirements in the energy range of interest. At intermediate neutrino energies, perhaps the most widely used technique is the detection of Cherenkov radiation from the charged leptons liberated by the neutrino interaction and moving in the detector medium. The Super-Kamiokande [9] detector for instance uses a large cylindrical tank (39 m in diameter and 42 m of height) filled with 50,000 tons of ultra pure water surrounded by 13,000 photomultipliers, whose purpose is to detect the Cherenkov light from electrons or muons. Electrons and muons are observed as circular light patterns on the detector walls from the Cherenkov cone, and are distinguished by the sharp (muons) or fuzzy (electrons) edges in the light ring, an effect caused by multiple scattering.

Exploiting Cherenkov light is also the basis for the technique used by the Sudbury Neutrino Observatory (SNO) [11], which employed heavy water (^2H_2O) in a spherical vessel contained within a large cavity filled with ultra pure ordinary water. The photograph in Fig. 11.9 shows the photomultipliers viewed from the bottom of the spherical vessel. The heavy water allows the neutrino-deuterium (D) reactions:

$$\nu_e + D \rightarrow p + p + e^- \quad \text{(charged current)}, \tag{11.15}$$

$$\nu_x + D \rightarrow p + n + \nu_x, \quad \text{(neutral current)}, \tag{11.16}$$

in which the deuterium nucleus is broken up into separate nucleons. The charged current reaction is observed through the Cherenkov light of the escaping electron, while the neutral current, produced by all neutrino flavours, can be detected via photons produced by capture of the the thermalised neutron.

Fig. 11.9 Photomultipliers of the SNO acrylic vessel, viewed from the bottom of the spherical vessel with a fish-eye lens. Courtesy of Ernest Orlando Lawrence Berkeley National Laboratory. Printed from [10] with permission

There are many more experiments using Cherenkov light for neutrino detection and their discussion goes beyond the scope of this textbook, however instructive examples worth mentioning include the Antarctic Muon And Neutrino Detector Array (AMANDA) [12, 13] and its successor IceCube, the Astronomy with a Neutrino Telescope and Abyss environmental RESearch (ANTARES) [14] and the future KM3NeT, and MiniBooNE [15] at Fermi National Accelerator Laboratory.

Glossary

Cherenkov light Electromagnetic radiation emitted by a charged particle travelling in a medium at speed faster than the speed of light in that medium

Cherenkov angle Angle between the direction of motion of a particle and the direction of the emitted Cherenkov light
PID detector A particle detector capable of distinguishing charged particle types, e.g. electrons from hadrons
Transition radiation Electromagnetic radiation emitted by a charged particle traversing between two substances with different relative static permittivities
Radiator Sequence of different material interfaces arranged to produce detectable transition radiation
Charged current (neutrinos) Neutrino weak interaction process mediated by a W boson
Neutral current (neutrinos) Neutrino weak interaction process mediated by a Z boson
Radiochemical neutrino detection Neutrino detection method based on the measurement of the concentration of specific radioactive elements, produced in the interaction between the neutrino and the detector medium
Cherenkov neutrino detection Neutrino detection method based on the detection of Cherenkov light from charged leptons, liberated by the neutrino interaction, and moving in the detector medium

References

1. LHCb Collaboration, RICH Technical Design Report, CERN LHCC 2000-037 (2000), https://cds.cern.ch/record/494263?ln=en Accessed 9 December 2016
2. LHCb Collaboration, The LHCb Detector at the LHC, JINST 3 (2008) S08005
3. M.L. Cherry et al., Phys. Rev. D **10**, 3594 (1974)
4. M.L. Cherry, Phys. Rev. D **17**, 2245 (1978)
5. ATLAS TRT Collaboration, The ATLAS TRT Barrel Detector, J. Instrum. 3 (2008)
6. ATLAS Collaboration, ATLAS inner detector: Technical Design Report, 2, ATLAS-TDR-5; CERN-LHCC-97-017 (1997)
7. The ALICE Collaboration, ALICE TRD Technical Design Report, CERN/LHCC 2001–021. ALICE TRD **9**, 3 (2001)
8. M. Kweon, the ALICE TRD Collaboration, The Transition Radiation Detector for ALICE at LHC, Nucl. Phys. A **1–4**(830), 535c–538c (2009). doi:10.1016/j.nuclphysa.2009.10.047
9. J.F. Beacom, P. Vogel, Phys. Rev. D 58, 053010. Cosmic Ray Research, The University of Tokyo (1998), http://www-sk.icrr.u-tokyo.ac.jp/sk/gallery/index-e.html
10. SNO Image Catalog, http://www.sno.phy.queensu.ca/sno/images/, Accessed 9 Dec 2016
11. N. Jelley, A.B. McDonald, R.G. Hamish Robertson, Ann. Rev. Nucl. Part. Sci **59**, 431–465 (2009)
12. F. Halzen, J.E. Jacobsen, E. Zas, Phys. Rev. D **53**, 7359 (1996)
13. J. Ahrens et al., Astropart. Phys. **16**, 345 (2002)
14. A Deep Sea Telescope for High Energy Neutrinos, ANTARES Collaboration (1990), arXiV:astro-ph/9907432
15. E. Church et al., A proposal for an experiment to measure $\nu_\mu \rightarrow \nu_e$ oscillations and ν_μ disappearance at the Fermilab Booster: BooNE, FERMILAB-P-0898; I. Stancu et al., The MiniBooNE Detector Technical Design Report, http://www-boone.fnal.gov

Appendix
Statistics and Probability

Abstract Statistics and probability distributions are an important aspect of radiation interaction processes, owing to the intrinsic stochastic nature of interactions. This Appendix is intended to collect in one place a number of related definitions, concepts and formulae on statistics and probability distributions. The style is very compact but the four sections on probability and probability distribution functions can be used to recall the formalism, basic statistics and the mathematical form of common probability distributions.

A.1 Sample Space and Probability

Consider a certain phenomenon that might or might not happen. That is, the occurrence of such phenomenon could be one of three situations: *certain*, meaning the process does happen with certainty; it could be *impossible*, that is it does not happen at all; or *random*, meaning it might or might not happen. Examples of such phenomena can be recognised from throwing a dice: if we are interested in the outcome corresponding to the draw of any number between 1 and 6, then the phenomenon is certain. It is certain that in throwing a standard dice the outcome will be some number between 1 and 6. On the other hand, it is impossible to draw 0 from a standard dice, so this would be an example of an impossible outcome. Finally, if the outcome of interest is defined as obtaining the number 3 from one single trial, then such phenomenon is random. Statistics deals with random processes giving them a degree of predictability. The *Sample Space* is defined as the set of all possible outcomes of an experiment, a trial, or a phenomenon. The *Event* is defined as a subset of the sample space, namely it is a set of outcomes. Let *Elementary Event* be defined as one single outcome in the sample space. A couple of examples should help place in focus such definitions. Two coins are tossed and such trial could lead to four possible outcomes: two tails, two heads, or 1 tail and 1 head with the possible permutation of the order. These four outcomes define the sample space; each outcome is an elementary event, while the outcomes with at least one head (3 elementary events) define an event.

L. Cerrito, *Radiation and Detectors*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-53181-6

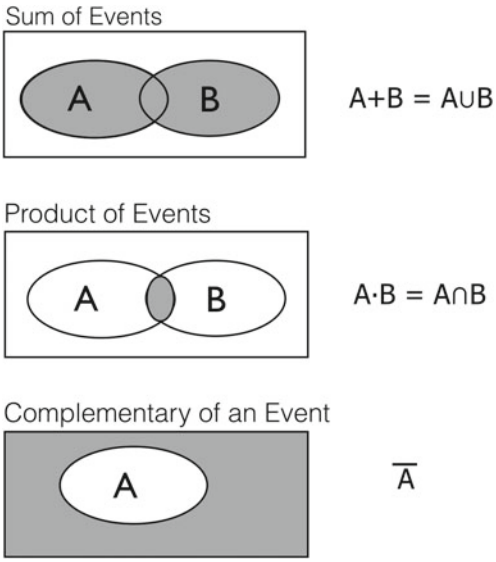


Fig. A.1 Algebra operations of sample spaces A and B into C, showing the sum, the product and the complementary. In shaded colour is indicated the set of elementary events belonging to the resulting event space, C

A second example is given by rolling one dice; the possible outcomes are integer numbers between 1 and 6, therefore such set defines the sample space. If the outcome of interest is the drawing of any number greater than 3, this defines an event, composed of three elementary events (the numbers 4, 5 and 6).

It is possible to define an Algebra of events, with common operations defined as follows. The space C, *Sum* of two sample spaces, A and B, is defined as the set of elementary events that belong to A or B. The space C, *Product* of two sample spaces, A and B, is defined as the set of elementary events that belong to A and B. The *Complementary* of a sample space A (indicated with $\bar{A}$) is the set of elementary events that do not belong to A. In mathematical formalism (see also Fig. A.1):

$$C = A + B = A \cup B, \tag{A.1}$$

$$C = A \cdot B = A \cap B. \tag{A.2}$$

The definition of Probability is not unique and there are several ways to approach it. There are at least three classes of definitions which are mostly relevant. In the so-called *classical* interpretation, considering a sample space with N elementary events and an event A consisting of a number of elementary events, n, then the probability of A is defined as the ratio

$$P(A) = \frac{n}{N}. \tag{A.3}$$

For example, the probability for obtaining the number 3 from a fair dice would be, $n = 1, N = 6, P(A) = 1/6$. There are clear limitations to such definition. It assumes

a finite number of outcomes and it considers only elementary events as being equally *likely*, besides the problem that the meaning of such likelihood is not itself defined.

A better definition is given by the *frequentist* interpretation of probability. Considering an event A part of a sample space, the frequency of A is defined as

$$F(A) = \frac{m}{M}, \tag{A.4}$$

where m is the number of cases in which A is obtained in a repetition of the outcomes and M is the number of repetitions, or trials. The probability of A is then defined as the limit of the frequency of A for infinite repetitions, namely

$$P(A) = \lim_{M \to \infty} F(A). \tag{A.5}$$

Compared to the classical interpretation of (A.3), the frequentist definition of (A.5) does not require equally likely elementary events, nor the concept of likelihood itself, but it does need an infinite number of repetitions in order to define the probability.

The third class of probability definitions is the *axiomatic basis* (Kolmogorov 1933 [1]), which can be summarised in the following statements: given a sample space S and events A with $A \in S$ (the symbol $\in$ means that A belongs to S):

- For any A, there is one real number $\mathrm{P}(A)$ associated with A such that $P(A) \geq 0$,
- $\mathrm{P}(S) = 1$ and $\mathrm{P}(\oslash) = 0$, where $\oslash$ indicates the impossible event
- If $A \cap B = \oslash$ (the events are mutually exclusive) then $\mathrm{P}(A + B) = \mathrm{P}(A) + P(B)$

These definitions allow deriving a number of important properties routinely used in statistics, whose proof goes beyond the scope of this Appendix, for instance

- if C is the certain event, $\mathrm{P}(C) = 1$
- $\mathrm{P}(\bar{A}) = 1 - \mathrm{P}(A)$
- $\mathrm{P}(A \cup \bar{A}) = 1$.

Moreover, if A and B are compatible and overlapping events then the probability of their sum is

$$P(A + B) = P(A) + P(B) - P(A \cdot B), \tag{A.6}$$

where $A \cdot B$ is the overlapping part of the event.

Conditional probability is also particularly important, calculated as

$$P(A|B) = \frac{P(A \cdot B)}{P(B)}, \tag{A.7}$$

where $P(A|B)$ indicates the probability of event A *after* event B has occurred or else within the set of elementary events belonging to event B. The expression in (A.7) is general but if the events A and B are independent then by definition of independent events the conditional probability of A is simply the probability of A, therefore

$$P(A|B) = P(A) \quad \text{(independent events)}, \tag{A.8}$$

since event A cannot bear any knowledge of B. By comparing (A.7) with (A.8) if follows a useful expression for calculating the probability of overlapping independent events

$$P(A \cdot B) = P(A) \cdot P(B) \quad \text{(independent events)}. \tag{A.9}$$

A.2 Probability Distribution Function

Consider a sample space consisting of a set of N events x_i ($i = 1, \ldots, N$), each with probabilities P(x_i). The collective probabilities P(x_i) are the *probability distribution function* (PDF) of the sample space under consideration. For example, given a sample space of 100 individuals, each being born in one of 4 quarters, January-March (Q_1), April-June (Q_2), July-September (Q_3), October-December (Q_4), the probabilities associated with the birth in each quarter (the events), indicated by P(Q_i) is a probability distribution function. An histogram with $Q_{1,\ldots,4}$ on the horizontal axis and P(Q_i) on the vertical axis would be a graphical representation of this probability distribution function.

Two quantities are extremely useful in the description of a PDF defined as above. Consider once more the sample space X consisting of a set of N events with values x_i, ($i = 1, .., N$) with probabilities P(x_i). The first quantity of interest is the *expectation value*, indicated with $E[X]$

$$E[X] = \sum_{i=1}^{N} x_i P(x_i), \tag{A.10}$$

that is the sum of the products of the event value x_i multiplied by its probability. Notice how (A.10) resembles the definition of average value of a repeatable sequence

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i, \tag{A.11}$$

and it is in fact a conceptual extension of the same idea (or rather it justifies the definition of average), representing the "center of gravity" of the distribution.

The second quantity useful to describe a PDF is the *variance*, defined as

$$var[X] = E\left[(X - E(X))^2\right], \tag{A.12}$$

which addresses the spread of the distribution. This resembles the definition of root-mean-square (RMS) of a sequence of N numbers,

$$RMS^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 \tag{A.13}$$

and this is conceptually the same idea (or, rather, the variance underpins the definition of root-mean-square). It is easy to show that

$$var[X] = E[X^2] - E^2[X], \tag{A.14}$$

which is sometimes a simpler expression to compute than what given in (A.13).

A.3 Binomial and Poisson Distributions

The PDF described in Sect. A.2 are discrete probability distribution functions, meaning that the N events take a discrete set of values x_i. An example of such distribution was discussed with the probability of the quarter of birth in a given population. Some PDF are however so common that have been studied in detail. One such PDF is the *Binomial* probability distribution function,

$$P_{N,p}(n) = \frac{N!}{n!(N-n)!}p^n(1-p)^{N-n} \quad \text{(Binomial distribution)}, \tag{A.15}$$

where N and n are integers, N is the total number of events, n is the value each event takes and p is a parameter of the distribution (a real number). Statistical processes described by a binomial distribution are extremely common, for instance (A.15) describes the probability of obtaining n tails from tossing one coin N times (or equivalently N coins once) and p is the single-trial probability. It can be shown that from the definitions of expectation value and variance applied to the distribution of (A.15), the binomial distribution has the following properties,

$$E[n] = \sum_{n=0}^{N} n P_{N,p}(n) = Np, \tag{A.16}$$

$$var[n] = Np(1-p). \tag{A.17}$$

A discrete PDF of exceptional importance is also the *Poisson* distribution; given a total of N events, the probability of each n event is given by

$$P_\lambda(n) = \frac{\lambda e^{-\lambda}}{n!} \quad \text{(Poisson distribution)}, \tag{A.18}$$

where λ is a real number and a parameter of the distribution. The expectation value and the variance of the Poisson distribution can be derived from the definitions given in (A.10) and (A.12), yielding

$$E[n] = \lambda, \tag{A.19}$$

$$var[n] = \lambda. \tag{A.20}$$

The Poisson distribution describes for instance the number of radioactive decays from a large number of atoms, given a fixed decay probability $p = \lambda/N$. Moreover, the mathematical expression of (A.18) can be derived as the limit of a binomial distribution for large N ($N \to \infty$), small probability p ($p \to 0$) but with their product being a finite number, $Np = \lambda$.

A.4 Continuous Probability Distribution Functions

The probability distributions functions listed in Sect. A.3 are built on an event space consisting of integers (for instance the number of times a tossed coin yields a head, the radioactive decay counts, etc.). The event space can be also a continuum of real numbers, x, and the probability of the events X is defined in such case as

$$P(x < X < x + \mathrm{d}x) = f(x)\mathrm{d}x, \tag{A.21}$$

where $f(x)$ is called the *Probability Density Function*. Clearly, the probability density function must be integrable,

$$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1, \tag{A.22}$$

in order to satisfy the condition that the total probability across the full event space, $P(S)$, equals 1.

A simple example of a continuous probability density function is the *Uniform* distribution, which is a constant in the range of real numbers between boundaries a and b

$$f(x) = \text{constant} = K \quad \text{(Uniform distribution)}, \tag{A.23}$$

where K is the constant value and a parameter of the distribution. The condition of (A.22) yields the following relationships

$$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1 \Rightarrow K(b-a) = 1 \tag{A.24}$$

$$K = \frac{1}{b-a}, \tag{A.25}$$

and therefore a uniform distribution between boundaries a and b has form

$$f(x) = \text{constant} = \frac{1}{b-a}. \tag{A.26}$$

The expectation value and the variance can be derived from the definitions given in (A.10) and (A.12), yielding

$$E[x] = \int_a^b x f(x) \mathrm{d}x = \frac{1}{2}(b+a), \tag{A.27}$$

$$var[x] = \int_a^b (x - E[x])^2 f(x) \mathrm{d}x = \frac{1}{12}(b-a)^2. \tag{A.28}$$

However, arguably the most important probability density function is the *Gaussian* distribution

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{(Gaussian distribution)}, \tag{A.29}$$

where σ and μ are parameters of the function (real numbers) and the normalisation condition is

$$\int_{-\infty}^{+\infty} G(x) \mathrm{d}x = 1. \tag{A.30}$$

The Gaussian function has the characteristic bell shape and some of its features are directly derived from the expectation value and the variance:

$$E[x] = \int_{-\infty}^{+\infty} G(x) dx = \mu \tag{A.31}$$

$$var[x] = \int_{-\infty}^{+\infty} (\mu - E[x])^2 G(x) dx = \sigma^2, \tag{A.32}$$

so the curve is centered around μ and its width is parametrised by σ. The Full Width at Half Max (FWHM), sometimes also a useful quantity, is positioned at values of x

$$x = \mu \pm \sigma\sqrt{2\ln 2} \simeq \mu \pm 1.17\,\sigma. \tag{A.33}$$

The integrals of the curve of (A.29) between two arbitrary edges give the probability intervals for the Gaussian distribution, and can be obtained in statistic tables. Most useful numbers are the integrals between one σ around the mean (called one standard deviation), which corresponds to an area of 68% of the total, or 0.68 if the Gaussian function is normalised to unit area. The area between two or three standard deviations around the mean corresponds to 95.5 and 99.7%, respectively. Finally, it can be shown that a Poisson distribution tends to a Gaussian shape in the large λ limit, and a Binomial distribution also tends to a Gaussian shape in the large N limit. This is part of a more general behaviour known as the *Central Limit* theorem.

One final PDF that is worth mentioning is the Chi-squared distribution

$$f(\chi^2) = \frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)} \left(\chi^2\right)^{(n/2-1)} e^{-\chi^2/2} \quad \text{(Chi} - \text{squared distribution)}, \tag{A.34}$$

where χ^2 is the continuous event variable, n is a parameter of the distribution (also known as the degrees of freedom) and Γ is a Gamma function. This distribution is approached by the squared sum of variables with Gaussian distributions and has expectation value and variance:

$$E[\chi^2] = n \tag{A.35}$$
$$var[\chi^2] = 2n \tag{A.36}$$

Furthermore, for large values of n the Chi-squared tends to a Gaussian shape.

A.5 Multivariate and Estimators

The probability distribution functions described so far are relative to one quantity, x, in the event space to which it is associated a probability or a probability density. Consider now two or more quantities, x and y each with its own PDF, $P_X(x)$ and $P_Y(y)$, where $P_X(x)$ and $P_Y(y)$ can be different. The event defined by the sum $z = x + y$ will be characterised by a PDF, indicated by $P_Z(z)$ which is not known apriori. However, given a linear combination of N variables x_i independent and characterised by the same, unknown probability distribution function, the sum z

$$z = a_1 x_1 + a_2 x_2 + a_3 x_3 + ..., \tag{A.37}$$

where a_i are real numbers, has the property that for N growing to infinity the probability distribution function of z tends to a Gaussian. This behaviour makes the Gaussian distribution particularly important in experimental physics. It implies that the sum of many independent measurements with the same probability distribution (not necessarily Gaussian) has approximately a Gaussian distribution. If N is not large then z does not have a Gaussian PDF, however the expectation value of the sum of two random variables, x and y, is

$$E[z] = E[x] + E[y], \tag{A.38}$$

and the variance, provided x and y are independent, is

$$var[z] = var[x] + var[y]. \tag{A.39}$$

The (A.38) and (A.39) underly the expressions used in the treatment of experimental data when dealing with repeated measurements of the same quantity x, each measurement being interpreted as a random variable. The expectation value of the mean (A.11) is the same as the expectation value of each measurement of x, while the variance of the mean is related to the variance of each measurement as

$$E[\bar{x}] = \frac{1}{N} E\left[\sum x_i\right] = E[x] \tag{A.40}$$

$$var[\bar{x}] = var\left[\frac{1}{N}\sum x_i\right] = \frac{1}{N} var[x] \tag{A.41}$$

Moreover, in the estimate of the properties of a distribution from a data sampling, the mean $\bar{x}$ is an estimator for the expectation value of the probability density function of x

$$\bar{x} \to E[x] \tag{A.42}$$

and the RMS is an estimator for the variance

$$RMS_B^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2 \to var[x], \tag{A.43}$$

where (A.43) differs from (A.13) only for the $N - 1$ in the denominator (Bessel correction).

Reference

1. A. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Springer, Berlin, 1933). English translation by N. Morrison Foundations of the Theory of Probability (Chelsea, New York) in 1950

Index