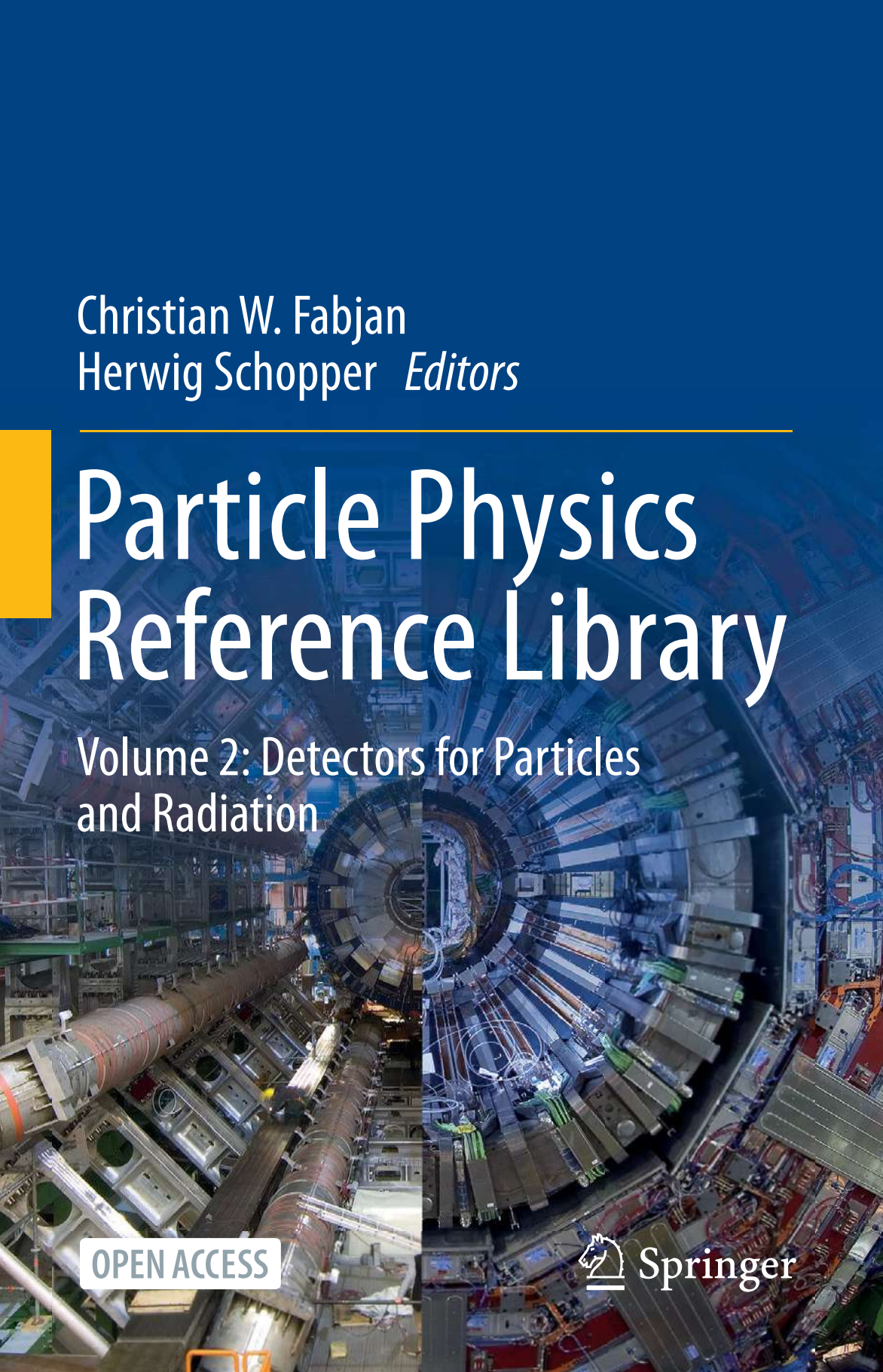
Christian W. Fabjan
Herwig Schopper *Editors*

Particle Physics Reference Library

Volume 2: Detectors for Particles and Radiation



OPEN ACCESS

Springer

Particle Physics Reference Library

Christian W. Fabjan • Herwig Schopper
Editors

Particle Physics Reference Library

Volume 2: Detectors for Particles and Radiation

Springer Open

Editors
Christian W. Fabjan
Austrian Academy of Sciences and
University of Technology
Vienna, Austria

Herwig Schopper
CERN
Geneva, Switzerland

ISBN 978-3-030-35317-9 ISBN 978-3-030-35318-6 (eBook)
https://doi.org/10.1007/978-3-030-35318-6

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

Preface

For many years, the *Landolt-Börnstein—Group I Elementary Particles, Nuclei and Atoms*, Vol. 21A (*Physics and Methods. Theory and Experiments*, 2008), Vol. 21B1 (*Elementary Particles. Detectors for Particles and Radiation*. Part 1: *Principles and Methods*, 2011), Vol. 21B2 (*Elementary Particles. Detectors for Particles and Radiation*. Part 2: *Systems and Applications*), and Vol. 21C (*Elementary Particles. Accelerators and Colliders*, 2013), has served as a major reference work in the field of high-energy physics.

When, not long after the publication of the last volume, open access (OA) became a reality for HEP journals in 2014, discussions between Springer and CERN intensified to find a solution for the "Labö" which would make the content available in the same spirit to readers worldwide. This was helped by the fact that many researchers in the field expressed similar views and their readiness to contribute.

Eventually, in 2016, on the initiative of Springer, CERN and the original Labö volume editors agreed in tackling the issue by proposing to the contributing authors a new OA edition of their work. From these discussions a compromise emerged along the following lines: transfer as much as possible of the original material into open access; add some new material reflecting new developments and important discoveries, such as the Higgs boson; and adapt to the conditions due to the change from copyright to a CC BY 4.0 license.

Some authors were no longer available for making such changes, having either retired or, in some cases, deceased. In most such cases, it was possible to find colleagues willing to take care of the necessary revisions. A few manuscripts could not be updated and are therefore not included in this edition.

We consider that this new edition essentially fulfills the main goal that motivated us in the first place—there are some gaps compared to the original edition, as explained, as there are some entirely new contributions. Many contributions have been only minimally revised in order to make the original status of the field available as historical testimony. Others are in the form of the original contribution being supplemented with a detailed appendix relating recent developments in the field. However, a substantial fraction of contributions has been thoroughly revisited by their authors resulting in true new editions of their original material.

We would like to express our appreciation and gratitude to the contributing authors, to the colleagues at CERN involved in the project, and to the publisher, who has helped making this very special endeavor possible.

Vienna, Austria — Christian W. Fabjan
Geneva, Switzerland — Stephen Myers
Geneva, Switzerland — Herwig Schopper
July 2020

Contents

About the Editors

Christian W. Fabjan is an experimental particle physicist, who spent the major part of his career at CERN, with leading involvement in several of the major CERN programs. At the Intersecting Storage Rings, he concentrated on strong interaction physics and the development of new experimental techniques and followed at the Super Synchrotron with experiments in the Relativistic Heavy Ion program. At the Large Hadron Collider, he focused on the development of several experimental techniques and participated in the ALICE experiment as Technical Coordinator. He is affiliated with the Vienna University of Technology and was, most recently, leading the institute of High Energy Physics of the Austrian Academy of Sciences.

Herwig Schopper joined as a research associate at CERN since 1966 and returned in 1970 as leader of the Nuclear Physics Division and went on to become a member of the directorate responsible for the coordination of CERN's experimental program. He was chairman of the ISR Committee at CERN from 1973 to 1976 and was elected as member of the Scientific Policy Committee in 1979. Following Léon Van Hove and John Adams' years as Director-General for research and executive Director-General, Schopper became the sole Director-General of CERN in 1981.

Schopper's years as CERN's Director-General saw the construction and installation of the Large Electron-Positron Collider (LEP) and the first tests of four detectors for the LEP experiments. Several facilities (including ISR, BEBC, and EHS) had to be closed to free up resources for LEP.

Chapter 1
Introduction

Christian W. Fabjan and Herwig Schopper

Enormous progress has been achieved during the last three decades in the understanding of the microcosm. This was possible by a close interplay between new theoretical ideas and precise experimental data. The present state of our knowledge has been summarised in Volume I/21A "Theory and Experiments". This Volume I/21B is devoted to detection methods and techniques and data acquisition and handling.

The rapid increase of our knowledge of the microcosm was possible only because of an astonishingly fast evolution of detectors for particles and photons. Since the early days of scintillation screens and Geiger counters a series of completely new detector concepts was developed. They are based on imaginative ideas, sometimes even earning a Nobel Prize, combined with sophisticated technological developments. It might seem surprising that the exploration of an utterly abstract domain like particle physics, requires the most advanced techniques, but this makes the whole field so attractive.

The development of detectors was above all pushed by the requirements of particle physics. In order to explore smaller structures one has to use finer probes, i.e. shorter wavelengths implying higher particle energies. This requires detectors for high-energy particles and photons. At the same time one has to cope with the quantum-mechanical principle that cross sections for particle interactions have a tendency to fall with increasing interaction energy. Therefore accelerators or colliders have to deliver not only higher energies but at the same time also higher collision rates. This implies that detectors must sustain higher rates. This problem is aggravated by the fact that the high-energy frontier is at present linked to hadron

C. W. Fabjan (✉)
Austrian Academy of Sciences and University of Technology, Vienna, Austria
e-mail: Chris.Fabjan@cern.ch

H. Schopper
CERN, Geneva, Switzerland

C. W. Fabjan, H. Schopper (eds.), *Particle Physics Reference Library*,
https://doi.org/10.1007/978-3-030-35318-6_1

collisions. Electron-positron colliders are characterised by events with relatively few outgoing particles since two pointlike particles collide and the strong interaction is negligible in such reactions. After the shutdown of LEP in 2000 the next electron-positron collider is far in the future and progress is now depending on proton-proton collisions at the LHC at CERN or heavy ion colliders, e.g. GSI, Germany, RHIC at BNL in the USA and also LHC. Protons are composite particles containing quarks and gluons and hence proton collisions produce very complicated events with many hundreds of particles. Consequently, detectors had to be developed which are able to cope with extremely high data rates and have to resist high levels of irradiation. Such developments were in particular motivated by the needs of the LHC experiments.

It seems plausible that accelerators and colliders have to grow in size with increasing energy. But why have detectors to be so large? Their task is to determine the direction of emitted particles, measure their momenta or energy and in some cases their velocity which together with the momentum allows to determine their mass and hence to identify the nature of the particle.

The most precise method to measure the momentum of charged particles is to determine their deflection in a magnetic field which is proportional to $B \cdot l$ where B is the magnetic field strength and l the length of the trajectory in the magnetic field. Of course, it is also determined by the spatial resolution of the detector to determine the track. To attain the highest possible precision superconducting coils are used in most experiments to produce a large B. Great efforts have been made to construct detectors with a spatial resolution down to the order of several microns. But even then track lengths l of the order of several meters are needed to measure momenta with a precision of about 1% of particles with momenta of several 100 GeV/c. This is the main reason why experiments must have extensions of several meters and weigh thousands of tons.

Another possibility to determine the energy of particles are so-called "calorimeters". This name is misleading since calorimeters have nothing to do with calorific measurements but this name became ubiquitous to indicate that the total energy of a particle is measured. The measurement is done in the following way. A particle hits the material of the detector, interacts with an atom, produces secondary particles which, if sufficiently energetic, generate further particles, leading to a whole cascade of particles of ever decreasing energies. The energy deposited in the detector material can be measured in various ways. If the material of the detector is a scintillator (crystal, liquid or gas), the scintillating light is approximately proportional to the deposited energy and it can be observed by, e.g., photomultipliers. Alternatively, the ionisation produced by the particle cascade can be measured by electrical means.

In principle two kinds of calorimeters can be distinguished. Electrons and photons produce a so-called electromagnetic cascade due to electromagnetic interactions. Such cascades are relatively small both in length and in lateral dimension. Hence electromagnetic calorimeters can consist of a homogenous detector material containing the whole cascade. Incident hadrons, however, produce in the cascade also a large number of neutrons which can travel relatively long ways before losing their energy and therefore hadronic cascades have large geometrical extensions even

in the densest materials (of the order few meters in iron). Therefore the detectors for hadronic cascades are composed of a sandwich of absorber material interspersed with elements to detect the deposited energy. In such a device, only a certain fraction of the total energy is sampled. The challenge of the design consists in making this fraction as much as possible proportional to the total energy. The main advantage of calorimeters, apart from the sensitivity to both charged and neutral particles, is that their size increases only logarithmically with the energy of the incident particle, hence much less than for magnetic spectrometers, albeit with an energy resolution inferior to magnetic spectrometers below about 100 GeV. They require therefore comparatively little space which is of paramount importance for colliders where the solid angle around the interaction area has to be covered in most cases as fully as possible.

Other detectors have been developed for particular applications, e.g. for muon and neutrino detection or the observation of cosmic rays in the atmosphere or deep underground/water. Experiments in space pose completely new problems related to mechanical stability and restrictions on power consumption and consumables.

The main aim in the development of all these detectors is higher sensitivity, better precision and less influence by the environment. Obviously, reduction of cost has become a major issue in view of the millions of detector channels in most modern experiments.

New and more sophisticated detectors need better signal processing, data acquisition and networking. Experiments at large accelerators and colliders pose special problems dictated by the beam properties and restricted space. Imagination is the key to overcome such challenges.

Experiments at accelerators/colliders and for the observation of cosmic rays have become big projects involving hundreds or even thousands of scientists and the time from the initial proposal to data taking may cover one to two decades. Hence it is sometimes argued that they are not well adapted for the training of students. However, the development of a new detector is subdivided in a large number of smaller tasks (concept of the detector, building prototypes, testing, computer simulations and preparation of the data acquisition), each lasting only a few years and therefore rather well suited for a master or PhD thesis. The final "mass production" of many detection channels in the full detector assembly, however, is eventually transferred to industry. These kinds of activities may in some cases have little to do with particle physics itself, but they provide an excellent basis for later employment in industry. Apart from specific knowledge, e.g., in vacuum, magnets, gas discharges, electronics, computing and networking, students learn how to work in the environment of a large project respecting time schedules and budgetary restrictions—and perhaps even most important to be trained to work in an international environment.

Because the development of detectors does not require the resources of a large project but can be carried out in a small laboratory, most of these developments are done at universities. Indeed most of the progress in detector development is due to universities or national laboratories. However, when it comes to plan a large experiment these originally individual activities are combined and coordinated

which naturally leads to international cooperation between scientists from different countries, political traditions, creeds and mentalities. To learn how to adapt to such an international environment represents a human value which goes much beyond the scientific achievements.

The stunning success of the "Standard Model of particle physics" also exhibits with remarkable clarity its limitations. The many open fundamental issues—origin of CP-violation, neutrino mass, dark matter and dark energy, to name just few—are motivating a vast, multi-faceted research programme for accelerator- and non-accelerator based, earth- and space-based experimentation. This has led to a vigorous R&D in detectors and data handling.

This revised edition provides an update on these developments over the past 7–9 years.

We gratefully acknowledge the very constructive collaboration with the authors of the first edition, in several cases assisted by additional authors. May this Open Access publication reach a global readership, for the benefit of science.

Chapter 2
The Interaction of Radiation with Matter

Hans Bichsel and Heinrich Schindler

2.1 Introduction

The detection of charged particles is usually based on their electromagnetic interactions with the electrons and nuclei of a detector medium. Interaction with the Coulomb field of the nucleus leads to deflections of the particle trajectory (multiple scattering) and to radiative energy loss (bremsstrahlung). Since the latter, discussed in Sect. 2.4.1, is inversely proportional to the particle mass squared, it is most significant for electrons and positrons.

"Heavy" charged particles (in this context: particles with a mass M exceeding the electron mass m) passing through matter lose energy predominantly through collisions with electrons. Our theoretical understanding of this process, which has been summarised in a number of review articles [1–7] and textbooks [8–13], is based on the works of some of the most prominent physicists of the twentieth century, including Bohr [14, 15], Bethe [16, 17], Fermi [18, 19], and Landau [20].

After outlining the quantum-mechanical description of single collisions in terms of the double-differential cross section $\mathrm{d}^2\sigma/(\mathrm{d}E\mathrm{d}q)$, where E and q are the energy transfer and momentum transfer involved in the collision, Sect. 2.3 discusses algorithms for the quantitative evaluation of the single-differential cross section

The author Hans Bichsel is deceased at the time of publication.

H. Bichsel · H. Schindler (✉)
CERN, Geneva, Switzerland
e-mail: Heinrich.Schindler@cern.ch

C. W. Fabjan, H. Schopper (eds.), *Particle Physics Reference Library*,
https://doi.org/10.1007/978-3-030-35318-6_2

$\mathrm{d}\sigma/\mathrm{d}E$ and its moments. The integral cross section (zeroth moment), multiplied by the atomic density N, corresponds to the charged particle's inverse mean free path λ^{-1} or, in other words, the average number of collisions per unit track length,

$$\lambda^{-1} = M_0 = N \int_{E_{\min}}^{E_{\max}} \frac{\mathrm{d}\sigma}{\mathrm{d}E} \mathrm{d}E. \tag{2.1}$$

The stopping power $\mathrm{d}E/\mathrm{d}x$, i.e. the average energy loss per unit track length, is given by the first moment,

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = M_1 = N \int_{E_{\min}}^{E_{\max}} E \frac{\mathrm{d}\sigma}{\mathrm{d}E} \mathrm{d}E. \tag{2.2}$$

The integration limits $E_{\min,\,\max}$ are determined by kinematics. Due to the stochastic nature of the interaction process, the number of collisions and the sum Δ of energy losses along a particle track are subject to fluctuations. Section 2.5 deals with methods for calculating the probability density distribution $f(\Delta, x)$ for different track lengths x. The energy transfer from the incident particle to the electrons of the medium typically results in excitation and ionisation of the target atoms. These observable effects are discussed in Sect. 2.6.

As a prologue to the discussion of charged-particle collisions, Sect. 2.2 briefly reviews the principal photon interaction mechanisms in the X-ray and gamma ray energy range.

Throughout this chapter, we attempt to write all expressions in a way independent of the system of units (cgs or SI), by using the fine structure constant $\alpha \sim 1/137$. Other physical constants used occasionally in this chapter include the Rydberg energy $\mathrm{Ry} = \alpha^2 mc^2/2 \sim 13.6\,\mathrm{eV}$, and the Bohr radius $a_0 = \hbar c/\left(\alpha mc^2\right) \sim 0.529\,\text{Å}$. Cross-sections are quoted in barn ($1\,\mathrm{b} = 10^{-24}\,\mathrm{cm}^2$).

2.2 Photon Interactions

Photons interact with matter via a range of mechanisms, which can be classified according to the type of target, and the effect of the interaction on the photon (absorption or scattering) [9, 21]. At energies beyond the ultraviolet range, the dominant processes are photoelectric absorption (Sect. 2.2.1), Compton scattering (Sect. 2.2.2), and pair production (Sect. 2.2.3). As illustrated in Fig. 2.1, photoabsorption constitutes the largest contribution to the total cross section at low photon energies, pair production is the most frequent interaction at high energies, and Compton scattering dominates in the intermediate energy range.

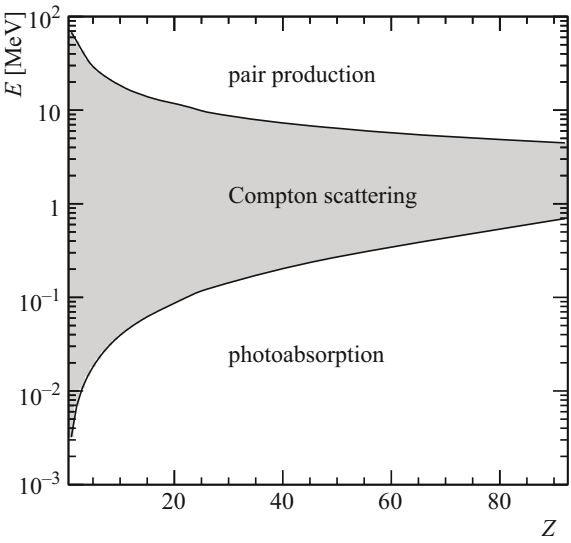


Fig. 2.1 The lower curve shows, as a function of the atomic number Z of the target material, the photon energy E below which photoelectric absorption is the most probable interaction mechanism, while the upper curve shows the energy above which pair production is the most important process. The shaded region between the two curves corresponds to the domain where Compton scattering dominates. The cross sections are taken from the NIST XCOM database [24]

Detailed descriptions of these processes can be found, for instance, in Refs. [8–10, 12, 22, 23]. The focus of this section is on photoabsorption, the description of which (as will be discussed in Sect. 2.3) is related to that of inelastic charged particle collisions in the regime of low momentum transfer.

2.2.1 Photoabsorption

In a photoelectric absorption interaction, the incident photon disappears and its energy is transferred to the target atom (or group of atoms). The intensity I of a monochromatic beam of photons with energy E thus decreases exponentially as a function of the penetration depth x in a material,

$$I(x) = I_0 \mathrm{e}^{-\mu x},$$

where the attenuation coefficient μ is proportional to the atomic density N of the medium and the photoabsorption cross section σ_γ,

$$\mu(E) = N\sigma_\gamma(E).$$

Let us first consider a (dipole-allowed) transition between the ground state $|0\rangle$ of an atom and a discrete excited state $|n\rangle$ with excitation energy E_n. The integral photoabsorption cross section of the line is given by

$$\int \sigma_\gamma^{(n)}(E)\,\mathrm{d}E = \frac{2\pi^2\alpha(\hbar c)^2}{mc^2} f_n.$$

The dimensionless quantity

$$f_n = \frac{2mc^2}{3\,(\hbar c)^2} E_n |\langle n| \sum_{j=1}^{Z} \mathbf{r}_i |0\rangle|^2, \tag{2.3}$$

with the sum extending over the electrons in the target atom, is known as the dipole oscillator strength (DOS). Similarly, transitions to the continuum are characterised by the dipole oscillator strength density $\mathrm{d}f/\mathrm{d}E$, and the photoionisation cross section $\sigma_\gamma\,(E)$ is given by

$$\sigma_\gamma\,(E) = \frac{2\pi^2\alpha\,(\hbar c)^2}{mc^2} \frac{\mathrm{d}f\,(E)}{\mathrm{d}E}. \tag{2.4}$$

The dipole oscillator strength satisfies the Thomas-Reiche-Kuhn (TRK) sum rule,

$$\sum_n f_n + \int \mathrm{d}E\, \frac{\mathrm{d}f\,(E)}{\mathrm{d}E} = Z. \tag{2.5}$$

For most gases, the contribution of excited states ($\sum f_n$) to the TRK sum rule is a few percent of the total, e.g. ~5% for argon and ~7% for methane [25, 26].

As can be seen from Fig. 2.2, the photoabsorption cross section reflects the atomic shell structure. Evaluated atomic and molecular photoabsorption cross

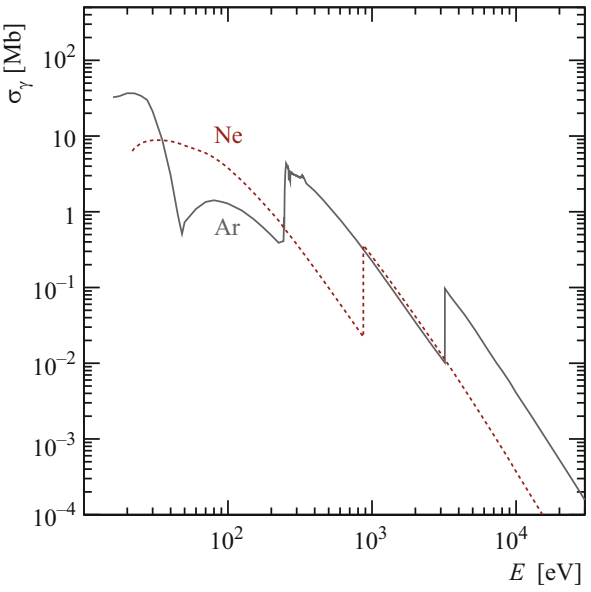


Fig. 2.2 Photoabsorption cross sections of argon (solid curve) and neon (dashed curve) as a function of the photon energy E [25, 26]

sections (both for discrete excitations as well as transitions to the continuum) for many commonly used gases are given in the book by Berkowitz [25, 26].

At energies sufficiently above the ionisation threshold, the molecular photoabsorption cross section is, to a good approximation, given by the sum of the photoabsorption cross sections of the constituent atoms. A comprehensive compilation of atomic photoabsorption data (in the energy range between ∼30 eV and 30 keV) can be found in Ref. [27]. Calculations for energies between 1 and 100 GeV are available in the NIST XCOM database [24]. Calculated photoionisation cross sections for individual shells can be found in Refs. [28–30]. At high energies, i.e. above the respective absorption edges, photons interact preferentially with inner-shell electrons. The subsequent relaxation processes (emission of fluorescence photons and Auger electrons) are discussed in Sect. 2.6.

The response of a solid with atomic number Z to an incident photon of energy $E = \hbar\omega$ is customarily described in terms of the complex dielectric function $\varepsilon(\omega) = \varepsilon_1(\omega) + \mathrm{i}\varepsilon_2(\omega)$. The oscillator strength density is related to $\varepsilon(\omega)$ by

$$\frac{\mathrm{d}f\,(E)}{\mathrm{d}E} = E\frac{2Z}{\pi\left(\hbar\Omega_p\right)^2}\frac{\varepsilon_2\,(E)}{\varepsilon_1^2\,(E)+\varepsilon_2^2\,(E)} = E\frac{2Z}{\pi\left(\hbar\Omega_p\right)^2}\mathrm{Im}\left(\frac{-1}{\varepsilon\,(E)}\right), \tag{2.6}$$

where

$$\hbar\Omega_p = \sqrt{\frac{4\pi\alpha\,(\hbar c)^3\,NZ}{mc^2}} \tag{2.7}$$

is the plasma energy of the material, which depends only on the electron density NZ. In terms of the dielectric loss function $\mathrm{Im}\,(-1/\varepsilon)$, the TRK sum rule reads

$$\int \mathrm{d}E\ \mathrm{Im}\left(\frac{-1}{\varepsilon\,(E)}\right)E = \frac{\pi}{2}\left(\hbar\Omega_p\right)^2. \tag{2.8}$$

Compilations of evaluated optical data for semiconductors are available in Ref. [32], and for solids in general in Ref. [31]. As an example, Fig. 2.3 shows the dielectric loss function of silicon, a prominent feature of which is the peak at ∼17 eV, corresponding to the plasma energy of the four valence (M-shell) electrons.

2.2.2 *Compton Scattering*

Compton scattering refers to the collision of a photon with a weakly bound electron, whereby the photon transfers part of its energy to the electron and is deflected with respect to its original direction of propagation. We assume in the following that the target electron is free and initially at rest, which is a good approximation if the photon energy E is large compared to the electron's binding energy. Due to

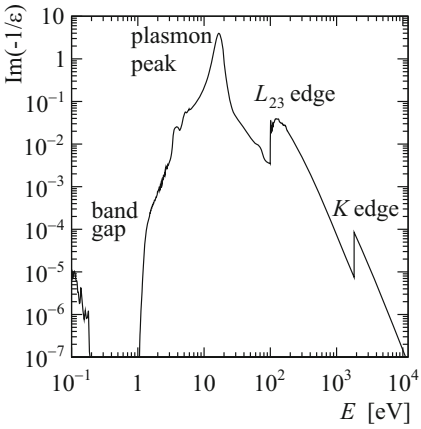


Fig. 2.3 Dielectric loss function $\mathrm{Im}\,(-1/\varepsilon\,(E))$ of solid silicon [31] as a function of the photon energy E

conservation of energy and momentum, the photon energy E' after the collision and the scattering angle θ of the photon are then related by

$$E' = \frac{mc^2}{1 - \cos\theta + (1/u)}, \tag{2.9}$$

where $u = E/\left(mc^2\right)$ is the photon energy (before the collision) in units of the electron rest energy.

The kinetic energy $T = E - E'$ imparted to the electron is largest for a head-on collision ($\theta = \pi$) and the energy spectrum of the recoil electrons consequently exhibits a cut-off (Compton edge) at

$$T_{\max} = E\frac{2u}{1+2u}.$$

The total cross section (per electron) for the Compton scattering of an unpolarised photon by a free electron at rest, derived by Klein and Nishina in 1929 [33], is given by

$$\sigma^{(\mathrm{KN})} = 2\pi\left(\frac{\alpha\hbar c}{mc^2}\right)^2\left(\frac{1+u}{u^2}\left[\frac{2\,(1+u)}{1+2u} - \frac{\ln\,(1+2u)}{u}\right] + \frac{\ln\,(1+2u)}{2u} - \frac{1+3u}{(1+2u)^2}\right). \tag{2.10}$$

At low energies ($u \ll 1$), the Klein-Nishina formula (2.10) is conveniently approximated by the expansion [34]

$$\sigma^{(\mathrm{KN})} = \underbrace{\frac{8\pi}{3}\left(\frac{\alpha\hbar c}{mc^2}\right)^2}_{\text{Thomson cross section}}\ \frac{1}{(1+2u)^2}\left(1 + 2u + \frac{6}{5}u^2 + \ldots\right),$$

while at high energies ($u \gg 1$) the approximation [8, 10, 22]

$$\sigma^{(\mathrm{KN})} \sim \pi \left(\frac{\alpha\hbar c}{mc^2}\right)^2 \frac{1}{u}\left(\ln(2u) + \frac{1}{2}\right)$$

can be used.

The angular distribution of the scattered photon is given by the differential cross section

$$\frac{\mathrm{d}\sigma^{(\mathrm{KN})}}{\mathrm{d}(\cos\theta)} = \pi \left(\frac{\alpha\hbar c}{mc^2}\right)^2 \left[\frac{1}{1+u(1-\cos\theta)}\right]^2 \left(\frac{1+\cos^2\theta}{2}\right) \times \left(1 + \frac{u^2(1-\cos\theta)^2}{(1+\cos^2\theta)[1+u(1-\cos\theta)]}\right),$$

which corresponds to a kinetic energy spectrum [22]

$$\frac{\mathrm{d}\sigma^{(\mathrm{KN})}}{\mathrm{d}T} = \pi \left(\frac{\alpha\hbar c}{mc^2}\right)^2 \frac{1}{u^2 mc^2}\left(2 + \left(\frac{T}{E-T}\right)^2 \left[\frac{1}{u^2} + \frac{E-T}{E} - \frac{2(E-T)}{uT}\right]\right)$$

of the target electron.

The cross section for Compton scattering off an atom scales roughly with the number of electrons in the atom and, assuming that the photon energy is large compared to the atomic binding energies, may be approximated by

$$\sigma^{(\mathrm{Compton})} \sim Z\sigma^{(\mathrm{KN})}.$$

Methods for including the effects of the binding energy and the internal motion of the orbital electrons in calculations of atomic Compton scattering cross sections are discussed, for instance, in Ref. [35].

2.2.3 *Pair Production*

For photon energies exceeding $2mc^2$, an interaction mechanism becomes possible where the incoming photon disappears and an electron-positron pair, with a total energy equal to the photon energy E, is created. Momentum conservation requires this process, which is closely related to bremsstrahlung (Sect. 2.4.1), to take place in the electric field of a nucleus or of the atomic electrons. In the latter case, kinematic constraints impose a threshold of $E > 4mc^2$.

At high photon energies, the electron-positron pair is emitted preferentially in the forward direction and the absorption coefficient due to pair production can be approximated by

$$\mu = N\sigma^{(\text{pair production})} = \frac{7}{9}\frac{1}{X_0},$$

where X_0 is a material-dependent parameter known as the radiation length (see Sect. 2.4.1). More accurate expressions are given in Ref. [8]. Tabulations of calculated pair-production cross sections can be found in Ref. [36] and are available online [24].

2.3 Interaction of Heavy Charged Particles with Matter

The main ingredient for computing the energy loss of an incident charged particle due to interactions with the electrons of the target medium is the single-differential cross section with respect to the energy transfer E in a collision. In this section, we discuss the calculation of $\mathrm{d}\sigma/\mathrm{d}E$ and its moments for "fast", point-like particles. To be precise, we consider particles with a velocity that is large compared to the velocities of the atomic electrons, corresponding to the domain of validity of the first-order Born approximation.

In the limit where the energy transfer E is large compared to the atomic binding energies, $\mathrm{d}\sigma/\mathrm{d}E$ approaches the cross section for scattering off a free electron. For a spin-zero particle with charge ze and speed βc, the asymptotic cross section (per electron) towards large energy transfers is given by [8]

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E} = \underbrace{\frac{2\pi z^2 (\alpha\hbar c)^2}{mc^2\beta^2}\frac{1}{E^2}}_{\text{Rutherford cross section}}\left(1-\beta^2\frac{E}{E_{\max}}\right) = \frac{\mathrm{d}\sigma_R}{\mathrm{d}E}\left(1-\beta^2\frac{E}{E_{\max}}\right). \tag{2.11}$$

Similar expressions have been derived for particles with spin 1 and spin 1/2 [8]. The maximum energy transfer is given by the kinematics of a head-on collision between a particle with mass M and an electron (mass m) at rest,

$$E_{\max} = 2mc^2\beta^2\gamma^2\left(1+2\gamma\frac{m}{M}+\left(\frac{m}{M}\right)^2\right)^{-1}, \tag{2.12}$$

which for $M \gg m$ becomes $E_{\max} \sim 2mc^2\beta^2\gamma^2$.

These so-called "close" or "knock-on" collisions, in which the projectile interacts with a single atomic electron, contribute a significant fraction (roughly half) to the average energy loss of a charged particle in matter but are rare compared to "distant" collisions in which the particle interacts with the atom as a whole or with a group of

atoms. For an accurate calculation of $\mathrm{d}\sigma/\mathrm{d}E$, the electronic structure of the target medium therefore needs to be taken into account.

In the non-relativistic first-order Born approximation, the transition of an atom from its ground state to an excited state $|n\rangle$ involving a momentum transfer $\mathbf{q}$ is characterised by the matrix element (inelastic form factor)

$$F_{n0}(\mathbf{q}) = \langle n| \sum_{j=1}^{Z} \exp\left(\frac{\mathrm{i}}{\hbar}\mathbf{q}\cdot\mathbf{r}_j\right)|0\rangle,$$

which is independent of the projectile. The differential cross section with respect to the recoil parameter $Q = q^2/(2m)$, derived by Bethe in 1930 [16], is given by [1–3, 16]

$$\frac{\mathrm{d}\sigma_n}{\mathrm{d}Q} = \frac{2\pi z^2 (\alpha\hbar c)^2}{mc^2\beta^2}\frac{1}{Q^2}|F_{n0}(\mathbf{q})|^2 = \frac{2\pi z^2 (\alpha\hbar c)^2}{mc^2\beta^2}\frac{f_n(q)}{QE_n},$$

where $f_n(q)$ denotes the generalised oscillator strength (GOS). In the limit $q \to 0$ it becomes the dipole oscillator strength f_n discussed in Sect. 2.2.1. The double-differential cross section for transitions to the continuum (i.e. ionisation) is given by

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}E\mathrm{d}Q} = \frac{2\pi z^2 (\alpha\hbar c)^2}{mc^2\beta^2}\frac{1}{QE}\frac{\mathrm{d}f(E,q)}{\mathrm{d}E}, \tag{2.13}$$

where $\mathrm{d}f(E,q)/\mathrm{d}E$ is the generalised oscillator strength density. The GOS is constrained by the Bethe sum rule [2, 16] (a generalisation of the TRK sum rule),

$$\sum_n f_n(q) + \int \mathrm{d}E \frac{\mathrm{d}f(E,q)}{\mathrm{d}E} = Z, \quad \forall q. \tag{2.14}$$

Closed-form expressions for the generalised oscillator strength (density) exist only for very simple systems such as the hydrogen atom (Fig. 2.4). Numerical calculations are available for a number of atoms and molecules (see e.g. Ref. [37]). A prominent feature of the generalised oscillator strength density is the so-called "Bethe ridge": at high momentum transfers $\mathrm{d}f(E,q)/\mathrm{d}E$ is concentrated along the free-electron dispersion relation $Q = E$.

In order to calculate $\mathrm{d}\sigma/\mathrm{d}E$, we need to integrate the double-differential cross-section over Q,

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E} = \int_{Q_{\min}}^{Q_{\max}} \mathrm{d}Q \frac{\mathrm{d}^2\sigma}{\mathrm{d}E\mathrm{d}Q}, \qquad Q_{\min} \sim \frac{E^2}{2m\beta^2c^2}. \tag{2.15}$$

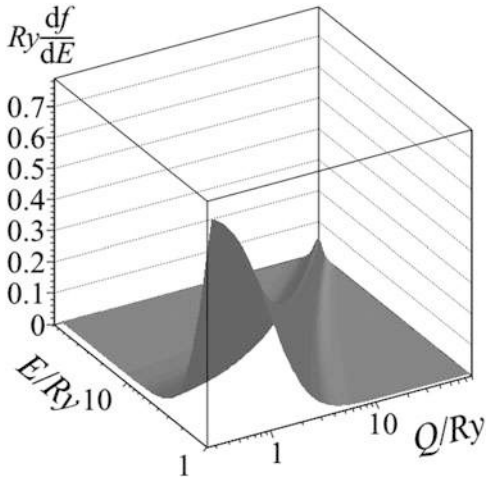


Fig. 2.4 Generalised oscillator strength density $\mathrm{d}f\,(E,q)\,/\mathrm{d}E$ of atomic hydrogen [2, 3, 16], for transitions to the continuum

For this purpose, it is often sufficient to use simplified models of the generalised oscillator strength density, based on the guidelines provided by model systems like the hydrogen atom, and using (measured) optical data in the low-Q regime.

Equation (2.13) describes the interaction of a charged particle with an isolated atom, which is a suitable approximation for a dilute gas. In order to extend it to dense media and to incorporate relativistic effects, it is convenient to use a semi-classical formalism [19, 38]. In this approach, which can be shown to be equivalent to the first-order quantum mechanical result, the response of the medium to the incident particle is described in terms of the complex dielectric function.

2.3.1 Dielectric Theory

Revisiting the energy loss of charged particles in matter from the viewpoint of classical electrodynamics, we calculate the electric field of a point charge ze moving with a constant velocity $\boldsymbol{\beta}c$ through an infinite, homogeneous and isotropic medium, that is we solve Maxwell's equations

$$\nabla\cdot\mathbf{B}=0\,,\qquad \nabla\times\mathbf{E}=-\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t},$$
$$\nabla\times\mathbf{B}=\frac{1}{c}\frac{\partial\mathbf{D}}{\partial t}+\frac{4\pi}{c}\mathbf{j}\,,\qquad \nabla\cdot\mathbf{D}=4\pi\rho,$$

for source terms

$$\rho=ze\delta^3\left(\mathbf{r}-\boldsymbol{\beta}ct\right),\qquad \mathbf{j}=\boldsymbol{\beta}c\rho.$$

The perturbation due to the moving charge is assumed to be weak enough such that there is a linear relationship between the Fourier components of the electric field $\mathbf{E}$ and the displacement field $\mathbf{D}$,

$$\mathbf{D}(\mathbf{k},\omega) = \varepsilon(\mathbf{k},\omega)\,\mathbf{E}(\mathbf{k},\omega),$$

where $\varepsilon(\mathbf{k},\omega) = \varepsilon_1(\mathbf{k},\omega) + \mathrm{i}\varepsilon_2(\mathbf{k},\omega)$ is the (generalized) complex dielectric function.

The particle experiences a force $ze\mathbf{E}(\mathbf{r} = \boldsymbol{\beta}ct, t)$ that slows it down, and the stopping power is given by the component of this force parallel to the particle's direction of motion,

$$\frac{\mathrm{d}E}{\mathrm{d}x} = ze\mathbf{E}\cdot\frac{\boldsymbol{\beta}}{\beta}.$$

Adopting the Coulomb gauge $\mathbf{k}\cdot\mathbf{A} = 0$, one obtains after integrating over the angles (assuming that the dielectric function ε is isotropic),

$$\frac{\mathrm{d}E}{\mathrm{d}x} = -\frac{2z^2e^2}{\beta^2\pi}\int \mathrm{d}\omega \int \mathrm{d}k \times\left[\frac{\omega}{kc^2}\mathrm{Im}\left(\frac{-1}{\varepsilon(k,\omega)}\right) + \omega k\left(\beta^2 - \frac{\omega^2}{k^2c^2}\right)\mathrm{Im}\left(\frac{1}{-k^2c^2 + \varepsilon(k,\omega)\,\omega^2}\right)\right]. \tag{2.16}$$

The first term in the integrand represents the non-relativistic contribution to the energy loss which we would have obtained by considering only the scalar potential ϕ. It is often referred to as the longitudinal term. The second term, known as the transverse term, originates from the vector potential $\mathbf{A}$ and incorporates relativistic effects.

On a microscopic level, the energy transfer from the particle to the target medium proceeds through discrete collisions with energy transfer $E = \hbar\omega$ and momentum transfer $q = \hbar k$. Comparing Eq. (2.2) with the macroscopic result (2.16), one obtains

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}E\mathrm{d}q} = \frac{2z^2\alpha}{\beta^2\pi\hbar cN} \times\left[\frac{1}{q}\mathrm{Im}\left(\frac{-1}{\varepsilon(q,E)}\right) + \frac{1}{q}\left(\beta^2 - \frac{E^2}{q^2c^2}\right)\mathrm{Im}\left(\frac{1}{-1 + \varepsilon(q,E)\,E^2/\left(q^2c^2\right)}\right)\right]. \tag{2.17}$$

The loss function $\mathrm{Im}\left(-1/\varepsilon\left(q,E\right)\right)$ and the generalized oscillator strength density are related by

$$\frac{\mathrm{d}f\left(E,q\right)}{\mathrm{d}E}=E\frac{2Z}{\pi\left(\hbar\Omega_p\right)^2}\mathrm{Im}\left(\frac{-1}{\varepsilon\left(q,E\right)}\right). \tag{2.18}$$

Using this identity, we see that the longitudinal term (first term) in Eq. (2.17) is equivalent to the non-relativistic quantum mechanical result (2.13). As is the case with the generalized oscillator strength density, closed-form expressions for the dielectric loss function $\mathrm{Im}\left(-1/\varepsilon\left(q,E\right)\right)$ can only be derived for simple systems like the ideal Fermi gas [39, 40]. In the following (Sects. 2.3.2 and 2.3.3), we discuss two specific models of $\mathrm{Im}\left(-1/\varepsilon\left(q,E\right)\right)$ (or, equivalently, $\mathrm{d}f\left(E,q\right)/\mathrm{d}E$).

2.3.2 Bethe-Fano Method

The relativistic version of Eq. (2.13) or, in other words, the equivalent of Eq. (2.17) in oscillator strength parlance, is [1, 41]

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}E\mathrm{d}Q}=\frac{2\pi z^2\left(\alpha\hbar c\right)^2}{mc^2\beta^2}Z\left[\frac{\left|F\left(E,\mathbf{q}\right)\right|^2}{Q^2\left(1+\frac{Q}{2mc^2}\right)^2}+\frac{\left|\boldsymbol{\beta}_t\cdot\mathbf{G}\left(E,\mathbf{q}\right)\right|^2}{\left[Q\left(1+\frac{Q}{2mc^2}\right)-\frac{E^2}{2mc^2}\right]^2}\right]\left(1+\frac{Q}{mc^2}\right) \tag{2.19}$$

where $Q\left(1+Q/2mc^2\right)=q^2/2m$, $\boldsymbol{\beta}_t$ is the component of the velocity perpendicular to the momentum transfer $\mathbf{q}$, and $F\left(E,\mathbf{q}\right)$ and $\mathbf{G}\left(E,\mathbf{q}\right)$ represent the matrix elements for longitudinal and transverse excitations.

Depending on the type of target and the range of momentum transfers involved, we can use Eqs. (2.13), (2.19) or (2.17) as a starting point for evaluating the single-differential cross section. Following the approach described by Fano [1], we split $\mathrm{d}\sigma/\mathrm{d}E$ in four parts. For small momentum transfers ($Q<Q_1\sim 1\,\mathrm{Ry}$), we can use the non-relativistic expression (2.13) for the longitudinal term and approximate the generalised oscillator strength density by its dipole limit,

$$\frac{\mathrm{d}\sigma^{(1)}}{\mathrm{d}E}=\frac{2\pi z^2\left(\alpha\hbar c\right)^2}{mc^2\beta^2}\frac{1}{E}\frac{\mathrm{d}f\left(E\right)}{\mathrm{d}E}\int_{Q_{\min}}^{Q_1}\frac{\mathrm{d}Q}{Q}=\frac{2\pi z^2\left(\alpha\hbar c\right)^2}{mc^2\beta^2}\frac{1}{E}\frac{\mathrm{d}f\left(E\right)}{\mathrm{d}E}\ln\frac{Q_1 2mc^2\beta^2}{E^2}. \tag{2.20}$$

In terms of the dielectric loss function, one obtains

$$\frac{\mathrm{d}\sigma^{(1)}}{\mathrm{d}E}=\frac{z^2\alpha}{\beta^2\pi\hbar cN}\mathrm{Im}\left(\frac{-1}{\varepsilon\left(E\right)}\right)\ln\frac{Q_1 2mc^2\beta^2}{E^2}.$$

For high momentum transfers ($Q > Q_2 \sim 30\,\mathrm{keV}$), i.e. for close collisions where the binding energy of the atomic electrons can be neglected, the longitudinal and transverse matrix elements are strongly peaked at the Bethe ridge $Q = E$. Using [1]

$$|F(E,\mathbf{q})|^2 \sim \frac{1 + Q/\left(2mc^2\right)}{1 + Q/\left(mc^2\right)}\delta(E-Q),$$

$$|\boldsymbol{\beta}_t \cdot \mathbf{G}(E,\mathbf{q})|^2 \sim \beta_t^2 \frac{1 + Q/\left(2mc^2\right)}{1 + Q/\left(mc^2\right)}\delta(E-Q)$$

and

$$\beta_t^2 = \frac{1}{1 + Q/\left(2mc^2\right)} - \left(1-\beta^2\right)$$

one obtains (for longitudinal and transverse excitations combined),

$$\frac{\mathrm{d}\sigma^{(h)}}{\mathrm{d}E} = \frac{2\pi z^2 (\alpha\hbar c)^2}{mc^2\beta^2}\frac{Z}{E}\left(1 - \frac{E\left(1-\beta^2\right)}{2mc^2}\right). \tag{2.21}$$

In the intermediate range, $Q_1 < Q < Q_2$, numerical calculations of the generalised oscillator strength density are used. An example of $\mathrm{d}f(E,q)/\mathrm{d}E$ is shown in Fig. 2.5. Since the limits Q_1, Q_2 do not depend on the particle velocity, the integrals

$$\frac{\mathrm{d}\sigma^{(2)}}{\mathrm{d}E} = \frac{2\pi z^2 (\alpha\hbar c)^2}{mc^2\beta^2}\frac{1}{E}\int_{Q_1}^{Q_2}\frac{\mathrm{d}Q}{Q}\frac{\mathrm{d}f(E,q)}{\mathrm{d}E}$$

need to be evaluated only once for each value of E. The transverse contribution can be neglected[1] [41].

The last contribution, described in detail in Ref. [1], is due to low-Q transverse excitations in condensed matter. Setting $\mathrm{Im}\left(-1/\varepsilon(E,q)\right) = \mathrm{Im}\left(-1/\varepsilon(E)\right)$ in the second term in Eq. (2.17) and integrating over q gives

$$\frac{\mathrm{d}\sigma^{(3)}}{\mathrm{d}E} = \frac{z^2\alpha}{\beta^2\pi N\hbar c} \times\left[\mathrm{Im}\left(\frac{-1}{\varepsilon(E)}\right)\ln\frac{1}{\left|1-\beta^2\varepsilon(E)\right|} + \left(\beta^2 - \frac{\varepsilon_1(E)}{|\varepsilon(E)|^2}\right)\left(\frac{\pi}{2} - \arctan\frac{1-\beta^2\varepsilon_1(E)}{\beta^2\varepsilon_2(E)}\right)\right]. \tag{2.22}$$

[1] For particle speeds $\beta < 0.1$, this approximation will cause errors, especially for M_0.

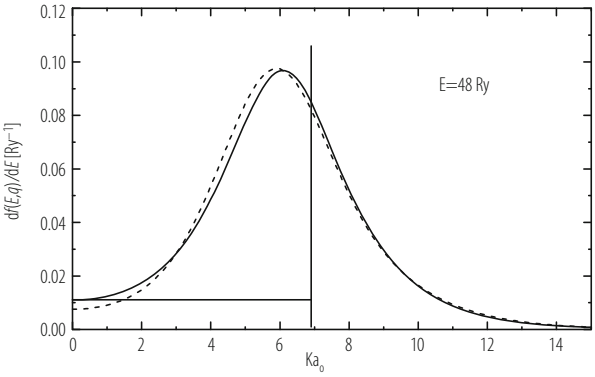


Fig. 2.5 Generalized oscillator strength density for Si for an energy transfer $E = 48\,\mathrm{Ry}$ to the 2p-shell electrons [41–44], as function of ka_0 (where $k^2a_0^2 = Q/\mathrm{Ry}$). Solid line: calculated with Herman-Skilman potential, dashed line: hydrogenic approximation [45, 46]. The horizontal and vertical line define the FVP approximation (Sect. 2.3.3)

We will discuss this term in more detail in Sect. 2.3.3. The total single-differential cross section,

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E} = \frac{\mathrm{d}\sigma^{(1)}}{\mathrm{d}E} + \frac{\mathrm{d}\sigma^{(2)}}{\mathrm{d}E} + \frac{\mathrm{d}\sigma^{(3)}}{\mathrm{d}E} + \frac{\mathrm{d}\sigma^{(h)}}{\mathrm{d}E},$$

is shown in Fig. 2.6 for particles with $\beta\gamma = 4$ in silicon which, at present, is the only material for which calculations based on the Bethe-Fano method are available.

2.3.3 *Fermi Virtual-Photon (FVP) Method*

In the Bethe-Fano algorithm discussed in the previous section, the dielectric function $\varepsilon\,(q, E)$ was approximated at low momentum transfer by its optical limit $\varepsilon\,(E)$. In the Fermi virtual-photon (FVP) or Photoabsorption Ionisation (PAI) model [6, 47, 48], this approximation is extended to the entire domain $q^2 < 2mE$. Guided by the shape of the hydrogenic GOS, the remaining contribution to $\mathrm{Im}\,(-1/\varepsilon\,(q, E))$ required to satisfy the Bethe sum rule

$$\int_0^\infty E\,\mathrm{Im}\left(\frac{-1}{\varepsilon\,(q, E)}\right)\mathrm{d}E = \frac{\pi}{2}\left(\hbar\Omega_p\right)^2 \quad \forall q, \tag{2.23}$$

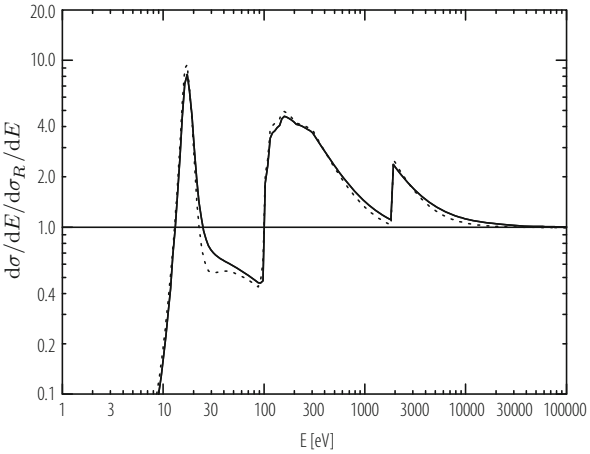


Fig. 2.6 Differential cross section $\mathrm{d}\sigma/\mathrm{d}E$, divided by the Rutherford cross section $\mathrm{d}\sigma_R/\mathrm{d}E$, for particles with $\beta\gamma = 4$ in silicon, calculated with two methods. The abscissa is the energy loss E in a single collision. The Rutherford cross section is represented by the horizontal line at 1.0. The solid line was obtained [41] with the Bethe-Fano method (Sect. 2.3.2). The cross section calculated with the FVP method (Sect. 2.3.3) is shown by the dotted line. The functions all extend to $E_{\max} \sim 16\,\mathrm{MeV}$. The moments are $M_0 = 4$ collisions/µm and $M_1 = 386\,\mathrm{eV/\mu m}$ (Table 2.2)

is attributed to the scattering off free electrons (close collisions). This term is thus of the form $C\delta\left(E - q^2/(2m)\right)$, with the factor C being determined by the normalisation (2.23),

$$C = \frac{1}{E}\int_0^E E'\mathrm{Im}\left(\frac{-1}{\varepsilon\,(E')}\right)\mathrm{d}E'.$$

Combining the two terms, the longitudinal loss function becomes

$$\mathrm{Im}\left(\frac{-1}{\varepsilon\,(q,\,E)}\right) = \mathrm{Im}\left(\frac{-1}{\varepsilon\,(E)}\right)\Theta\left(E - \frac{q^2}{2m}\right) + \frac{\delta\left(E - \frac{q^2}{2m}\right)}{E}\int_0^E E'\mathrm{Im}\left(\frac{-1}{\varepsilon\,(E')}\right)\mathrm{d}E'.$$

In the transverse term, the largest contribution to the integral comes from the region $E \sim qc/\sqrt{\varepsilon}$, i.e. from the vicinity of the (real) photon dispersion relation, and one consequently approximates $\varepsilon\,(q,\,E)$ by $\varepsilon\,(E)$ throughout.

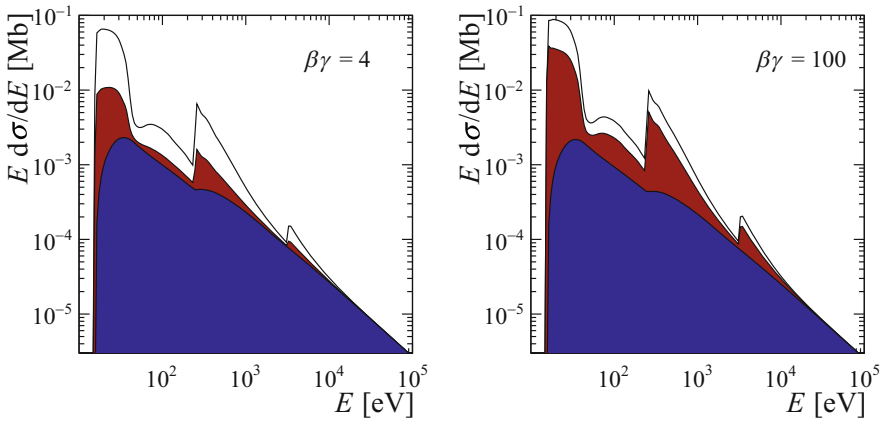


Fig. 2.7 Differential cross section $\mathrm{d}\sigma/\mathrm{d}E$ (scaled by the energy loss E) calculated using the FVP algorithm, for particles with $\beta\gamma = 4$ (left) and $\beta\gamma = 100$ (right) in argon (at atmospheric pressure, $T = 20\,^{\circ}\mathrm{C}$). The upper, unshaded area corresponds to the first term in Eq. (2.24), i.e. to the contribution from distant longitudinal collisions. The lower area corresponds to the contribution from close longitudinal collisions, given by the second term in Eq. (2.24). The intermediate area corresponds to the contribution from transverse collisions

The integration over q can then be carried out analytically and one obtains for the single-differential cross section $\mathrm{d}\sigma/\mathrm{d}E$

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E} = \frac{z^2\alpha}{\beta^2\pi N\hbar c}\left[\mathrm{Im}\left(\frac{-1}{\varepsilon\,(E)}\right)\ln\frac{2mc^2\beta^2}{E} + \frac{1}{E^2}\int_0^E E'\mathrm{Im}\left(\frac{-1}{\varepsilon\,(E')}\right)\mathrm{d}E'\right] + \frac{z^2\alpha}{\beta^2\pi N\hbar c}$$
$$\times\left[\mathrm{Im}\left(\frac{-1}{\varepsilon\,(E)}\right)\ln\frac{1}{\left|1-\beta^2\varepsilon\,(E)\right|} + \left(\beta^2 - \frac{\varepsilon_1\,(E)}{|\varepsilon\,(E)|^2}\right)\left(\frac{\pi}{2} - \arctan\frac{1-\beta^2\varepsilon_1\,(E)}{\beta^2\varepsilon_2\,(E)}\right)\right] \tag{2.24}$$

The relative importance of the different terms in Eq. (2.24) is illustrated in Fig. 2.7. The first two terms describe the contributions from longitudinal distant and close collisions. The contribution from transverse collisions (third and fourth term) is identical to $\mathrm{d}\sigma^{(3)}/\mathrm{d}E$ in the Bethe-Fano algorithm. As can be seen from Fig. 2.7, its importance grows with increasing $\beta\gamma$. The third term incorporates the relativistic density effect, i.e. the screening of the electric field due to the polarisation of the medium induced by the passage of the charged particle. In the transparency region $\varepsilon_2\,(E) = 0$, the fourth term can be identified with the cross section for the emission of Cherenkov photons. It vanishes for $\beta < 1/\sqrt{\varepsilon}$; above this threshold it becomes

$$\frac{\mathrm{d}\sigma^{(\mathrm{C})}}{\mathrm{d}E} = \frac{\alpha}{N\hbar c}\left(1 - \frac{1}{\beta^2\varepsilon}\right) \sim \frac{\alpha}{N\hbar c}\sin^2\theta_{\mathrm{C}},$$

where

$$\cos\theta_{\mathrm{C}} = \frac{1}{\beta\sqrt{\varepsilon}}.$$

Cherenkov detectors are discussed in detail in Chap. 7 of this book.

In the formulation of the PAI model by Allison and Cobb [6], the imaginary part ε_2 of the dielectric function is approximated by the photoabsorption cross section σ_γ,

$$\varepsilon_2(E) \sim \frac{N\hbar c}{E}\sigma_\gamma(E) \tag{2.25}$$

and the real part ε_1 is calculated from the Kramers-Kronig relation

$$\varepsilon_1(E) - 1 = \frac{2}{\pi}\mathsf{P}\int_0^\infty \frac{E'\varepsilon_2(E')}{E'^2 - E^2}\mathrm{d}E'.$$

In addition, the approximation $|\varepsilon(E)|^2 \sim 1$ is used. These are valid approximations if the refractive index[2] is close to one ($\mathsf{n} \sim 1$) and the attenuation coefficient k is small. For gases, this requirement is usually fulfilled for energies above the ionisation threshold.

Requiring only optical data as input, the FVP/PAI model is straightforward to implement in computer simulations. In the HEED program [49], the differential cross section $\mathrm{d}\sigma/\mathrm{d}E$ is split into contributions from each atomic shell, which enables one to simulate not only the energy transfer from the projectile to the medium but also the subsequent atomic relaxation processes (Sect. 2.6). The GEANT4 implementation of the PAI model is described in Ref. [50]. For reasons of computational efficiency, the photoabsorption cross section $\sigma_\gamma(E)$ is parameterised as a fourth-order polynomial in $1/E$. FVP calculations for Ne and Ar/CH_4 (90:10) are discussed in Ref. [51].

2.3.4 *Integral Quantities*

For validating and comparing calculations of the differential cross section, it is instructive to consider the moments M_i of $N\mathrm{d}\sigma/\mathrm{d}E$, in particular the inverse mean free path M_0 and the stopping power M_1.

[2]The complex refractive index and the dielectric function are related by $\mathsf{n} + \mathrm{i}\mathsf{k} = \sqrt{\varepsilon}$.

2.3.4.1 Inverse Mean Free Path

In the relativistic first-order Born approximation, the inverse mean free path for ionising collisions has the form [1, 2]

$$M_0 = \frac{2\pi z^2 (\alpha\hbar c)^2}{mc^2\beta^2} N \left[M^2 \left(\ln\left(\beta^2\gamma^2\right) - \beta^2 \right) + C \right], \tag{2.26}$$

where

$$M^2 = \int \frac{1}{E} \frac{\mathrm{d}f\,(E)}{\mathrm{d}E} \mathrm{d}E, \qquad C = M^2 \left(\ln \tilde{c} + \ln \frac{4}{\alpha^2} \right),$$

and $\tilde{c}$ is a material-dependent parameter that can be calculated from the generalised oscillator strength density. Calculations can be found, for example, in Refs. [53, 54]. As in the Bethe stopping formula (2.28) discussed below, a correction term can be added to Eq. (2.26) to account for the density effect [55].

The inverse mean free path for dipole-allowed discrete excitations is given by [2]

$$M_0^{(n)} = \frac{2\pi z^2 (\alpha\hbar c)^2}{mc^2\beta^2} N \frac{f_n}{E_n} \left[\ln\left(\beta^2\gamma^2\right) - \beta^2 + \ln \tilde{c}_n + \ln \frac{4}{\alpha^2} \right].$$

We can thus obtain a rough estimate of the relative frequencies of excitations and ionising collisions from optical data. In argon, for instance, the ratio of $\sum f_n/E_n$ and M^2 is ∼20% [25].

For gases, M_0 can be determined experimentally by measuring the inefficiency of a gas-filled counter operated at high gain ("zero-counting method"). Results (in the form of fit parameters M^2, C) from an extensive series of measurements, using electrons with kinetic energies between 0.1 and 2.7 MeV, are reported in Ref. [52]. Other sets of experimental data obtained using the same technique can be found in Refs. [56, 57]. Table 2.1 shows a comparison between measured and calculated values (using the FVP algorithm) of M_0 for particles with $\beta\gamma = 3.5$ at a temperature of 20 °C and atmospheric pressure. The inverse mean free path is

Table 2.1 Measurements [52] and calculations (using the FVP algorithm as implemented in HEED [49]) of M_0 for $\beta\gamma = 3.5$ at $T = 20\,^{\circ}\mathrm{C}$ and atmospheric pressure

	M_0 [cm^{-1}]	
Gas	Measurement	FVP
Ne	10.8	10.5
Ar	23.0	25.4
Kr	31.5	31.0
Xe	43.2	42.1
CO_2	34.0	34.0
CF_4	50.9	51.8
CH_4	24.6	29.4
iC_4H_{10}	83.4	90.9

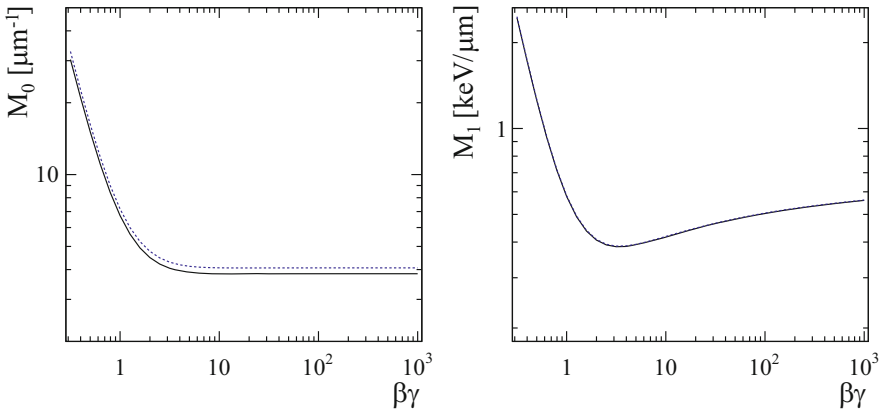


Fig. 2.8 Inverse ionisation mean free path (left) and stopping power (right) of heavy charged particles in silicon as a function of $\beta\gamma$, calculated using the Bethe-Fano algorithm (solid line) and the FVP model (dashed line). The two stopping power curves are virtually identical

sensitive to the detailed shape of the differential cross section $\mathrm{d}\sigma/\mathrm{d}E$ at low energies and, consequently, to the optical data used.

Figure 2.8(left) shows M_0 in solid silicon as a function of $\beta\gamma$, calculated using the Bethe-Fano and FVP algorithms. The difference between the results is $\sim 6-8\%$, as can also be seen from Table 2.2. Owing to the more detailed (and more realistic) modelling of the generalised oscillator strength density at intermediate Q, the Bethe-Fano algorithm can be expected to be more accurate than the FVP method.

2.3.4.2 Stopping Power

Let us first consider the average energy loss of a non-relativistic charged particle in a dilute gas, with the double-differential cross section given by Eq. (2.13),

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{2\pi z^2 (\alpha\hbar c)^2}{mc^2\beta^2} N \int_{E_{\min}}^{E_{\max}} \mathrm{d}E \int_{Q_{\min}}^{Q_{\max}} \frac{\mathrm{d}Q}{Q} \frac{\mathrm{d}f(E,q)}{\mathrm{d}E}.$$

As an approximation, we assume that the integrations over Q and E can be interchanged and the integration limits $Q_{\min}, Q_{\max}$ (which depend on E) be replaced by average values $\overline{Q}_{\min} = I^2/\left(2m\beta^2c^2\right)$, $\overline{Q}_{\max} = E_{\max}$ [58]. Using the Bethe sum rule (2.23), we then obtain

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{2\pi z^2 (\alpha\hbar c)^2}{mc^2\beta^2} NZ \ln\frac{2mc^2\beta^2 E_{\max}}{I^2},$$

where the target medium is characterised by a single parameter: the "mean ionisation energy" I, defined by

$$\ln I = \frac{1}{Z}\int \mathrm{d}E \ln E \frac{\mathrm{d}f\,(E)}{\mathrm{d}E}$$

in terms of the dipole oscillator strength density, or

$$\ln I = \frac{2}{\pi\left(\hbar\Omega_p\right)^2}\int \mathrm{d}E\; E\,\mathrm{Im}\left(\frac{-1}{\varepsilon\,(E)}\right)\ln E. \tag{2.27}$$

in terms of the dielectric loss function.

In the relativistic case, one finds the well-known Bethe stopping formula

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{2\pi z^2\,(\alpha\hbar c)^2}{mc^2\beta^2} NZ\left[\ln\frac{2mc^2\beta^2\gamma^2 E_{\max}}{I^2} - 2\beta^2 - \delta\right], \tag{2.28}$$

where δ is a correction term accounting for the density effect [59].

Sets of stopping power tables for protons and alpha particles are available in ICRU report 49 [60] and in the PSTAR and ASTAR online databases [61]. Tables for muons are given in Ref. [62]. These tabulations include stopping power contributions beyond the first-order Born approximation, such as shell corrections [42, 45, 46] and the Barkas-Andersen effect [63–65].

The stopping power in silicon obtained from the Bethe-Fano algorithm (Sect. 2.3.2) has been found to agree with measurements within ±0.5% [41]. As can be seen from Table 2.2 and Fig. 2.8, FVP and Bethe-Fano calculations for M_1 in silicon are in close agreement, with differences <1%.

In addition to M_0, M_1, Table 2.2 also includes the most probable value of the energy loss spectrum in an 8 µm thick layer of silicon. For thin absorbers, as will be discussed in Sect. 2.5, the stopping power $\mathrm{d}E/\mathrm{d}x$ is not a particularly meaningful quantity for characterising energy loss spectra. Because of the asymmetric shape of the differential cross section $\mathrm{d}\sigma/\mathrm{d}E$, the most probable value Δ_p of the energy loss distribution is typically significantly smaller than the average energy loss $\langle\Delta\rangle = M_1 x$.

2.4 Electron Collisions and Bremsstrahlung

The formalism for computing the differential cross section $\mathrm{d}\sigma/\mathrm{d}E$ for collisions of heavy charged particles with the electrons of the target medium, discussed in Sect. 2.3, is also applicable to electron and positron projectiles, except that the asymptotic close-collision cross section (2.11) is replaced by the Møller and Bhabha cross sections respectively [8, 66]. When evaluating the inverse mean free path M_0 or the stopping power M_1, we further have to take into account that the energy loss

Table 2.2 Integral properties of collision cross sections for Si calculated with Bethe-Fano (B-F) and FVP algorithms

	M_0 [μm^{-1}]		M_1 [eV/μm]		Δ_p/x [eV/μm]	
$\beta\gamma$	B-F	FVP	B-F	FVP	B-F	FVP
0.316	30.325	32.780	2443.72	2465.31	1677.93	1722.92
0.398	21.150	22.781	1731.66	1745.57	1104.90	1135.68
0.501	15.066	16.177	1250.93	1260.18	744.60	765.95
0.631	11.056	11.840	928.70	935.08	520.73	536.51
0.794	8.433	9.010	716.37	720.98	381.51	394.03
1.000	6.729	7.175	578.29	581.79	294.54	304.89
1.259	5.632	5.996	490.84	493.65	240.34	249.25
1.585	4.932	5.245	437.34	439.72	207.15	215.02
1.995	4.492	4.771	406.59	408.70	187.39	194.60
2.512	4.218	4.476	390.95	392.89	176.30	183.06
3.162	4.051	4.296	385.29	387.12	170.70	177.16
3.981	3.952	4.189	386.12	387.89	168.59	174.81
5.012	3.895	4.127	391.08	392.80	168.54	174.63
6.310	3.865	4.094	398.54	400.24	169.62	175.60
7.943	3.849	4.076	407.39	409.07	171.19	177.10
10.000	3.842	4.068	416.91	418.58	172.80	178.66
12.589	3.839	4.064	426.63	428.29	174.26	180.06
15.849	3.839	4.063	436.30	437.96	175.45	181.24
19.953	3.839	4.063	445.79	447.44	176.36	182.14
25.119	3.840	4.063	455.03	456.68	177.04	182.79
31.623	3.840	4.064	463.97	465.63	177.53	183.28
39.811	3.841	4.064	472.61	474.27	177.86	183.61
50.119	3.842	4.065	480.93	482.58	178.09	183.83
63.096	3.842	4.065	488.90	490.55	178.22	183.95
79.433	3.842	4.065	496.52	498.17	178.32	184.06
100.000	3.842	4.066	503.77	505.42	178.38	184.10
125.893	3.843	4.066	510.66	512.31	178.43	184.15
158.489	3.843	4.066	517.20	518.84	178.44	184.17
199.526	3.843	4.066	523.40	525.05	178.47	184.18
251.189	3.843	4.066	529.29	530.94	178.48	184.18
316.228	3.843	4.066	534.91	536.56	178.48	184.21
398.107	3.843	4.066	540.28	541.92	178.48	184.22
501.187	3.843	4.066	545.43	547.08	178.48	184.22
630.958	3.843	4.066	550.40	552.05	178.48	184.22
794.329	3.843	4.066	555.21	556.86	178.48	184.22
1000.000	3.843	4.066	559.89	561.54	178.48	184.22

The third column shows the most probable value Δ_p of the energy loss spectrum divided by the track length x, for $x = 8\,\mu m$. The minimum values for M_0 are at $\beta\gamma \sim 18$, for M_1 at $\beta\gamma \sim 3.2$, for Δ_p at $\beta\gamma \sim 5$. The relativistic rise for M_0 is 0.1%, for M_1 it is 45%, for Δ_p it is 6%

of an electron in an ionising collision is limited to half of its kinetic energy,

$$E_{\max} = \frac{1}{2} mc^2 (\gamma - 1) , \tag{2.29}$$

as primary and secondary electron are indistinguishable. Stopping power tables for electrons are available in ICRU report 37 [67] and in the ESTAR database [61].

The other main mechanism by which fast electrons and positrons lose energy when traversing matter is the emission of radiation (bremsstrahlung) due to deflections in the electric field of the nucleus and the atomic electrons.

2.4.1 Bremsstrahlung

Let us first consider electron-nucleus bremsstrahlung, the first quantum-mechanical description of which was developed by Bethe and Heitler [68]. The differential cross section (per atom) for the production of a bremsstrahlung photon of energy E by an incident electron of kinetic energy T is given by [8, 68]

$$\frac{\mathrm{d}\sigma_{\mathrm{rad}}}{\mathrm{d}E} = 4\alpha^3 \left(\frac{\hbar c}{mc^2} \right)^2 Z^2 \frac{F(u, T)}{E}, \tag{2.30}$$

where $u = E/\left(\gamma mc^2\right)$ denotes the ratio of the photon energy to the projectile energy. Expressions for the function $F(u, T)$ are reviewed in Ref. [69] and can be fairly complex. Amongst other parameters, $F(u, T)$ depends on the extent to which the charge of the nucleus is screened by the atomic electrons. In the first-order Born approximation and in the limit of complete screening, applicable at high projectile energies, one obtains [8, 68, 69]

$$F(u) = \left(1 + (1-u)^2 - \frac{2}{3}(1-u) \right) \ln \frac{183}{Z^{1/3}} + \frac{1}{9}(1-u) . \tag{2.31}$$

The theoretical description of electron-electron bremsstrahlung is similar to the electron-nucleus case, except that the differential cross section is proportional to Z instead of Z^2. To a good approximation, we can include electron-electron bremsstrahlung in Eq. (2.30) by replacing the factor Z^2 by $Z(Z+1)$.

The inverse mean free path for the emission of a bremsstrahlung photon with energy $E > E_{\mathrm{cut}}$ is given by

$$\lambda^{-1} = M_0 = N \int_{E_{\mathrm{cut}}}^{T} \frac{\mathrm{d}\sigma_{\mathrm{rad}}}{\mathrm{d}E} \mathrm{d}E .$$

If we neglect the term $(1-u)/9$ in Eq. (2.31), we find for the radiative stopping power at $T \gg mc^2$

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = M_1 = N \int_0^T E \frac{\mathrm{d}\sigma_{\mathrm{rad}}}{\mathrm{d}E} \mathrm{d}E \sim \frac{T}{X_0}, \tag{2.32}$$

where the parameter X_0, defined by

$$\frac{1}{X_0} = 4\alpha^3 \left(\frac{\hbar c}{mc^2}\right)^2 NZ\,(Z+1) \ln \frac{183}{Z^{1/3}}, \tag{2.33}$$

is known as the radiation length. Values of X_0 for many commonly used materials can be found in Ref. [70] and on the PDG webpage [71]. Silicon, for instance, has a radiation length of $X_0 \sim 9.37\,\mathrm{cm}$ [71].

Being approximately proportional to the kinetic energy of the projectile, the radiative stopping power as a function of T increases faster than the average energy loss due to ionising collisions given by Eq. (2.28). At high energies—more precisely, above a so-called critical energy ($\sim$38 MeV in case of silicon [71])—bremsstrahlung therefore represents the dominant energy loss mechanism of electrons and positrons.

2.5 Energy Losses Along Tracks: Multiple Collisions and Spectra

Consider an initially monoenergetic beam of identical particles traversing a layer of material of thickness x. Due to the randomness both in the number of collisions and in the energy loss in each of the collisions, the total energy loss Δ in the absorber will vary from particle to particle. Depending on the use case, the kinetic energy of the particles, and the thickness x, different techniques for calculating the probability distribution $f(\Delta, x)$—known as "straggling function" [72]—can be used.

Our focus in this section is on scenarios where the average energy loss in the absorber is small compared to the kinetic energy T of the incident particle (as is usually the case in vertex and tracking detectors), such that the differential cross section $\mathrm{d}\sigma/\mathrm{d}E$ and its moments do not change significantly between the particle's entry and exit points in the absorber. The number of collisions n then follows a Poisson distribution

$$p(n, x) = \frac{\langle n \rangle^n}{n!} \mathrm{e}^{-\langle n \rangle}, \tag{2.34}$$

with mean $\langle n \rangle = x M_0$. The probability $f^{(1)}(E)\,\mathrm{d}E$ for a particle to lose an amount of energy between E and $E + \mathrm{d}E$ in a single collision is given by the normalised

differential cross section,

$$f^{(1)}(E) = \frac{1}{M_0} N \frac{\mathrm{d}\sigma}{\mathrm{d}E},$$

and the probability distribution for a total energy loss Δ in n collisions is obtained from n-fold convolution of $f^{(1)}$,

$$f^{(n)}(\Delta) = \underbrace{\left(f^{(1)} \otimes f^{(1)} \otimes \cdots \otimes f^{(1)}\right)}_{n \text{ times}}(\Delta) = \int \mathrm{d}E \; f^{(n-1)}(\Delta - E)\, f^{(1)}(E),$$

as illustrated in Figs. 2.9 and 2.10.

The probability distribution for a particle to suffer a total energy loss Δ over a fixed distance x is given by [72, 73]

$$f(\Delta, x) = \sum_{n=0}^{\infty} p(n, x)\, f^{(n)}(\Delta), \tag{2.35}$$

where $f^{(0)}(\Delta) = \delta(\Delta)$. Equation (2.35) can be evaluated in a stochastic manner (Sect. 2.5.1), by means of direct numerical integration (Sect. 2.5.2), or by using integral transforms (Sect. 2.5.3).

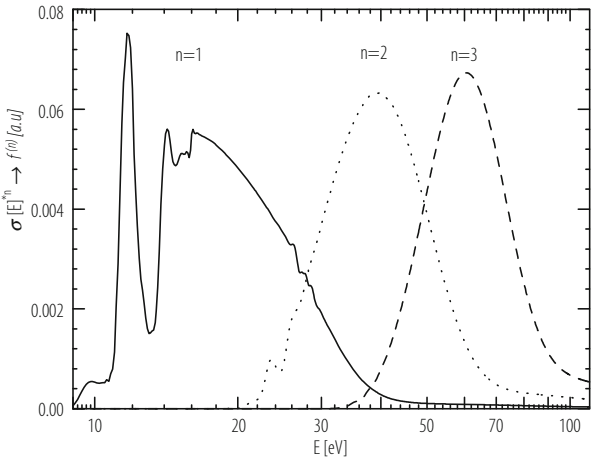


Fig. 2.9 Distributions $f^{(n)}$ of the energy loss in n collisions (n-fold convolution of the single-collision energy loss spectrum) for Ar/CH_4 (90:10)

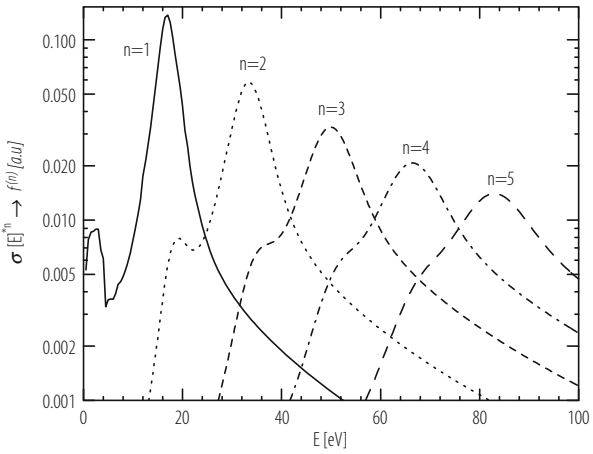


Fig. 2.10 Distributions $f^{(n)}$ of the energy loss in n collisions for solid silicon. The plasmon peak at ∼17 eV appears in each spectrum at $E \sim n \times 17$ eV, and its FWHM is proportional to $\sqrt{n}$. The structure at ∼2 eV appears at $2 + 17(n - 1)$ eV, but diminishes with increasing n. For $n = 6$ (not shown) the plasmon peak (at 102 eV) merges with the L-shell energy losses at 100 eV, also see Fig. 2.12

2.5.1 Monte Carlo Method

In a detailed Monte Carlo simulation, the trajectory of a single incident particle is followed from collision to collision. The required ingredients are the inverse mean free path $M_0^{(i)}$ and the cumulative distribution function,

$$\Phi^{(i)}(E) = \frac{1}{M_0^{(i)}} \int_0^E N \frac{\mathrm{d}\sigma^{(i)}}{\mathrm{d}E'} \mathrm{d}E', \tag{2.36}$$

for each interaction process i (electronic collisions, bremsstrahlung, etc.) to be taken into account in the simulation. The distance Δx between successive collisions follows an exponential distribution and is sampled according to

$$\Delta x = -\frac{\ln r}{\lambda^{-1}},$$

where $r \in (0, 1]$ is a uniformly distributed random number and

$$\lambda^{-1} = \sum_i M_0^{(i)}$$

is the total inverse mean free path. After updating the coordinates of the particle, the collision mechanism to take place is chosen based on the relative frequencies $M_0^{(i)}/\lambda^{-1}$. The energy loss in the collision is then sampled by drawing another uniform random variate $u \in [0, 1]$, and determining the corresponding energy loss E from the inverse of the cumulative distribution,

$$E = \Phi^{-1}(u) .$$

In general, the new direction after the collision will also have to be sampled from a suitable distribution. The above procedure is repeated until the particle has left the absorber. The spectrum $f(\Delta, x)$ is found by simulating a large number of particles and recording the energy loss Δ in a histogram. Advantages offered by the Monte Carlo approach include its straightforward implementation, the possibility of including interaction mechanisms other than inelastic scattering (bremsstrahlung, elastic scattering etc.), and the fact that it does not require approximations to the shape of $\mathrm{d}\sigma/\mathrm{d}E$ to be made.

For thick absorbers, detailed simulations can become unpractical due to the large number of collisions, and the need to update the inverse mean free path M_0 and the cumulative distribution $\Phi(E)$ following the change in velocity of the particle.

In "mixed" simulation schemes, a distinction is made between "hard" collisions which are simulated individually, and "soft" collisions (e.g. elastic collisions with a small angular deflection of the projectile, or emission of low-energy bremsstrahlung photons) the cumulative effect of which is taken into account after each hard scattering event. Details on the implementation of mixed Monte Carlo simulations can be found, for example, in the PENELOPE user guide [74].

2.5.2 Convolutions

For short track segments, one can calculate the distributions $f^{(n)}$ explicitly by numerical integration and construct $f(\Delta, x)$ directly using Eq. (2.35). A computationally more efficient approach is the absorber doubling method [41, 75], which proceeds as follows. Consider a step x that is small compared to the mean free path such that $\langle n \rangle \ll 1$ (in practice: $\langle n \rangle < 0.01$ [76]). Expanding Eq. (2.35) in powers of $\langle n \rangle$ and retaining only constant and linear terms gives

$$f(E, x) \sim (1 - \langle n \rangle) f^{(0)}(E) + \langle n \rangle f^{(1)}(E) .$$

The straggling function for a distance $2x$ is then calculated using

$$f(\Delta, 2x) = \int_0^{\Delta} f(\Delta - E, x) f(E, x) \, \mathrm{d}E .$$

This procedure is carried out k times until the desired thickness $2^k x$ is reached. Because of the tail of $f^{(1)}(E)$ towards large energy transfers, the numerical convolution is performed on a logarithmic grid. More details of the implementation can be found in Refs. [75, 76].

2.5.3 Laplace Transforms

In the Laplace domain, Eq. (2.35) becomes

$$F(s,x) = \mathcal{L}\{f(\Delta,x)\} = \mathrm{e}^{-\langle n\rangle}\sum_{n=0}^{\infty}\frac{\langle n\rangle^n}{n!}\mathcal{L}\{f^{(1)}(\Delta)\}^n$$
$$= \exp\left[-Nx\int_0^{\infty}\mathrm{d}E\left(1-\mathrm{e}^{-sE}\right)\frac{\mathrm{d}\sigma}{\mathrm{d}E}\right].$$

Following Landau [20], we split the integral in the exponent in two parts,

$$Nx\int_0^{\infty}\mathrm{d}E\left(1-\mathrm{e}^{-sE}\right)\frac{\mathrm{d}\sigma}{\mathrm{d}E} = Nx\int_0^{E_1}\mathrm{d}E\left(1-\mathrm{e}^{-sE}\right)\frac{\mathrm{d}\sigma}{\mathrm{d}E} + Nx\int_{E_1}^{\infty}\mathrm{d}E\left(1-\mathrm{e}^{-sE}\right)\frac{\mathrm{d}\sigma}{\mathrm{d}E},$$

where E_1 is chosen to be large compared to the ionisation threshold while at the same time satisfying $sE_1 \ll 1$. For energy transfers exceeding E_1, the differential cross section is assumed to be given by the asymptotic expression for close collisions (2.11); for $E < E_1$, it is not specified.

Using $\exp(-sE) \sim 1 - sE$, we obtain for the first term

$$I_1 = Nx\int_0^{E_1}\mathrm{d}E\frac{\mathrm{d}\sigma}{\mathrm{d}E}\left(1-\mathrm{e}^{-sE}\right) \sim Nxs\int_0^{E_1}\mathrm{d}E\frac{\mathrm{d}\sigma}{\mathrm{d}E}E.$$

We can therefore evaluate I_1 by subtracting the contribution due to energy transfers between E_1 and $E_{\max}$ according to Eq. (2.11) from the total average energy loss $x\mathrm{d}E/\mathrm{d}x = \langle\Delta\rangle$,

$$I_1 \sim s\langle\Delta\rangle - s\xi\left(\ln\frac{E_{\max}}{E_1} - \beta^2\right),$$

where we have introduced the variable

$$\xi = x\frac{2\pi z^2(\alpha\hbar c)^2 NZ}{mc^2\beta^2}.$$

For evaluating the second integral, we approximate $\mathrm{d}\sigma/\mathrm{d}E$ by the Rutherford cross section $\mathrm{d}\sigma_R/\mathrm{d}E \propto 1/E^2$. Because of the rapid convergence of the integral for $sE \gg 1$, we further assume that the upper integration limit can be extended to infinity (instead of truncating $\mathrm{d}\sigma/\mathrm{d}E$ at $E_{\max}$). Integrating by parts and substituting $z = sE$ yields

$$I_2 = \xi \int_{E_1}^{\infty} \mathrm{d}E \frac{1-\mathrm{e}^{-sE}}{E^2} = \xi \underbrace{\frac{1-\mathrm{e}^{-sE_1}}{E_1}}_{\sim s} + \xi s \int_{sE_1}^{\infty} \mathrm{d}z \frac{\mathrm{e}^{-z}}{z} \sim \xi s \left(1 + \int_{sE_1}^{1} \frac{\mathrm{d}z}{z} - C\right),$$

where $C \sim 0.577215665$ is Euler's constant.[3] Combining the two terms I_1 and I_2, one obtains

$$F(s,x) = \exp\left[-\xi s \left(1 - C + \frac{\langle\Delta\rangle}{\xi} - \ln sE_{\max} + \beta^2\right)\right],$$

and, applying the inverse Laplace transform,

$$f(\Delta, x) = \mathcal{L}^{-1}\{F(s,x)\} = \frac{1}{\xi}\phi_L(\lambda), \tag{2.37}$$

where

$$\phi_L(\lambda) = \frac{1}{2\pi\mathrm{i}} \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} \mathrm{d}u\, \mathrm{e}^{u\ln u + \lambda u}. \tag{2.38}$$

is a universal function of the dimensionless variable

$$\lambda = \frac{\Delta - \langle\Delta\rangle}{\xi} - (1-C) - \beta^2 - \ln\frac{\xi}{E_{\max}}.$$

The maximum of $\phi_L(\lambda)$ is located at $\lambda \sim -0.222782$ and the most probable energy loss is, consequently, given by

$$\Delta_{\mathrm{p}} \sim \langle\Delta\rangle + \xi\left(0.2 + \beta^2 + \ln\kappa\right), \tag{2.39}$$

[3] $$-C = \int_0^1 \mathrm{d}z \frac{\mathrm{e}^{-z}-1}{z} + \int_1^{\infty} \mathrm{d}z \frac{\mathrm{e}^{-z}}{z} = -0.577215665\ldots$$

where $\kappa = \xi/E_{\max}$. The full width at half maximum (FWHM) of the Landau distribution[4] (2.37) is approximately 4.02ξ.

A somewhat unsatisfactory aspect of $\phi_L(\lambda)$ is that its mean is undefined (a consequence of allowing arbitrarily large energy transfers $E > E_{\max}$). This deficiency was overcome by Vavilov [77] who, taking account of the kinematically allowed maximum energy transfer $E_{\max}$ and using the differential cross section (2.11) in I_2, obtained

$$f(\Delta, x) = \frac{1}{\xi}\phi_V(\lambda), \qquad \phi_V(\lambda) = \frac{1}{2\pi \mathrm{i}} \mathrm{e}^{\kappa\left(1+\beta^2 C\right)} \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} \exp\left(\psi(u) + \lambda u\right) \mathrm{d}u,$$

where

$$\psi(u) = u \ln \kappa + \left(u + \beta^2 \kappa\right) \left(\int_{u/\kappa}^{\infty} \frac{\mathrm{e}^{-t}}{t} \mathrm{d}t + \ln \frac{u}{\kappa} \right) - \kappa \mathrm{e}^{-u/\kappa}.$$

For small values of κ ($\kappa < 0.01$ [77]) the Vavilov distribution tends to the Landau distribution, while for $\kappa \gg 1$ it approaches a Gaussian distribution with $\sigma^2 = \xi E_{\max}\left(1 - \beta^2/2\right)$ [78]. Algorithms for the numerical evaluation of ϕ_L and ϕ_V and for drawing random numbers from these distributions are discussed e.g. in Refs. [78–81] and are implemented in ROOT [82].

Attempts have been made [83, 84] to improve the Landau-Vavilov method with respect to the treatment of distant collisions by including the second order term in the expansion of $\exp(-sE)$ in I_1. The results are akin to convolving ϕ_L or ϕ_V with a Gaussian distribution (expressions for estimating the standard deviation σ of the Gaussian are reviewed in Ref. [41]).

2.5.4 Examples

Let us first consider track segments for which the projectile suffers on average only tens of collisions. At the minimum of M_0, $\langle n \rangle = 10$ corresponds to a track length $x \sim 4\,\mathrm{mm}$ for argon (at atmospheric pressure, $T = 20\,^{\circ}\mathrm{C}$) and $x \sim 2\,\mu\mathrm{m}$ for silicon (Tables 2.1 and 2.2). As can be seen from Figs. 2.11 and 2.12, the features of the differential cross section $\mathrm{d}\sigma/\mathrm{d}E$ are clearly visible in the straggling functions $f(\Delta, x)$. These spectra cannot be described by a Landau distribution (or variants thereof) and need to be calculated using Monte Carlo simulation or numerical convolution.

[4] In high-energy physics parlance, the term "Landau distribution" is sometimes used for energy loss spectra $f(\Delta, x)$ in general. In this chapter, it refers only to the distribution given by Eq. (2.37).

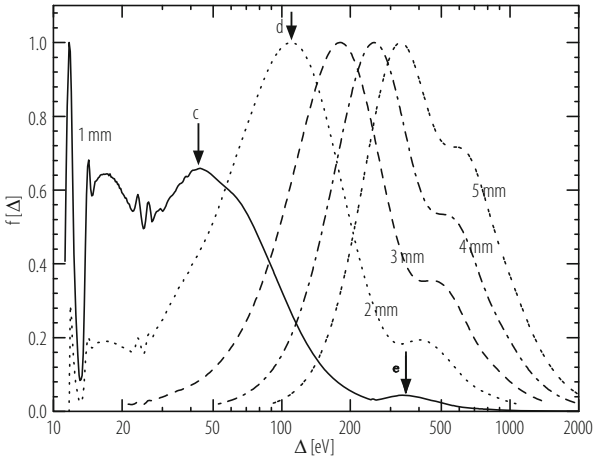


Fig. 2.11 Straggling functions for singly charged particles with $\beta\gamma = 4.48$ traversing segments of length $x = 1 \ldots 5$ mm in Ar. The inverse mean free path M_0 is 30 collisions/cm. The functions are normalised to unity at the most probable value. The broad peak at ∼17 eV is due to single collisions, see Fig. 2.9. For two collisions it broadens and shifts to about 43 eV, marked c, and for $n = 3$ it can be seen at d. It may be noted that the peak at 11.7 eV (if the function is normalised to unit area) is exactly proportional to $\langle n \rangle \exp(-\langle n \rangle)$, as expected from Eq. (2.35). Energy losses to L-shell electrons of Ar (with a binding energy of ∼250 eV) appear at e, for $x = 1$ mm they have an amplitude of 0.04. For $x > 2$ mm, peak c disappears, and peak d becomes the dominant contribution defining the most probable energy loss Δ_p. The buildup for peak e at 440–640 eV is the contribution from L-shell collisions. It appears roughly at 250 eV+Δ_p. The inverse mean free path for collisions with $E > 250$ eV is only 1.7 collisions/cm, thus the amplitude of the peak e is roughly proportional to x. The Bethe mean energy loss is 250 eV/mm

With increasing number of collisions, the detailed features of the differential cross section become "washed out" and the energy loss spectra $f(\Delta, x)$ tend to the Landau shape but are typically broader, as shown in Figs. 2.13 and 2.14. Reasonable agreement with measured energy loss spectra for thin absorbers can often be achieved by fit functions based on the convolution of a Landau/Vavilov distribution and a Gaussian distribution. For a predictive calculation of $f(\Delta, x)$, however, numerical convolution or a Monte Carlo simulation are usually needed.

2.5.5 Methods for Thick Absorbers

In order to compute the energy loss distribution for a layer of material in which the kinetic energies T of the traversing particles change considerably (i.e. by more than 5–10% [86]), we divide the absorber in segments of length x that are sufficiently small such that the straggling function $f(\Delta, x)$ can be calculated using the methods for thin absorbers described above. Let $\phi(y, T)$ be the distribution of

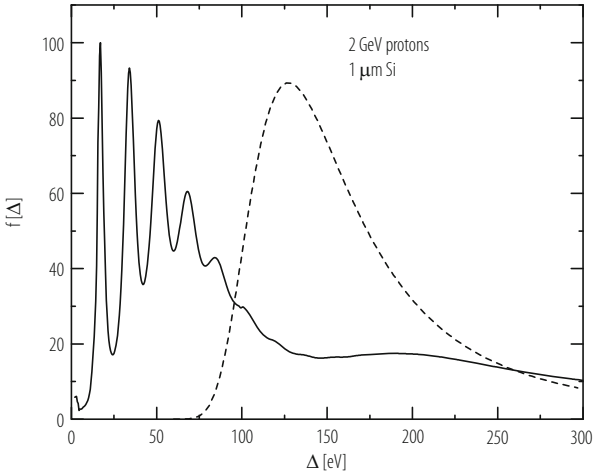


Fig. 2.12 Straggling in 1 μm of Si ($\langle n \rangle = 4$) for particles with $\beta\gamma = 2.1$, compared to the Landau function (dashed line). The Bethe mean energy loss is $\langle \Delta \rangle = 400$ eV. Measured straggling functions of this type are given in Ref. [85]

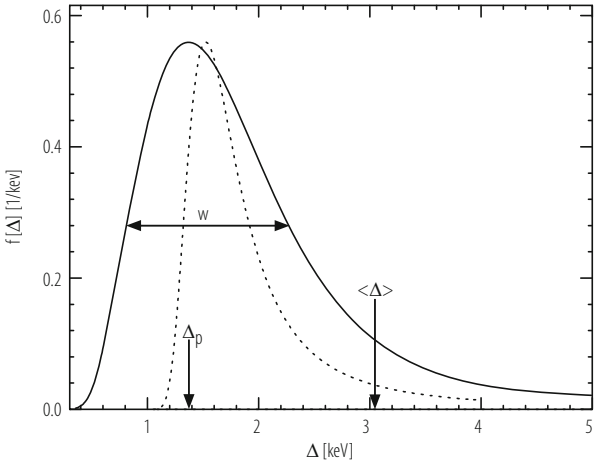


Fig. 2.13 Straggling function $f(\Delta)$ for particles with $\beta\gamma = 3.6$ traversing 1.2 cm of Ar gas ($\langle n \rangle = 36$) calculated using the convolution method (solid line) compared to the Landau distribution (dashed line). Parameters describing $f(\Delta)$ are the most probable energy loss Δ_p, i.e. the position of the maximum of the straggling function, at 1371 eV, and the full width at half maximum (FWHM) $w = 1463$ eV. The Bethe mean energy loss is $\langle \Delta \rangle = 3044$ eV. The peak of the Landau function is at 1530 eV

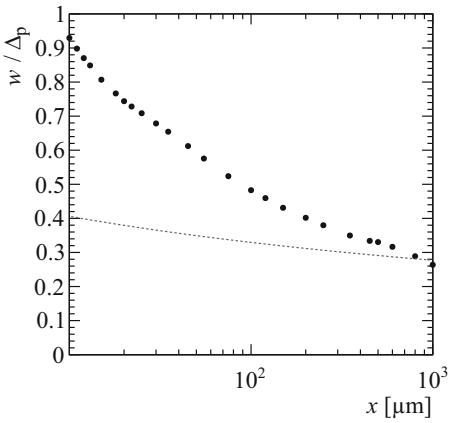


Fig. 2.14 Relative width (full width at half maximum w divided by the most probable value Δ_p) of the straggling spectrum $f(\Delta, x)$ as function of the absorber thickness x, for particles with $\beta\gamma \sim 3.16$ in silicon. The dashed line corresponds to the relative width of the Landau distribution. Circles represent results of a Monte Carlo simulation using the Bethe-Fano differential cross section

kinetic energies at a distance y in the absorber. If $f(\Delta, x)$ is known for all T, the spectrum of kinetic energies at $y + x$ can be calculated using

$$\phi(y + x, T) = \int \phi(y, T + \Delta)\, f(\Delta, x; T + \Delta)\, \mathrm{d}\Delta .$$

Scaling relations, discussed in Ref. [51], can be used to limit the number of thin-absorber distributions $f(\Delta, x; T)$ that need to be tabulated.

In a "condensed history" Monte Carlo simulation [87], the energy loss spectrum is calculated stochastically by sampling the energy loss over a substep x from a suitable thin-absorber distribution (e.g. a Vavilov function), and updating the kinetic energy T of the projectile after each substep.

2.6 Energy Deposition

Leaving the emission of Cherenkov radiation and other collective effects aside, charged-particle collisions with electrons in matter result in the promotion of one of the electrons in the target medium to a bound excited state or to the continuum. Both effects (excitation and ionisation) can be exploited for particle detection purposes. In scintillators, discussed in Chap. 3 of this book, part of the energy transferred to excitations is converted to light. Detectors based on ionisation measurement in gases and semiconductors are discussed in Chaps. 4 and 5. In the following we briefly review the main mechanisms determining the number of electron-ion pairs (in gases) or electron-hole pairs (in semiconductors) produced in the course of an ionising collision, along with their spatial distribution.

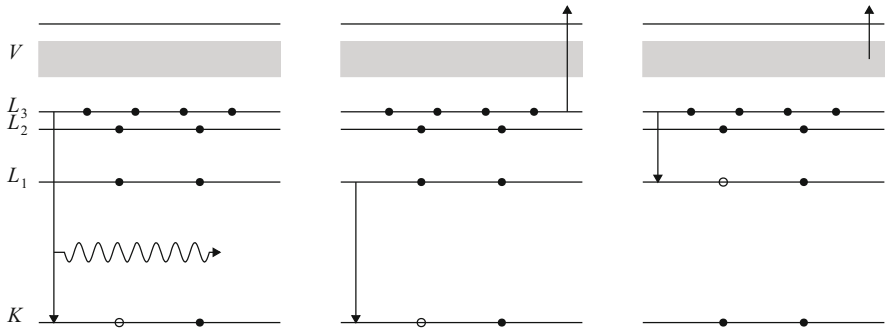


Fig. 2.15 After the ejection of an inner-shell electron, the resulting vacancy is filled by an electron from a higher shell. The energy released in the transition can either be carried away by a fluorescence photon (left) or be transferred to an electron in a higher shell (Auger process, middle). Coster-Kronig transitions (right) are Auger processes in which the initial vacancy is filled by an electron from the same shell

2.6.1 Atomic Relaxation

If a charged-particle collision (or a photoabsorption interaction) ejects an inner-shell electron from an atom, the resulting vacancy will subsequently be filled by an electron from a higher shell, giving rise to a relaxation chain which can proceed either radiatively, i.e. by emission of a fluorescence photon, or radiation-less (Auger effect). The two processes are illustrated schematically in Fig. 2.15. Fluorescence photons can in turn ionise another atom in the medium or, with a probability depending on the geometry of the device, escape from the detector. The fluorescence yield, i.e. the probability for a vacancy to be filled radiatively, increases with the atomic number Z: in silicon, for example, the average fluorescence yield is $\sim$5%, compared to $\sim$54% in germanium [88]. Compilations of fluorescence yields can be found in Refs. [88–91]. Tabulations of transition probabilities are available in the EADL database [92, 93].

2.6.2 Ionisation Statistics

The "primary" ionisation electron knocked out in a collision (and also the Auger electrons) may have kinetic energies exceeding the ionisation threshold of the medium and thus undergo further ionising collisions along their path. Electrons with a kinetic energy T that is large compared to the ionisation threshold are referred to as "delta" electrons; their energy distribution follows approximately the close-collision differential cross section, given by Eq. (2.11) for spin-zero particles. The number of electrons n_e produced in the energy degradation cascade of a delta electron with initial kinetic energy T is subject to fluctuations. The mean and variance of the

distribution of n_e are described by the average energy W required to produce an electron-ion (electron-hole) pair,

$$\langle n_e \rangle = \frac{T}{W}, \tag{2.40}$$

and the Fano factor F [94],

$$\sigma^2 = \langle (n - \langle n \rangle)^2 \rangle = F \langle n_e \rangle, \tag{2.41}$$

respectively. Both W and F are largely determined by the relative importance of ionising and non-ionising inelastic collisions, the latter including e.g. excitations or phonon scattering. If the cross sections for these processes are known, the distribution of n_e can be calculated using detailed Monte Carlo simulations. An example is the MAGBOLTZ program [95, 96], which includes the relevant cross sections for many commonly used detection gases. Inelastic cross sections of delta electrons in solids can be calculated based on the dielectric formalism discussed in Sect. 2.3.1 (in its non-relativistic version), often making using of optical data and a suitable model of the q-dependence of $\mathrm{Im}\,(-1/\varepsilon\,(q, E))$ as, for instance, in the Penn algorithm described in Ref. [97].

Measurements of W for electrons in gases as a function of the electron's initial kinetic energy are reported in Refs. [98–100, 102, 103]. As can be seen from Fig. 2.16, which shows measurements and calculations for CO_2, W increases towards low kinetic energies, while in the keV range and above it depends only weakly on T. For most gases and semiconductors typically used as sensitive media in particle detectors, the asymptotic (high-energy) W values are fairly well established. A compilation of recommended average W values, based on experimental data until 1978, is given in ICRU report 31 [101]. Critical reviews of W values and Fano factors including also more recent data can be found in Ref. [104]

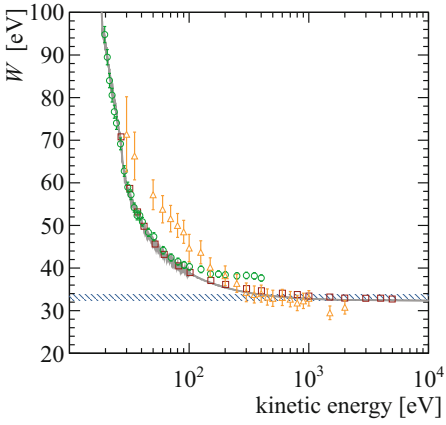


Fig. 2.16 W value for electrons in CO_2 as a function of the electron's initial kinetic energy according to measurements by Combecher [98] (circles), Smith and Booz [99] (triangles), and Waibel and Grosswendt [100] (squares). The grey band represents results of a Monte Carlo calculation using the cross sections implemented in MAGBOLTZ [96]. The hatched band corresponds to the high-energy value recommended in Ref. [101]

Table 2.3 Asymptotic W values and Fano factors for different gases and for solid silicon (at $T = 300\,\mathrm{K}$)

	W [eV]	F
Ne	35.4 ± 0.9 [101]	0.13–0.17 [104]
Ar	26.4 ± 0.5 [101]	0.15–0.17 [104]
Kr	24.4 ± 0.3 [101]	0.17–0.21 [104]
Xe	22.1 ± 0.1 [101]	0.124–0.24 [104]
CO_2	33.0 ± 0.7 [101]	0.32 [104]
CH_4	27.3 ± 0.3 [101]	0.22–0.26 [104]
iC_4H_{10}	23.4 ± 0.4 [101]	0.261 [106]
CF_4	34.3 [107]	
Si	3.67 ± 0.02 [108]	<0.1 [104]

Except for CF_4, the values shown are for measurements using electrons

and, with emphasis on noble gases, in Ref. [105]. Parameters for silicon and some commonly used gases are listed in Table 2.3.

Analogously to Eqs. (2.40) and (2.41) one can define W values and Fano factors characterising the distribution of the number of electrons produced by a heavy charged particle (provided that it is stopped completely in the medium) or by the absorption of a photon. The asymptotic W values for electrons and photons at high energies are in general very similar.

In gas mixtures without excitation transfers, the W value and Fano factor are, to a good approximation, given by the values in the pure gases, weighted by their respective concentrations. In mixtures where one of the components has excited states with energies exceeding the ionisation threshold of another component, excitation transfer can lead to a significant reduction of W and F with respect to the pure gases ("Jesse effect" [109]). Results for a number of binary gas mixtures from measurements with α particles can be found in Ref. [110].

2.6.3 Range

The spatial distribution of secondary ionisations produced by a delta electron can be characterised in terms of the electron range, i.e. the typical path length travelled by an electron before its energy falls below the ionisation threshold. In the literature, a number of different definitions of "range" exist, two of which—the fractional ionisation range R_x and the practical range R_p—are illustrated in Fig. 2.17. If the cross sections (including those for elastic scattering) are known, the range of delta electrons and, more generally, the ionisation pattern produced by a charged-particle collision, can be calculated using Monte Carlo techniques. As an example, Fig. 2.18 shows measurements of the 95% range in CH_4 as a function of the primary electron energy [102], together with calculated values based on the cross sections implemented in MAGBOLTZ.

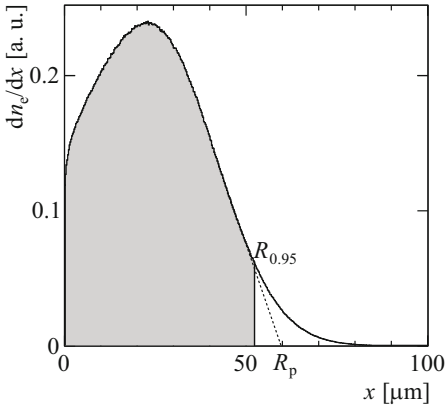


Fig. 2.17 Distribution of the coordinates (projected on the electron's initial direction) of ionising collisions by a $T = 1$ keV electron and its secondaries in methane (at atmospheric pressure, $T = 20\,^{\circ}\mathrm{C}$), calculated using the cross sections implemented in MAGBOLTZ. The fractional ionisation range R_x is defined as the projected distance along the electron's initial direction within which the fraction x of the total ionisation is produced [102]. The practical range R_p is determined by linear extrapolation from the region of steepest descent to the horizontal axis

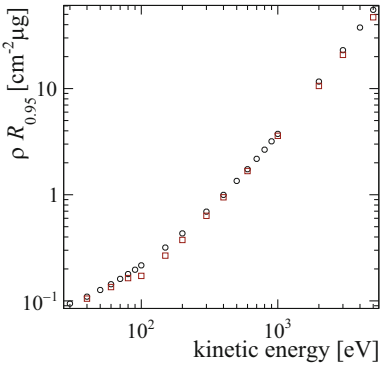


Fig. 2.18 Measurements [102] (squares) and MAGBOLTZ calculations (circles) of the 95% fractional ionisation range of electrons in methane (at atmospheric pressure)

In the absence of a detailed calculation, the semi-empirical formula by Kobetich and Katz [111, 112] can be used to estimate the practical range,

$$\rho R_p\,(T) = AT\left(1 - \frac{B}{1 + CT}\right),$$

where the parameters A, B, C are given by [112]

$$A = \left(0.81Z^{-0.38} + 0.18\right) \times 10^{-3}\,\mathrm{g\,cm^{-2}\,keV^{-1}},$$
$$B = 0.21Z^{-0.055} + 0.78,$$
$$C = \left(1.1Z^{0.29} + 0.21\right) \times 10^{-3}\,\mathrm{keV^{-1}}.$$

References

1. U. Fano, Annu. Rev. Nucl. Sci. **13**, 1 (1963).
2. M. Inokuti, Rev. Mod. Phys. **43**, 297 (1971).
3. M. Inokuti, Rev. Mod. Phys. **50**, 23 (1978).
4. H. Bichsel, Passage of Charged Particles Through Matter, in *Amer. Instrum. Phys. Handbook*, edited by D. E. Gray, McGraw Hill, 1972.
5. R. J. Gould, Physica **62**, 555 (1972).
6. W. W. M. Allison and J. H. Cobb, Annu. Rev. Nucl. Part. Sci. **30**, 253 (1980).
7. S. P. Ahlen, Rev. Mod. Phys. **52**, 121 (1980).
8. B. Rossi, *High-Energy Particles*, Prentice-Hall, New York, NY, 1952.
9. R. D. Evans, *The Atomic Nucleus*, McGraw Hill, 1967.
10. W. Heitler, *The Quantum Theory of Radiation, second edition*, Oxford University Press, 1949.
11. P. Sigmund, *Particle Penetration and Radiation Effects*, Springer, 2006.
12. N. J. Carron, *An Introduction to the Passage of Energetic Particles through Matter*, Taylor and Francis, 2007.
13. B. McParland, *Medical radiation dosimetry: Theory of charged particle collision energy loss*, Springer, 2014.
14. N. Bohr, Philos. Mag. **25**, 10 (1913).
15. N. Bohr, Philos. Mag. **30**, 581 (1915).
16. H. Bethe, Ann. Phys. **5**, 325 (1930).
17. H. Bethe, Z. Phys. **76**, 293 (1932).
18. E. Fermi, Z. Phys. **29**, 3157 (1924).
19. E. Fermi, Phys. Rev. **57**, 485 (1940).
20. L. D. Landau, J. Phys. USSR **8**, 201 (1944).
21. U. Fano, L. V. Spencer, and M. J. Berger, Penetration and Diffusion of X Rays, in *Handbuch der Physik*, edited by S. Flügge, volume 38/2, Springer, 1959.
22. W. Rösch, E. Tochilin, and F. H. Attix, editors, *Radiation Dosimetry; 2nd ed.*, Academic Press, 1968.
23. F. Salvat and J. M. Fernandez-Varea, Metrologia **46**, S112 (2009).
24. M. J. Berger et al., XCOM: Photon Cross Sections Database, https://www.nist.gov/pml/xcom-photon-cross-sections-database.
25. J. Berkowitz, *Atomic and molecular photo absorption*, Academic Press, 2002.
26. N. Sakamoto et al., Oscillator Strength Spectra and Related Quantities of 9 Atoms and 23 Molecules Over the Entire Energy Region, NIFS-DATA-109, National Institute for Fusion Science, 2010, available online at http://www.nifs.ac.jp/report/nifsdata.html.
27. B. L. Henke, E. M. Gullikson, and J. C. Davis, Atomic Data and Nucl. Data Tables **54**, 181 (1993), available online at http://henke.lbl.gov/optical_constants/.
28. J.-J. Yeh and I. Lindau, Atomic Data and Nucl. Data Tables **32**, 1 (1985).

29. M. B. Trzhaskovskaya, V. I. Nefedov, and V. G. Yarzhemsky, Atomic Data and Nucl. Data Tables **77**, 97 (2001).
30. M. B. Trzhaskovskaya, V. I. Nefedov, and V. G. Yarzhemsky, Atomic Data and Nucl. Data Tables **82**, 257 (2002).
31. E. D. Palik and E. J. Prucha, *Handbook of optical constants of solids*, Academic Press, 1998.
32. S. Adachi, *Optical constants of crystalline and amorphous semiconductors*, Kluwer Academic Publishers, 1999.
33. O. Klein and Y. Nishina, Z. Physik **52**, 853 (1929).
34. J. H. Hubbell, Photon Cross Sections, Attenuation Coefficients, and Energy Absorption Coefficients from 10 keV to 100 GeV, NSRDS-NBS 29, National Bureau of Standards, 1969, available online at https://www.nist.gov/srd/national-standard-reference-data-series.
35. P. M. Bergstrom and R. H. Pratt, Radiation Physics and Chemistry **50**, 3 (1997).
36. J. H. Hubbell, H. A. Gimm, and I. Øverbø, J. Phys. Chem. Ref. Data **9**, 1023 (1980).
37. D. Bote and F. Salvat, Phys. Rev. A **77**, 042701 (2008).
38. L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskii, *Electrodynamics of Continuous Media*, Butterworth, 1984.
39. J. Lindhard, Mat. Fys. Medd. Dan. Vid. Selsk. **28**, 1 (1954).
40. J. Lindhard and A. Winther, Mat. Fys. Medd. Dan. Vid. Selsk. **34**, 1 (1964).
41. H. Bichsel, Rev. Mod. Phys. **60**, 663 (1988).
42. H. Bichsel, Phys. Rev. A **65**, 052709 (2002).
43. H. Bichsel, Phys. Rev. A **46**, 5761 (1992).
44. H. Bichsel, K. M. Hanson, and M. E. Schillaci, Phys. Med. Biol. **27**, 959 (1982).
45. M. C. Walske, Phys. Rev. **88**, 1283 (1952).
46. M. C. Walske, Phys. Rev. **101**, 1940 (1956).
47. V. A. Chechin, L. P. Kotenko, G. I. Merson, and V. C. Yermilova, Nucl. Instr. Meth. **98**, 577 (1972).
48. F. Lapique and F. Piuz, Nucl. Instr. Meth. **175**, 297 (1980).
49. I. B. Smirnov, Nucl. Instr. Meth. A **554**, 474 (2005).
50. J. Apostolakis et al., Nucl. Instr. Meth. A **453**, 597 (2000).
51. H. Bichsel, Nucl. Instrum. Meth. A **562**, 154 (2006).
52. F. Rieke and W. Prepejchal, Phys. Rev. A **6**, 1507 (1972).
53. R. P. Saxon, Phys. Rev. A **8**, 839 (1973).
54. M. Inokuti, R. P. Saxon, and J. L. Dehmer, Int. J. Radiat. Phys. Chem. **7**, 109 (1975).
55. B. Sitar, G. I. Merson, V. A. Chechin, and Y. A. Budagov, *Ionization Measurements in High Energy Physics*, Springer, 1993.
56. G. W. McClure, Phys. Rev. **90**, 796 (1953).
57. G. Malamud, A. Breskin, R. Chechik, and A. Pansky, J. Appl. Phys. **74**, 3645 (1993).
58. H. A. Bethe and R. W. Jackiw, *Intermediate quantum mechanics*, Benjamin, 1986.
59. R. M. Sternheimer, M. J. Berger, and S. M. Seltzer, Atomic Data and Nucl. Data Tables **30**, 261 (1984).
60. *Stopping powers and ranges for protons and alpha particles*, International Commission on Radiation Units and Measurements, Washington, DC, 1993, ICRU Report 49.
61. M. J. Berger, J. S. Coursey, M. A. Zucker, and J. Chang, ESTAR, PSTAR, and ASTAR: Computer Programs for Calculating Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions, http://physics.nist.gov/Star.
62. D. E. Groom, N. V. Mokhov, and S. I. Striganoff, Atomic Data and Nucl. Data Tables **78**, 183 (2001).
63. W. H. Barkas, J. N. Dyer, and H. H. Heckman, Phys. Rev. Lett. **11**, 26 (1963).
64. H. H. Andersen, H. Simonsen, and H. Sørensen, Nuclear Physics A **125**, 171 (1969).
65. H. Bichsel, Phys. Rev. A **41**, 3642 (1990).
66. E. A. Uehling, Annu. Rev. Nucl. Sci. **4**, 315 (1954).
67. *Stopping powers for electrons and positrons*, International Commission on Radiation Units and Measurements, Washington, DC, 1979, ICRU Report 37.
68. H. Bethe and W. Heitler, Proc. Roy. Soc. A **146**, 83 (1934).

69. H. W. Koch and J. W. Motz, Rev. Mod. Phys. **31**, 920 (1959).
70. M. Tanabashi et al., Phys. Rev. D **98**, 030001 (2018).
71. D. E. Groom, Atomic and Nuclear Properties of Materials, http://pdg.lbl.gov/AtomicNuclearProperties.
72. E. J. Williams, Proc. Roy. Soc. A **125**, 420 (1929).
73. H. Bichsel and R. P. Saxon, Phys. Rev. A **11**, 1286 (1975).
74. F. Salvat, J. M. Fernández-Varea, and J. Sempau, PENELOPE-2014: A Code System for Monte Carlo Simulation of Electron and Photon Transport.
75. A. M. Kellerer, *Mikrodosimetrie*, Strahlenbiologisches Institut der Universität München, 1968, G.S.F. Bericht B-1.
76. A. M. Kellerer, Fundamentals of Microdosimetry, in *The Dosimetry of Ionizing Radiation*, edited by K. R. Kase, B. E. Bjängard, and F. H. Attix, Academic Press, 1985.
77. P. V. Vavilov, Sov. Phys. JETP **5**, 749 (1957).
78. M. J. Berger and S. M. Seltzer, *Studies in Penetration of Charged Particles in Matter*, The National Academies Press, Washington, DC, 1964.
79. W. Boersch-Supan, J. Res. Nat. Bur. Stand. **65B**, 245 (1961).
80. B. Schorr, Computer Phys. Comm. **7**, 215 (1974).
81. A. Rotondi and P. Montagna, Nucl. Instr. Meth. B **47**, 215 (1990).
82. R. Brun et al., Nucl. Instrum. Meth. A **389**, 81 (1997), http://root.cern.ch.
83. O. Blunck and S. Leisegang, Z. Phys. **128**, 500 (1950).
84. P. Shulek et al., Sov. J. Nucl. Phys. **4**, 400 (1967).
85. J. P. Perez, J. Sevely, and B. Jouffrey, Phys. Rev. A **16**, 1061 (1977).
86. C. Tschalär, Nucl. Instrum. Meth. **61**, 141 (1968).
87. M. J. Berger, Meth. Comput. Phys. **1**, 135 (1963).
88. M. O. Krause, J. Phys. Chem. Ref. Data **8**, 307 (1979).
89. W. Bambynek et al., Rev. Mod. Phys. **44**, 716 (1972).
90. J. H. Hubbell et al., J. Phys. Chem. Ref. Data **23**, 339 (1994).
91. J. H. Hubbell et al., J. Phys. Chem. Ref. Data **33**, 621 (2004).
92. S. T. Perkins et al., Tables and Graphs of Atomic Subshell and Relaxation Data Derived from the LLNL Evaluated Atomic Data Library (EADL), $Z = 1 - 100$, UCRL-50400, Lawrence Livermore National Laboratory, Livermore, CA, 1991.
93. D. E. Cullen, EPICS2017. Electron and Photon Interaction Cross Sections, https://www-nds.iaea.org/epics/.
94. U. Fano, Phys. Rev. **72**, 26 (1947).
95. S. F. Biagi, Nucl. Instr. Meth. A **421**, 234 (1999).
96. S. F. Biagi, Magboltz 8, http://magboltz.web.cern.ch/magboltz.
97. D. R. Penn, Phys. Rev. B **35**, 482 (1987).
98. D. Combecher, Radiation Research **84**, 189 (1980).
99. B. G. R. Smith and J. Booz, Experimental results on W-values and transmission of low energy electrons in gases, in *Sixth Symposium on Microdosimetry*, page 759, 1978.
100. E. Waibel and B. Grosswendt, Nucl. Instr. Meth. B **53**, 239 (1991).
101. *Average energy required to produce an ion pair*, International Commission on Radiation Units and Measurements, Washington, DC, 1979, ICRU Report 31.
102. E. Waibel and B. Grosswendt, Nucl. Instr. Meth. **211**, 487 (1983).
103. E. Waibel and B. Grosswendt, Study of W-values, practical ranges, and energy dissipation profiles of low-energy electrons in N_2, in *Eighth Symposium on Microdosimetry*, page 301, 1983.
104. Atomic and Molecular Data for Radiotherapy and Radiation Research, IAEA TECDOC 799, 1995.
105. I. Krajcar Bronić, Hoshasen: ionizing radiation **24**, 101 (1998).
106. D. Srdoč, B. Obelić, and I. Krajcar Bronić, J. Phys. B: At. Mol. Phys. **20**, 4473 (1987).
107. G. F. Reinking, L. G. Christophorou, and S. R. Hunter, J. Appl. Phys. **60**, 499 (1986).
108. R. H. Pehl, F. S. Goulding, D. A. Landis, and M. Lenzlinger, Nucl. Instr. Meth. **59**, 45 (1968).
109. W. P. Jesse and J. Sadauskis, Phys. Rev. **88**, 417 (1952).

110. T. E. Bortner, G. S. Hurst, M. Edmundson, and J. E. Parks, Alpha particle ionization of argon mixtures – further study of the role of excited states, ORNL-3422, Oak Ridge National Laboratory, 1963.
111. E. J. Kobetich and R. Katz, Phys. Rev. **170**, 391 (1968).
112. E. J. Kobetich and R. Katz, Nucl. Instr. Meth. **71**, 226 (1969).

Chapter 3
Scintillation Detectors for Charged Particles and Photons

P. Lecoq

3.1 Basic Detector Principles and Scintillator Requirements

3.1.1 Interaction of Ionizing Radiation with Scintillator Material

As any radiation detector, a scintillator is an absorbing material, which has the additional property to convert into light a fraction of the energy deposited by ionizing radiation. Charged and neutral particles interact with the scintillator material through the well-known mechanisms of radiation interactions in matter described by many authors [1, 2]. Charged particles continuously interact with the electrons of the scintillator medium through Coulomb interactions, resulting in atomic excitation or ionization. Neutral particles will first have to undergo a direct interaction with the nucleus producing recoil protons or spallation fragments, which will then transfer their energy to the medium in the same way as primary charged particles.

The rate of energy loss $(-\mathrm{d}E/\mathrm{d}x)$ for charged particles is strongly energy dependant. It is well described by the Bethe-Bloch formula (see Chap. 2) for incoming particles in the MeV-GeV range, with atomic shell corrections at lower energy and radiative loss corrections at higher energy. For heavy materials currently used as scintillators with a density of 6–8 g/cm^3, it is typically of the order of 10 MeV/cm for a minimum ionizing particle but it can be a factor up to 100 more at very low or very high energy (radiative losses).

P. Lecoq (✉)
CERN, Geneva, Switzerland
e-mail: Paul.Lecoq@cern.ch

C. W. Fabjan, H. Schopper (eds.), *Particle Physics Reference Library*,
https://doi.org/10.1007/978-3-030-35318-6_3

In the case of X- or γ- rays, the three fundamental mechanisms of electromagnetic interaction are [3]:

- Photo-absorption
- Compton scattering
- Electron-positron pair production

The dominant process at low energy (up to a few hundred keV for heavy materials) is the photoelectric absorption. The interacting photon transfers its energy to an electron from one of the electron shells of the absorber atom (usually from a deep shell). The resulting photoelectron is ejected with a kinetic energy corresponding to the incident photon energy minus the binding energy of the electron on its shell. This is followed by a rapid reorganization of the electron cloud to fill the electron vacancy, which results in the emission of characteristic X-Rays or Auger electrons. The photoelectric process has the highest probability when the incident photon has an energy comparable to the kinetic energy of the electron on its shell. This is the origin of the typical peaks observed in the cross-section curve corresponding to resonances for the different electron shells (Fig. 3.1). The general trend of this cross-section is a rapid decrease with energy and a strong dependence on the atomic number Z of the absorber explaining the preponderance of high-Z materials for X- or γ-rays detection and shielding:

$$\sigma_{\text{ph}} \propto \frac{Z^5}{E_\gamma^{7/2}} \tag{3.1}$$

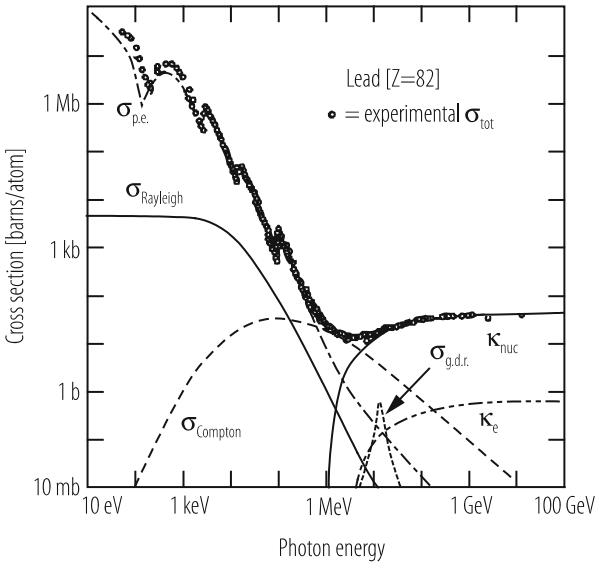


Fig. 3.1 Energy dependence of photon total cross sections in Lead (from Particle Data Group)

At energies above a few hundred keV, Compton scattering becomes predominant. In this case, the incident photon transfers only part of its initial energy E_γ to an electron of the atomic shells and is scattered at an angle θ with respect to its original direction. The recoil electron is then rapidly absorbed by the scintillator and releases an energy according to the formula:

$$E_\mathrm{e} = E_\gamma - E'_\gamma - E_\mathrm{ebinding} \tag{3.2}$$

where E'_γ is the energy of the scattered photon given by (with m_0 the rest mass of the electron):

$$E'_\gamma = \frac{E_\gamma}{1 + \frac{E_\gamma}{m_0 c^2}(1 - \cos\theta)} \tag{3.3}$$

The energy released in the scintillator by the recoil electron is distributed on a continuum between zero and a maximum up to $E_\gamma - m_0c^2/2 = E_\gamma - 256$ keV (for gamma energy large compared to the rest mass of the electron).

The probability of Compton scattering is related to the electron density in the medium and increases linearly with the atomic number of the absorber, favouring therefore high Z materials.

Above a threshold of 1.02 MeV (twice the rest mass of the electron), the mechanism of e^+e^- pair production can take place, predominantly in the electric field of the nuclei, and to a lesser extent in the electric field of the electron cloud (respectively κ_nuc and κ_e in Fig. 3.1). Similarly to photo-absorption and Compton scattering this process has a higher probability for high Z materials as the cross section is approximately given by the formula [4]:

$$\sigma_\mathrm{pair} \propto Z^2 \ln\left(2E_\gamma\right) \tag{3.4}$$

Below the threshold of electron-positron pair production electrons will continue to loose energy mainly through Coulomb scattering.

In the case of an ordered material like a crystal another mechanism takes place at this stage. In the process of energy degradation the electrons in the keV range start to couple with the electrons of the atoms of the lattice and excite the electrons from the occupied valence or core bands to different levels in the conduction band. Each of these interactions results in an electron-hole pair formation. If the energy of the electron is high enough to reach the ionization threshold free carriers are produced, which will move randomly in the crystal until they are trapped by a defect or recombine on a luminescent centre. In the case the ionization threshold is not reached the electron and hole release part of their energy by coupling to the lattice vibration modes until they reach the top of the valence band for the hole and the bottom of the conduction band for the electron. They can also be bound and form an exciton whose energy is in general slightly smaller than the bandgap between the valence and the conduction bands. At this stage the probability is maximum

for their relaxation on luminescent centres through an energy or a charge transfer mechanism.

For a material to be a scintillator it must contain luminescent centres. They are either extrinsic, generally doping ions, or intrinsic i.e. molecular systems of the lattice or of defects of the lattice, which possess a radiative transition between an excited and a lower energy state. Moreover, the energy levels involved in the radiative transition must be smaller than the forbidden energy bandgap, in order to avoid re-absorption of the emitted light or photo-ionization of the centre.

In a way, a scintillator can be considered as a wavelength shifter. It converts the energy (or wavelength) of an incident particle or energetic photon (UV, X-ray or gamma-ray) into a number of photons of much lower energy (or longer wavelength) in the visible or near visible range, which can be detected by photomultipliers, photodiodes or avalanche photodiodes.

3.1.2 Important Scintillator Properties

Scintillators are among the most popular ionizing radiation detectors.

There are two main classes of scintillators: inorganic and organic. For the inorganic systems (generally ionic crystals), scintillation arises from thermalized electrons and holes, moved to the bottom of the conduction band or the top of the valence band respectively, by scattering from the initially produced fast charge carriers. For the organic systems, scintillation arises upon transition between an excited molecular level and the corresponding electronic ground state. Inorganic scintillators are generally brighter but with a slower decay time than organic ones. However no "ideal" material exists and the choice of a scintillator depends on the application, as it is generally driven by a trade-off between a number of physico-chemical and optical parameters such as density, scintillation properties and radiation hardness. The production and processing cost is also an important issue taking into consideration the very large volumes required for some applications.

3.1.2.1 Physico-chemical Properties

Physico-chemical properties are related to the material composition, structure and density, as well as to its chemical stability when exposed to different environmental conditions: air, humidity, ionizing radiation.

Frequently the density and hence the compactness of the detector is essential in order to reduce the detector volume and cost. This is achieved by using high stopping power and therefore high density materials. This reduces the size of the shower for high energy γ's and electrons as well as the range of Compton scattered photons for lower energy γ-rays. A dense material also reduces the lateral spread of the shower, which is particularly important for the majority of High Energy Physics detectors.

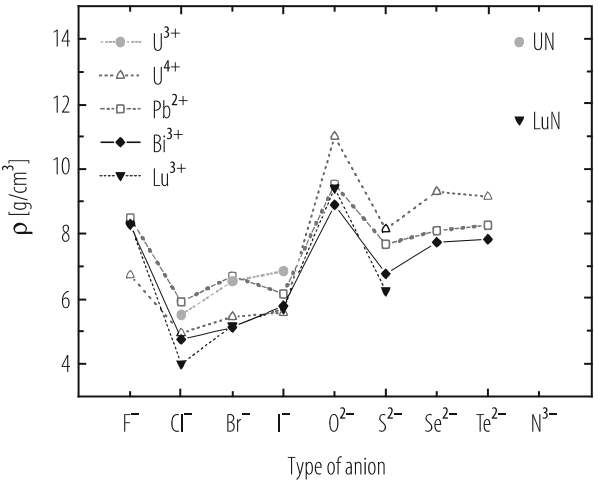


Fig. 3.2 Density for various binary compounds as a function of the binding anion (courtesy P. Derenbos, from ref. [5])

Crystals with a density higher than 8 g/cm^3 are currently available, such as Lead Tungstate (PWO: 8.28 g/cm^3) or Lutetium Aluminium Perovskite (LuAP: 8.34 g/cm^3). Materials of even higher density in the range of 10 g/cm^3 are being identified and studied, such as: Lutetium Oxyde: Lu_2O_3, Lutetium Hafnate: $Lu_4Hf_3O_{12}$, Lutetium Tantalate: Lu_3TaO_7, Lutetium Lead Tantalate: $LuPb_2TaO_6$, Thorium Oxyde: ThO_2. Scintillators are wide bandgap ionic materials and high density implies the choice of anions and cations of high atomic number A (and therefore high Z), as well as small ionic radius to increase the ionic density in the crystal lattice. From this point of view, oxides are generally denser than iodides because of the much smaller ionic radius of the oxygen compared to the iodine ion and in spite of its lighter weight. Similarly, the oxidation potential of the anion is important as it allows reducing the number of anions (generally light) needed to compensate for the positive charge of the much heavier cation. For this reason oxygen is a better ligand than the slightly heavier fluorine ion because of its higher oxidation state (2 or 3 instead of 1). Figure 3.2 illustrates this effect for a number of binary compounds as a function of the anion type.

High Z materials are also preferred for low and medium energy spectroscopy because of the strong dependence of the photoelectric cross-section on Z (see Sect. 3.1.1). High density is also required at high energy to achieve a small radiation length X_0 (mean distance over which an electron loses 1/e of its energy) given as a function of the density ρ, atomic mass A and atomic number Z by:

$$X_0 = \frac{A}{\rho} \frac{716.4 \text{gcm}^{-2}}{Z\,(Z+1)\ln(287/Z)} \tag{3.5}$$

However, contrary to a common assumption, the optimum conditions are not necessarily achieved with the highest Z ions, because in addition to a small X_0, the density ρ should be high. This reduces the lateral shower size given by the Moliere radius:

$$R_{\mathrm{M}} \approx X_0 \cdot (Z + 1.2)\,/37.74 \sim 1/\rho \tag{3.6}$$

The stability of the physico-chemical parameters is also important for the detector design. Scintillation crystals are very stable materials, at least in the bulk, if grown under conditions allowing a good structural quality. This provides a high degree of internal symmetry in the material together with high energetic stability. However, the charge unbalance on the surface can be at the origin of different problems, such as a concentration of impurities or crystallographic defects. As a result, the material can interact with its environment and locally change its properties. The majority of halide crystals have the anions weakly bound to the cations at the surface. They are therefore easily replaced by OH^- radicals from the atmosphere, which have strong optical absorption bands in the visible spectrum. This causes a progressive brownish discoloration of the crystal surface, a well know feature of hygroscopic materials. Encapsulating the crystal in an inert atmosphere avoids this effect.

3.1.2.2 Optical Properties

Inorganic scintillators usually show wide emission bands because of multi-site emission centres differently distorted by the crystal field, as well as by temperature broadening of the optical transitions through vibronic coupling of the emission centres with the crystal lattice. These emission bands are situated in the optical window of the scintillator and produce light in the visible, near infrared or near ultraviolet part of the spectrum. One of the objectives of scintillator development is to design scintillators with emissions peaks matching the maximum quantum efficiency of photodetectors, typically 250–500 nm for photomultipliers and 450–900 nm for solid state photodetectors (pin diodes and avalanche photodiodes).

Light yield (LY) is an essential parameter for a scintillator as it directly influences the energy resolution at low or medium energy through the photostatistic term proportional to $(LY)^{-1/2}$ and the timing resolution proportional to $(\tau_{\mathrm{sc}}/LY)^{-1/2}$, with τ_{sc} being the scintillation decay time. The scintillation mechanism is a multi-step process, which will be described in detail in Sect. 3.2. The overall scintillation yield is determined by the product of efficiencies for all these steps. The dominant factor, which sets the fundamental limit on the light output of a given scintillator, is the number n_{eh} of thermalized electron-hole pairs (active for scintillation) produced in the ionization track of the incoming particle:

$$n_{\mathrm{eh}} = \frac{E_\alpha}{\beta \cdot E_{\mathrm{g}}} \tag{3.7}$$

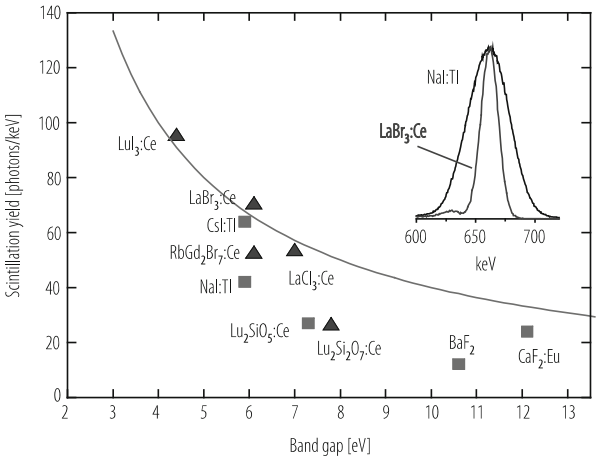


Fig. 3.3 Photon yield/keV of several scintillators as a function of the width of the forbidden band (courtesy P. Dorenbos)

where $\beta \cdot E_g$ is the mean energy necessary for the formation of one thermalized electron-hole pair in a medium with a forbidden zone of width E_g and E_α is the absorbed energy. For ionic crystals, the factor β is usually close to 2.3 and takes into account the energy loss through coupling with lattice phonons during the thermalization process [5]. As shown on Fig. 3.3 low bandgap materials have higher scintillation yields, although such materials are potentially more subject to trap induced quenching, re-absorption phenomena and photo-ionization of the luminescence centre. The ultimate light yield obtained for a material having a bandgap of 3 eV and an emission wavelength of about 600 nm is in the range of 140 photons/keV. The observed signal in photoelectrons/MeV is much smaller, due to losses in the light transport to the photodetector and the quantum efficiency of the photodetector.

The scintillation kinetics is another important consideration as a fast response and low dead time is frequently required for high detection rates. It is related to the rate of decrease of the population of the excited luminescent centres. For a simple process, with only one radiating centre and no interaction between luminescent centres and traps, the decay is exponential and characterized by a time constant τ_{sc}, the time after which the population has decreased by a factor e. For two independent radiating centres the same description with two exponentials holds. Real cases are however very often more complex, involving energy transfer between centres and quenching mechanisms, and the resulting light emission is strongly non-exponential. It is nevertheless common practice to describe this complex emission curve by a sum of exponentials with different time constants. This has in most of the cases no physical justification but simplifies the calculations. If we assume a very fast transfer of the electrons and holes to the luminescent centres the ultimate

limit for the scintillation decay time is given by the transition probability between its excited and ground states:

$$\Gamma = \frac{1}{\tau_{sc}} \propto \frac{n}{\lambda_{em}^3}\left(\frac{n^2+2}{3}\right)^2 \sum_f |\langle f|\ \mu\ |i\rangle|^2 \tag{3.8}$$

where n is the refractive index of the crystal, λ_{em} the emission wavelength of the transition, f and i the wave functions of the final and initial states respectively. The strength of the dipole operator μ connecting the initial and final state determines the decay time of the transition. This matrix element can only be sufficiently large for a transition between two states with different parity (parity allowed transition). This is in particular the case for the 5d to 4f transition in commonly used activators like Ce^{3+}, Pr^{3+}, Nd^{3+} and Eu^{3+}. Forbidden transitions are generally characterized by long decay times, unless a competitive non-radiative relaxation channel exists, which will contribute to the decrease of the population of excited states:

$$\frac{dn_e}{dt} = -\frac{n_e}{\tau} - \alpha n_e e^{-\frac{E}{kT}} \tag{3.9}$$

Here n_e represents the electronic density of the excited state, which is depopulated through two competing decay channels, the first one radiative with a rate $1/\tau$ and the second one, non-radiative, through a thermal quenching mechanism. E is the thermal energy barrier and α expresses the balance between the two channels. Fast scintillation can therefore be obtained for intrinsically slow transitions at the expense of a loss in light output. This is the case of Lead Tungstate (PWO) with a low light yield but 10 ns decay time at room temperature to be compared to a 25 times larger light yield but 6 μs decay time at 80°K [6]). More details about thermal quenching will be given in Sect. 3.2.

Special attention must be given to afterglow, which limits the counting rate of scintillation detectors. Afterglow is a phosphorescence mechanism induced by the thermal release of charge carriers from traps. These carriers will eventually recombine on luminescence centres, causing a delayed luminescence, which can reach several percent after 1 ms for NaI(Tl) or CsI(Tl). Other crystals have a much lower level of afterglow, such as BGO (Bismuth Germanate): 0.005% after 3 ms, and CsF (Cesium Fluoride): 0.003% after 6 ms [7].

3.1.2.3 Radiation Hardness

Inorganic scintillators have in general a good stability of their scintillation properties even in the presence of intense ionizing radiation environment. This property is crucially important for detectors in space, oil well logging and high-energy physics experiments at high luminosity accelerators. The radiation hardness of the scintillation mechanism is related to the strong electrostatic field of the crystal

lattice, which shields the luminescent centres. However, the transport of light through the crystal may be affected by the production of colour centres, which absorb part of the scintillation light on its way to the photodetector. The formation of colour centres results from the trapping of electric charges by crystal structural defects or impurities and is therefore directly correlated to the quality of the raw material and of the growth technology. A large effort is needed to purify the raw materials to the required quality and to minimize the amount of structural defects during the crystal growth. However, in some cases, a specific doping of the crystal has proven to be an efficient and economical way of significantly increasing the radiation hardness [8].

3.1.3 Scintillator Requirements for Various Applications

The choice of a scintillator depends on the energy of the ionizing radiation to be detected and on constraints specific to the application. It is therefore tailored to the user requirements considering the relative importance of several parameters, such as density, light yield, scintillation kinetics, emission spectrum, radiation hardness. Ruggedness, hygroscopic behaviour and production cost are also important parameters. In practice, it is impossible to find a scintillator, which combines all the most desirable properties. Besides a number of industrial applications for process control, container inspection, thickness gauging, ore processing and oil well logging a large fraction of the scintillator market is driven by X-ray and γ-ray spectroscopy in the following areas:

- High and medium energy physics particle detectors;
- Astrophysics and space applications;
- Spectrometry of low energy γ-quanta;
- Medical imaging;
- Safety Systems and Homeland Security.

The most important user requirements for each of these categories are detailed below.

3.1.3.1 High and Medium Energy Physics Particle Detectors

Scintillators are used in High Energy Physics for compact, high precision, homogeneous electromagnetic calorimetry. The purpose is to measure with the highest achievable precision the energy of electrons and photons, generally the decay products of unstable heavier particles, over the widest possible energy range.

The first important requirement is a high density material. High energy implies a high particle multiplicity of the particle collisions and requires a high granularity with good lateral containment of the particle initiated showers in order to minimize overlapping showers and to ease event reconstruction. A small Moliere radius is

therefore required, which will also improve the electron identification and allow π^0 rejection with good efficiency in high multiplicity events. More generally, a high stopping power is mandatory to longitudinally contain high energy showers in a reasonable volume and cost (typically 20–25 X_0 are needed in high energy calorimeters to contain at least 95% of the shower). Total lateral and longitudinal containment of the showers is a prerequisite to minimize leakage fluctuations and to achieve good energy resolution.

Fast scintillation is also an important parameter. In the search for rare events, and at hadron colliders, one operates at high collision rates, which requires a short time response of the detectors. Decay times of the order of the bunch crossing time (typically 25 ns) or even less are necessary. Only optically allowed (inter-configuration) transitions (like the transition 5d $\rightarrow$ 4f for Ce^{3+}), cross-luminescence, which is intrinsically fast and temperature independent as observed in Barium Fluoride (BaF_2), and strongly quenched intrinsic luminescence (as for PWO) can give rise to a fast light signal.

The demand for a high light yield is less stringent at high energy (GeV range) than at low energy (MeV range), because of the high number of scintillation photons produced even by a poor scintillator, allowing a good signal detection above the electronic noise. Such low light yield scintillators can therefore also be used for calorimetric applications in magnetic spectrometers due the rapid development of silicon photomultipliers (SiPM), with a gain comparable to photomultiplier tubes (PMT) and with the additional advantages of being very compact and immune to strong magnetic fields.

However, the high track density and event pile-up at high luminosity colliders pose serious challenges for physics event reconstruction and analysis. At the Large Hadron Collider (LHC) at CERN up to 40 pile-up events and more can be produced at each bunch crossing at the design luminosity of 2.10^{34} cm^{-2} s^{-1}, which will reach 200 pile-up events when the luminosity will be increased to 10^{35} cm^{-2} s^{-1} at the High Luminosity LHC [9]. For a collision region of about 10 cm (bunch length) the collisions will be distributed over 300 ps (Fig. 3.4 left panel). Precise temporal association of collision tracks or jets would help mitigate the pile-up. If this can be done for charged particles at high transverse momentum with particle tracking detectors this approach will be much more difficult in the forward-backward region and even impossible for neutral particles. In this case only time-of-flight (TOF) techniques can be applied as shown on the right panel of Fig. 3.4, where the two crossing bunches are symbolized by blue and red bars while their overlapping area is represented by a white bar. Events generated in the centre of the detector ($z = 0$) will generate tracks arriving at the same time in the forward and backward regions. On the other hand, events generated at any time off-centre of the bunch-overlapping region will exhibit a TOF difference for the tracks generated in the forward and backward regions, as shown on Fig. 3.4 (t_2–t_4, t_5–t_7, t_8–t_{10}). A mitigation factor of one order of magnitude necessitates a TOF precision of at least 30 ps [9].

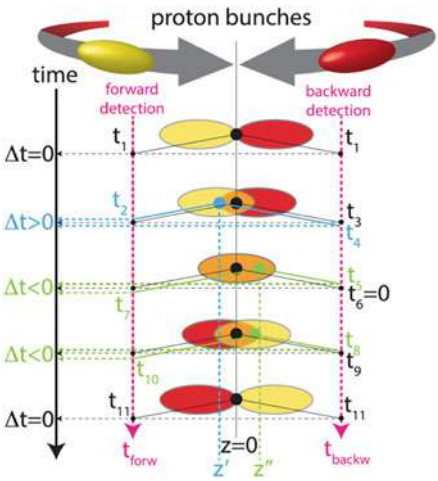


Fig. 3.4 Schematics of bunch crossing and TOF in the forward and backward directions of particles generated by events created in different positions of the overlap region

Excellent timing resolution is therefore needed. It can be shown [10] that it is related to the time density of the detected scintillation photons in the leading edge of the scintillation pulse, which is given by the following formula:

$$\sigma_t \approx \sqrt{\tau_r \tau_d \Big/ N_{pe}}$$

where τ_r and τ_d are the scintillator rise time and decay time respectively and N_{pe} is the number of photoelectrons produced in the photodetector, which is proportional to the light yield of the scintillator. A high light yield is therefore mandatory to minimize the photo-statistic fluctuation influencing the time jitter of the detector. An emission spectrum in the visible region is preferred as the quantum efficiency of the majority of photodetectors is higher and the light is generally less attenuated than in the UV region and hence more easily collected.

The energy resolution of the calorimeter is affected by all possible sources of non-uniformity. The light collection in a pointing geometry of tapered crystals introduces non-uniformity due to a focusing effect through the successive reflections of the light on the lateral faces, which depends on the refractive index of the crystal. Fluoride crystals and glasses, with low refractive index (around 1.5) have smaller non-uniformities (and therefore are easier to correct) than BGO (index 2.15) or PWO (index 2.3). The material can be intrinsically luminescent if it holds luminescent molecular complexes or ions, or is doped with a scintillating activator. Intrinsic scintillators are generally preferred, as it is easier to control the light yield uniformity in long crystals. On the other hand, a controlled distribution of the doping could help correcting for the non-uniformity caused by the light collection in a pointing geometry. Furthermore, the scintillation yield should be as independent as possible from temperature. Large temperature coefficients increase the complexity

of the detector design and of the software corrections, and temperature gradients between the front and back face of the crystals introduce non-uniformity affecting the resolution.

Finally, for large scintillator volumes cost considerations are of importance. The abundance of the raw materials, the facility to purify them against the most detrimental impurities to achieve good radiation hardness, a low temperature melting point to save on the energy cost, a high growing and mechanical processing yield are all parameters, which deserve particular attention.

3.1.3.2 Astrophysics and Space

Increasingly crystal-based calorimeters are embarked on satellites to study galactic and extra-galactic X- and γ-ray sources. This requires excellent energy resolution over a wide energy spectrum, typically from a few KeV to several TeV (see for instance Fig. 2.16 of ref. [11] for a list of different space missions with their respective energy range). One major aim of these measurements is the determination of the direction of the γ-ray source. Two classes of position sensitive devices have been developed in the last decades. These designs are using continuous scintillation crystal or pixilated detector geometries [12]. The required angular resolution is achieved with multilayer calorimeters or readout schemes to provide depth of interaction (DOI) information or using coded aperture masks.

The low orbit satellites are shielded by the earth magnetic field, relaxing therefore the requirement for radiation hardness of the scintillation material. Most of the scintillation materials can be used depending on the energy range of the detected γ-radiation. However, the payload is limiting the size of such detectors and not too dense materials are sometimes selected to reduce the weight.

In the interplanetary space the sun wind from charged particles strongly influences the detecting requirements of the scintillation materials. For these missions, high radiation hardness to ionizing radiation and low level of induced radioactivity are required. The same applies to detectors for planetary missions.

The general trend is to select high light yield, fast and not necessarily ultra-dense scintillators such as CsI or YAP. The very bright $LaBr_3$ is likely to find some applications in this domain because of its excellent low energy resolution (comparable to solid state detectors). BGO is very often used in veto counters for the rejection of Compton events.

3.1.3.3 Spectrometry of Low Energy γ-Quanta

This is probably the most important application domain for inorganic scintillators. The key requirement concerns energy resolution on the photopeak. It is therefore essential to maximize the photofraction and high Z materials are clearly preferred (see Sect. 3.1.1).

The energy resolution is driven by several factors and a detailed discussion is given in Sect. 3.1.1. However, two important parameters are playing an essential role. The first one is the light yield. One contribution to the energy resolution is the statistical fluctuation of the number of photoelectrons, n_{pe}, produced in the photodetector. Therefore a high light yield will reduce this statistical contribution like $(n_{pe})^{-1/2}$.

The second parameter concerns the deviations from the linearity of response at low energy. Most crystals exhibit a non-proportionality behaviour for energies below 100 keV. The relative light yield can show either relative increase with decreasing energy, as is the case for halide crystals, or a decrease, as for the majority of oxides and fluorides. Only few crystals have an almost linear response down to about 10 keV, such as $YAlO_3$ (YAP), $LuAlO_3$ (LuAP), $LuYAlO_3$ (LuYAP), $LaBr_3$. Given that the energy loss mechanisms—photoelectric, Compton scattering and pair production—are energy dependent, the total energy deposit in a crystal detector will be a mix of these contributions varying with energies. The non-linearity affects therefore the energy resolution, as is illustrated by the examples of Lutetium orthosilicate (LSO) and Lutetium Aluminium Perovskite (LuYAP). For the same detector volume, LuYAP achieves similar energy resolution (9%@511KeV) as LSO despite a three times lower light yield [13], as a result of a more linear response at low energy, as shown on Fig. 3.5.

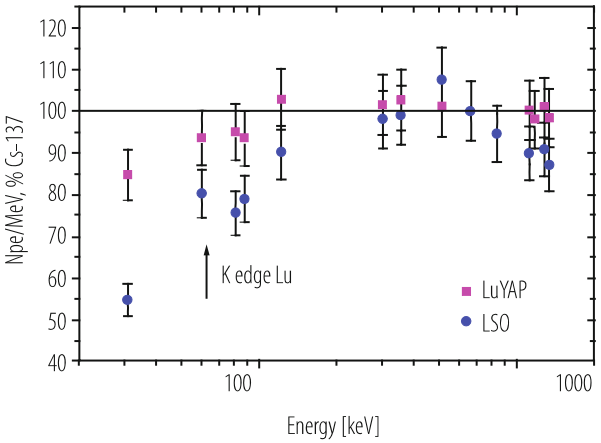


Fig. 3.5 Relative low energy response for LSO and LuYAP crystals, normalized to the ^{137}Cs energy peak (from ref [13])

3.1.3.4 Medical Imaging

Scintillators are widely used in medical imaging for X-ray radiology (digital radiography and CT scanners) and for emission tomography (PET and SPECT) with a market exceeding several hundred tons per year (see Sect. 20.1).

The choice of the scintillator for medical imaging devices is determined by the stopping power for the energy range of X and γ-rays to be considered, or more precisely the conversion efficiency. Materials with high Z and high density are favoured but the energy of the K-edge is also important as can be seen in Fig. 3.6. For low energy X-ray imaging (below 63 keV) the attenuation coefficient of Yttrium, Cesium and Iodine are quite high and crystals like YAP and CsI are good candidates for soft tissue X-ray imaging like mammography. Above the K-edge of Lu (63 keV) and Bismuth (90 keV) the situation is quite different and BGO and Lutetium based crystals are favored for bone, dental X-ray, ^{99}Tc (90 keV) SPECT and PET scanners (511 keV). Heavy scintillators have smaller thickness, reducing parallax errors in ring imagers and maintaining a good spatial resolution over the whole field of view (Sect. 7.1).

A high light yield is also mandatory for good energy resolution. Better energy resolution increases rejection of Compton events, improves the spatial resolution and the sensitivity. The sensitivity is a critical parameter as it determines the number of useful events per unit of injected dose. A higher sensitivity means a smaller injected dose or a better image contrast.

A short scintillation decay time reduces the dead time and therefore increases the maximum counting rate. In PET scanners for instance reducing the coincidence

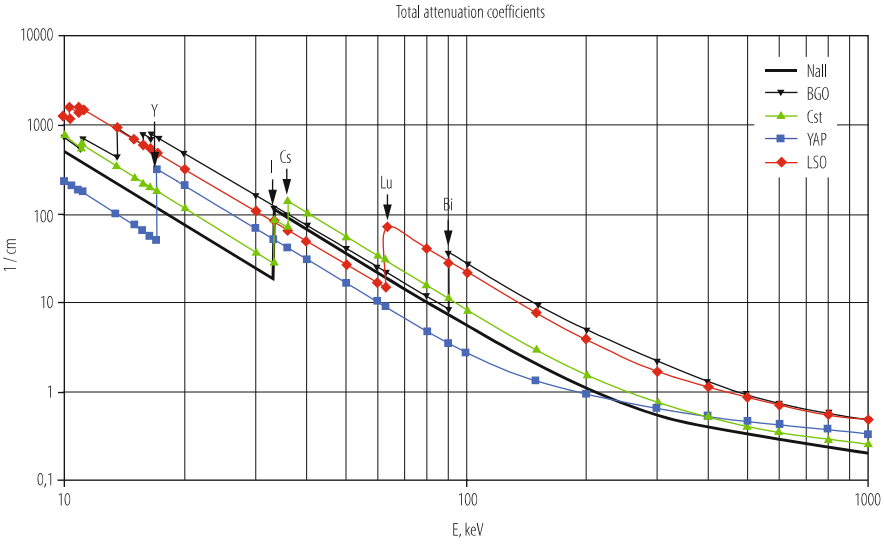


Fig. 3.6 Attenuation coefficients in several high Z materials

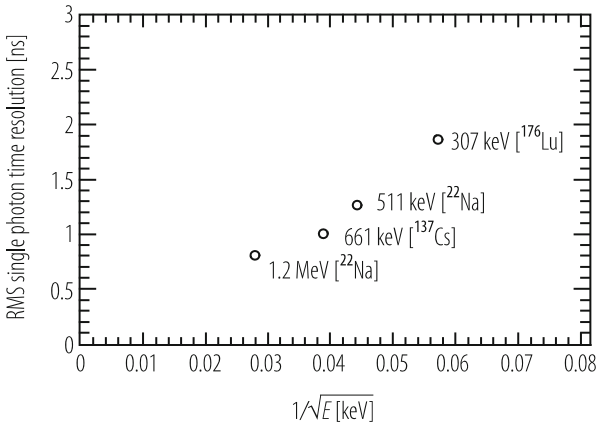


Fig. 3.7 Energy dependence of the timing resolution of a ClearPEM $2 \times 2 \times 20$ mm^3 LSO pixel coupled to an Hamamatsu avalanche photodiode (courtesy J. Varela)

window improves the signal to background ratio and increases the sensitivity and image contrast. Very fast scintillators open the way to scanners using the time-of-flight information, which helps reducing the background by selecting a narrow region of interest along the coincidence line. In the range of energies considered for medical imaging, the timing resolution is limited by the Poisson distribution of photons arrival time on the photodetector, even for bright scintillators like LSO. Figure 3.7 shows the $1/\sqrt{E}$ dependence of the timing resolution of a ClearPEM [14] detector head made of $2 \times 2 \times 20$ mm^3 LSO pixels coupled to a 32-channel Hamamatsu APD matrix, when excited by sources at different energies E.

Commercial PET scanners achieve about 500 ps FWHM coincidence time resolution (CTR) in the difference of detection time of the two 511 KeV gamma rays resulting from the positron annihilation. This allows a significant image quality improvement particularly for over-weighted patients. Ideally, one would like to achieve 100 ps FWHM CTR resolution, which would correspond to a centimetre resolution along the line of response (LOR) corresponding to the coincidence detection of the two gamma rays. It improves by an additional factor 5 the image signal-to-noise ratio. Thus a TOF-PET system with 100 ps CTR can either give a five times shorter examination time of the patient or a five times lower radiation dose at constant image quality.

As mentioned in Sect. 3.1.2.2, in first approximation (assuming single photon detection) the CTR for a scintillators with a scintillator rise time τ_r and a decay time $\tau_{d,}$ is given by:

$$CTR \propto \sqrt{\frac{\tau_r \tau_d}{N_{phe}}}$$

where N_{phe} is the number of photoelectrons readout from the crystal. Clearly, there is a premium for a high photon rate in the leading edge of the scintillation pulse, a high light yield as well as a short rise and decay times for improving the CTR.

3.1.3.5 Safety Systems and Homeland Security

Scintillators are used in three main types of equipment related to safety and homeland security: express control of luggage and passengers, search for explosive materials and remote detection of fissile materials.

Luggage inspection requires the highest possible throughput to quickly identify a suspect luggage in a few cubic meter large container moving across the inspection device. The spatial resolution is determined by the need to quickly localize and identify the suspect object in a large container. Fast scintillation kinetics with no afterglow is therefore the most important parameter.

For the remote detection of explosives the most attractive methods are based on the detection of natural or induced characteristic neutron and γ-rays under activation by a neutron source, either with fast neutrons from the ^{252}Cf radioisotope or fast-thermal neutrons from a pulsed electronic neutron generator. Neutrons initiate nuclear reactions in some elements, some of them producing characteristic γ-rays. Plastic explosives for instance are generally rich in nitrogen. The nitrogen (n,γ) reaction has a cross section of 75 mb and produces a characteristic γ-ray of 10.83 MeV.

For such applications, the most important scintillation crystal parameters are: high stopping power to improve the detector sensitivity; high light yield to improve the detector energy selectivity; fast scintillation decay time to allow time-of-flight analysis with pulsed neutron generators to increase the signal to noise ratio. Good stability of the scintillator parameters under ionizing and neutron irradiation allows the use of strong activation sources for a better sensitivity.

Remote detection and fissile materials warhead inspection has been for a long time restricted to the detection of neutrons, as the γ-channel would have easily revealed secret characteristics of the nuclear device. This has changed recently and opens new possibilities to detect the radiation emitted by Nuclear Explosive Devices (NED) based on enriched uranium or plutonium. The most useful energy range to detect fissile material is $E_\gamma \geq 3$ MeV because of (1) the absence in this range of natural radioactive sources and therefore an acceptable signal to background ratio; (2) the high penetration power of these energetic γ-quanta making the deliberate concealment of the intrinsic NED radiation more difficult.

Here, the most important parameters are sensitivity to allow detection at large distance (at least several meters) and good background rejection. High stopping power (and therefore high density) is mandatory. However, the crystals should be made from materials with very low natural radioactivity, which restricts the choice of heavy materials to the ones with no unstable isotopes. As the counting rates are usually low, there is no need for ultra-fast scintillators. A phoswich geometry based on two different crystals on top of each other can be an attractive solution for

improved low energy background rejection. A first thin scintillator layer detects (and rejects) the low energy background activity, whereas a thicker layer on the back will be mainly sensitive to the 3–10 MeV range of interest. The two scintillators must have different emission wavelength and/or decay time for a good identification of the hit source.

3.1.4 Organic Material, Glass and Condensed Gases

There is a particular class of scintillators, which does not require a regular lattice to produce scintillation light when excited by ionizing radiation. These are organic solid and liquid materials, condensed gases as well as scintillating glasses. A common feature of all these materials is that scintillation (also called fluorescence in this case) results from a direct excitation of a molecule and does not involve the transport of the excitation energy through the material. As the molecule is directly excited and the coupling with the host material is minimal, the fluorescence decay time is solely determined by the quantum numbers of the excited and ground states. If properly chosen the molecule will emit between two singlet states giving rise to a fast emission (usually not more than a few ns).

Different material combinations can be engineered, in particular in plastic scintillators, to meet specific requirements. The most popular one concerns wavelength shifters. Binary or even ternary solutions of different fluors can be dissolved in a plastic base containing aromatic molecules. After excitation by ionizing radiation, these aromatic rings will relax the stored energy by emitting UV photons. Properly chosen additional fluors can absorb these photons and reemit them at longer wavelength, e.g. to better match the quantum efficiency of a photodetector. As there are only energy transfer and no charge transfer mechanisms involved, the whole process is very fast.

Plastic scintillators can be easily machined in any shape, including in the form of fibres, one important advantage. However, these materials are intrinsically light (density around 1–1.2 g/cm^3) and therefore are not suitable for homogeneous calorimetry. They find a number of applications in sampling calorimetry and tracking. More information can be found in ref. [15].

3.2 Scintillation and Quenching Mechanisms in Inorganic Scintillators

3.2.1 The Five Steps in Scintillation Process

In contrast to luminescence (such as in lasers), where the excitation source is tuned to the energy levels of the luminescent centres, scintillation is the result of a

complex chain of processes, each of them characterized by a specific time constant and efficiency factors [16]. This is summarized in Fig. 3.8, where the valence and conduction bands of an insulator with a bandgap width E_g (forbidden band) are represented. The upper level core band (energy E_c and bandwidth ΔE_c) is also shown.

The sequence of processes is shown as a function of time and can be qualitatively divided into five main phases:

- The first one is the energy conversion phase and the subsequent production of primary excitations by interaction of ionizing particles with the material. For an incident particle energy in the keV range or higher, the excitations are essentially deep holes h created in inner core bands and hot electrons e in the conduction band. Subsequently, on a very short time scale (10^{-16}–10^{-14} s), a large number of secondary electronic excitations are produced through inelastic electron-electron scattering and Auger processes with creation of electrons in the conduction band and holes in core and valence bands. At the end of this stage, the multiplication of excitations stops. All electrons in the conduction band have an energy smaller than $2E_g$ (e-e scattering threshold) and all holes occupy the valence band if there is no core band lying above the Auger process threshold (general case).
- The second stage is the thermalization of electronic excitations through a phonon coupling mechanism with the crystal lattice, leading to low kinetic energy electrons in the bottom of the conduction band and of holes in the top of the valence band. This thermalization phase takes place in the sub-picosecond range, typically between 10^{-14} and 10^{-12} s.
- The next stage, between 10^{-12} and 10^{-10} s, is characterized by the localization of the excitations through their interaction with stable defects and impurities of the material. For example, electrons and holes can be captured by different traps or self-trapped in the crystal lattice. Excitons, self-trapped excitons, self-trapped holes (V_K centers) can be formed with emission of phonons. Localization of excitations can be sometimes accompanied by displacements of atoms (defect creation, photo-stimulated desorption).
- The transfer of excitations to the luminescent centres through the sequential capture of charge carriers or different energy transfer mechanisms takes place during the following 10^{-10} and 10^{-8} s.
- Finally, the radiative relaxation of the excited luminescent enters produces the light signal with an efficiency and time structure, which is given by the quantum selection rules of the transition. Parity allowed transitions with more than 3 eV energy gaps are generally preferred as they give rise to fast luminescence. However, smaller energy gaps (2–3 eV) are likely to favour higher light yield, as discussed in Sect. 3.1.2.2.

The scheme depicted in Fig. 3.8 describes the scintillation mechanisms in the case of ionic crystals with simple energy structures. However, important groups of scintillators exhibit a more complicated band structure.

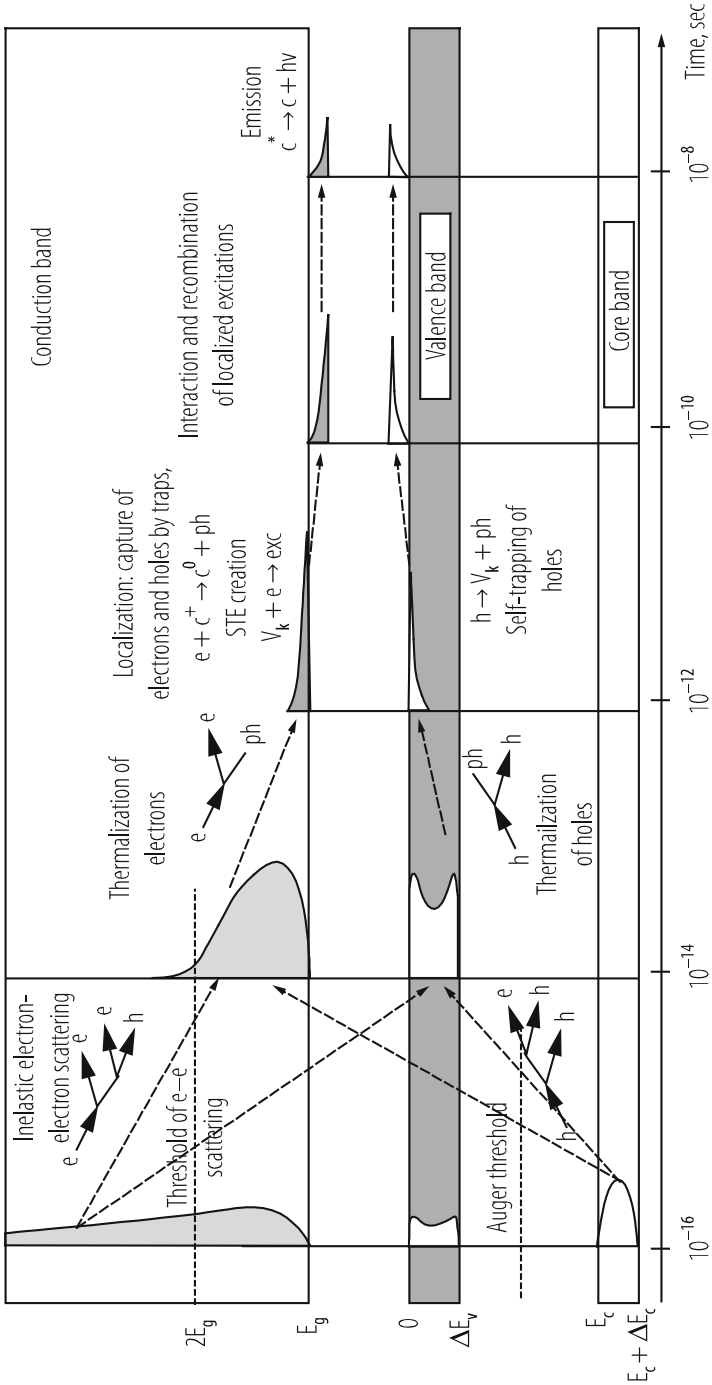


Fig. 3.8 Relaxation scheme for electronic excitations in an insulator: e, electrons; h, holes; ph, phonons; hν, photons; V_k, self-trapped holes; STE, self-trapped excitons; c^n, ionic centers with charge n. Density of states is represented by grey and white areas for electrons and holes respectively (courtesy A. Vasiliev)

One such case are so-called cross-luminescent materials, of which one well-known example is Barium Fluoride (BaF_2). Such systems are characterized by a specific configuration of the energy bands, such that the width of the forbidden gap (between the valence and conduction bands) is larger than the energy gap between the uppermost core band (5pBa in the case of BaF_2) and the bottom of the valence band. When a hole produced in this core band recombines with an electron of the valence band there is not enough energy available to eject an Auger electron from the valence to the conduction band. The core-valence transition can therefore only be radiative giving rise to a scintillation in the UV, which is usually very fast (sub-nanosecond).

3.2.2 Scintillation Efficiency

The overall scintillation efficiency η is generally given by the product of three terms:

$$\eta = \beta \cdot S \cdot Q \tag{3.10}$$

where β represents the conversion efficiency for the production of electron-hole pairs, S the excitation transport efficiency, including thermalization of electric carriers, localization and transfer to the luminescent centre, and Q is is the quantum efficiency of the radiative transition of the luminescent centre. If we consider, as discussed in Sect. 3.1.2.2, the number of 140,000 ph/MeV as an upper limit for the scintillation yield of an ideal scintillator with an emission peak around 600 nm the maximum scintillation efficiency is less than 30%. In reality, for the majority of existing scintillators it is less than 5%, mostly because of important losses during the thermalization and transport process.

At the end of the first phase of inelastic scattering the holes and electrons have reached an energy below the Auger and ionization thresholds respectively. Their thermalization to the top of the valence band for holes and to the bottom of the conduction band for electrons can only take place by heat dissipation through coupling to the phonon modes of the lattice. This is an unavoidable part of energy loss for the scintillation process. The energy gap between these two thresholds being of the order of $2.3E_g$ for ionic crystals one concludes that an ideal scintillator cannot convert more than 43% of the absorbed energy into light.

Another important loss is related to the transfer of the excitations to the luminescent centres. A frequent channel of excitation for acceptors is a charge transfer process with a sequential capture of charge carriers. In Ce^{3+}-doped crystals, the hole is first captured with its capture probability strongly depending on the position of the Ce^{3+} ground level (4f) in the forbidden band gap. In cerium-doped oxides and halides, this level is usually lying very low in the gap close to the top of the valence band, and these systems can lead to very efficient scintillation (LSO, LuAP, $LaCl_3$, etc.). On the other hand, Ce^{3+}-doped fluoride crystals cannot exhibit

very high light yield because the Ce^{3+} 4f is lying around 3–4 eV above the valence band, which strongly reduces the hole capture probability.

It is also important to avoid the delocalization of electrons from the activator excited state to the conduction band. This is achieved if the energy gap ΔE between the radiating level of the doping ion and the bottom of the conduction band is large enough. If $\Delta E >> kT$, or the radiative decay $\tau_\gamma << \tau_d$, where the delocalization time $\tau_d \approx (1/S)\exp(-\Delta E/kT)$, with S—the frequency factor, k—the Boltzman constant, and T—the temperature, the scintillation yield is not strongly dependent on the temperature. In the reverse case one can expect a reduction of the scintillation yield when the temperature increases (temperature quenching). Similarly, when the ground state is located in or very close to the valence band, the hole is weakly trapped and can be easily delocalized to the valence band.

Besides these different processes, a number of competing channels can limit the probability of charge carrier capture by the luminescent centres. Impurities or ions in the lattice can act as specific killer ions and compete with active ions for the capture of charge carriers and/or interact with them, inducing severe limitations in scintillation efficiency. For example, in cerium-doped crystals the presence of ions or molecular groups with two or more stable valence states is generally to be avoided. This is due to the fact that cerium has two stable valence states, Ce^{3+} and Ce^{4+}, but Ce^{3+} only gives rise to luminescence. If a possibility exists for Ce^{3+} to transfer one electron to these killers it will transform into Ce^{4+} and no longer scintillate. This is the case for Ce-doped tungstates and vanadates, which do not exhibit cerium scintillation because of such Ce-W and Ce-V interactions. For the same reason the good electron acceptor Yb^{3+} severely quenches the Ce^{3+} scintillation.

Self-trapping is also a very frequent source of efficiency loss in insulating materials. Indeed, some of the electrons and holes can be trapped by impurity or crystal defect related acceptors and cannot excite directly luminescent centres through sequential capture. If the trap is very shallow it will quickly release the charge carriers and will slightly delay scintillation. However, in deep traps strong quenching of the fast luminescence components is observed. Very long components in the fluorescence decay appear when the temperature is raised to the point, where trapped electrons can be released by thermal energy (glow peaks).

The interaction between closely spaced electronic excitations (in a few nanometre range) may lead to luminescence quenching, also-called local density-induced quenching. For electronic excitations created through the different mechanisms of photon absorption, the probability to produce excitations at such short distances is very low if the excitation source has a limited intensity. On the contrary, secondary electronic excitations created by inelastic scattering of photoelectrons or Auger decay of core holes can be quite closely spaced. In these clusters of high local e and h density, the interaction between excitations can modify their localization and can even create defects in crystals. In addition, these clusters can excite closely spaced luminescent centres, which can interact with each-others, giving rise to faster and

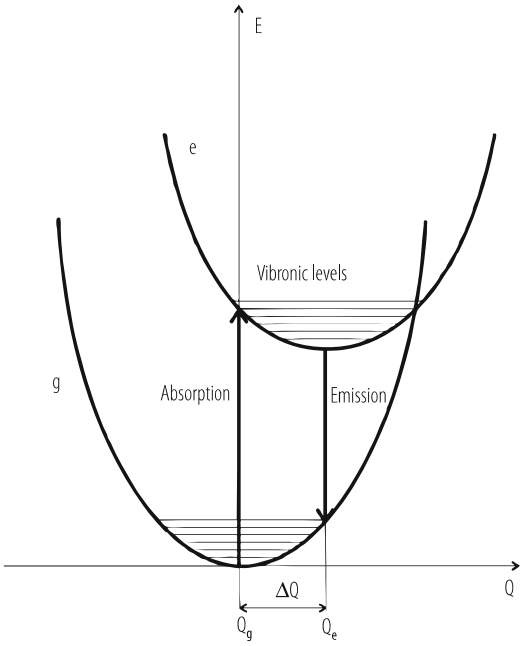


Fig. 3.9 The configurational coordinate diagram. The energy E is plotted versus the coordinate Q (configurational coordinate in the lattice). The ground state g and one excited state e are represented by potential curves with offset ΔQ. Absorption and emission transitions are indicated

non-exponential decay time and total or partial luminescence quenching. The first evidence of such effect was observed in CeF_3 [17].

Another type of thermal quenching can occur related to electron-phonon coupling. The different electronic configuration of the ground and excited states of the activator generally induces an exchange of phonons and the relaxation of the position of the activator ion when it is excited. As a result, the emission transition from the relaxed excited state is shifted towards lower energy than the absorption transition. This is the well-known Stokes shift illustrated in Fig. 3.9. The Stokes shift is a measure of the interaction between the emitting centre and the vibrating lattice. The stronger the electron-phonon coupling the larger the Stokes shift. For weak coupling, the potential curves are not significantly shifted and the emission spectra show narrow lines (case of f-f transitions of rare earth ions). In the case of intermediate coupling for which the parabolas are weakly shifted, vibronic spectra of broad emission lines are observed reflecting the progression in stretching vibration of the luminescent ion (case of uranyl pseudo-molecules in oxides, like UO_2^{2+}).

In the case of strong coupling (shown in Fig. 3.8) the relaxed excited state may decay non-radiatively to the ground state if the temperature is high enough to allow the excitation to reach the crossing of the two parabolas.

In practice, the relevant parameter is the light yield efficiency Y, which is the product of the scintillation yield η by the light transport and collection efficiency η_{col} to the photodetector. A number of parameters influence η_{col}: the crystal shape, its optical transparency to the scintillation wavelength, the presence of scatters and

different optical defects in the bulk of the crystal, the surface state and wrapping conditions of the faces of the crystal, the coupling face to the photodetector, the surface matching between the coupling face and the photodetector, the crystal index of refraction. Heavy scintillators generally have a high index of refraction (larger than 2 in many cases) and the light collection efficiency is limited to 10–30% for the majority of existing detectors. New approaches based on nanostructured surfaces, in particular photonic crystals, are presently being explored [18]. Significant light extraction gains of more than 50% have been obtained as well as a strong reduction of the photon transit time spread in the crystal associated to the higher extraction probability of the photons at their first hit on the coupling face to the photodetector (reduction of multiple bouncing) [19].

The fact that some self-activated scintillators, like $PbWO_4$, exhibit fast room temperature scintillation in the ns-range is only the consequence of a luminescence quenching mechanism competing with the radiative relaxation of the excitation. In this case the decay is non-exponential, which is a common signature of temperature quenched scintillators.

3.2.3 Response Linearity and Energy Resolution

The ultimate energy resolution (FWHM) of a perfect scintillator based detector is given by the well-known Poisson law:

$$R_{\text{lim}} = 2.35\sqrt{\frac{1 + v(PD)}{N_{\text{pe}}}} \tag{3.11}$$

where $v(PD)$ is the variance of the photodetector gain and N_{pe} is the number of photoelectrons emitted by the photodetector. As the number of photoelectrons is proportional to the number of photons N_{ph} produced by the scintillator, the resolution should be driven by the photostatistics of the scintillator light production. However, several other factors contribute to the practical resolution R:

$$R^2 = R_{\text{lim}}^2 + R_{\text{inh}}^2 + R_{\text{tr}}^2 + R_{\text{np}}^2 \tag{3.12}$$

where R_{inh} reflects homogeneities of the crystal, inducing local variations of the scintillations efficiency, R_{tr} is related to the light transport and collection by the photodetector and R_{np} is a factor of non-proportionality, which accounts for the fact that for some scintillators, the number of emitted photons is not strictly proportional to the incident energy.

Non-linear response has been first reported for NaI(Tl) and CsI(Tl); the response per unit deposited energy decreases continuously from X- and γ-rays to electrons, protons, α particles, and fission fragments. Moreover, this trend is strongly correlated with the ionization density dE/dx [20]. In other words, the response of

a scintillator depends not only on the total amount of energy but also on the mechanisms of the energy deposit. There is common agreement that this is related to the saturation of response of the luminescent centres in the presence of a high density of charge carriers. This is parameterized by Birks law, which postulates a non-radiative relaxation of excitons interacting with each others in the case of high ionization density:

$$N_{\mathrm{ph}}\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right) = \frac{N^0_{\mathrm{ph}}}{1 + a_{\mathrm{B}}\frac{\mathrm{d}E}{\mathrm{d}x}} \tag{3.13}$$

where $N^{\circ}{}_{\mathrm{ph}}$ is the light yield in the absence of saturation, N_{ph} is the actual light yield and a_{B} is the Birks parameter.

When combining the $1/\beta^2$ ionization density increase for low energy particles of decreasing velocity β (Bethe Bloch formula) with the Birks saturation law one obtains the typical scintillator non-linear response at low energy as illustrated on Fig. 3.10 in the case of NaI(Tl) [21].

It remains, however, to be explained why some scintillators are more affected by this saturation effect than others.

Each of the steps of the conversion process described in the previous section can be characterized by a certain degree of non-linearity. It seems, however, that last stages of thermalization and capture are the most affected by non-linear phenomena. Indeed, as long as the kinetic energy of electrons and holes is large relative to the bandgap E_{g} the excess energy will be used to produce secondary e-h pairs and this energy conversion process is intrinsically linear. On the other hand the stability of the thermalized excitons in its crystallographic environment is very much dependant on the energy band structure of the material as well as on the density of luminescent centres or defects. This stability is related to the correlation distance between the electron and hole, which is energy and temperature dependant.

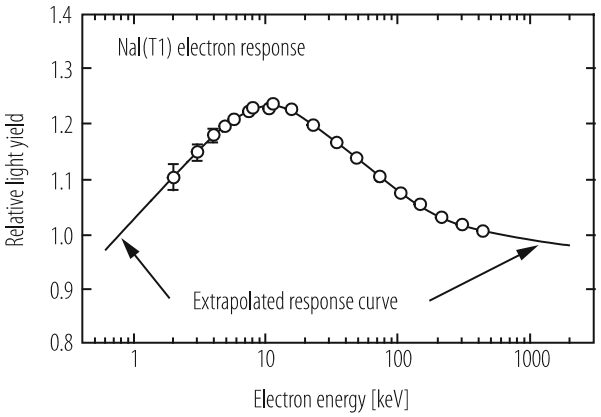


Fig. 3.10 Measured electron response for NaI(Tl) scintillator (from ref [21])

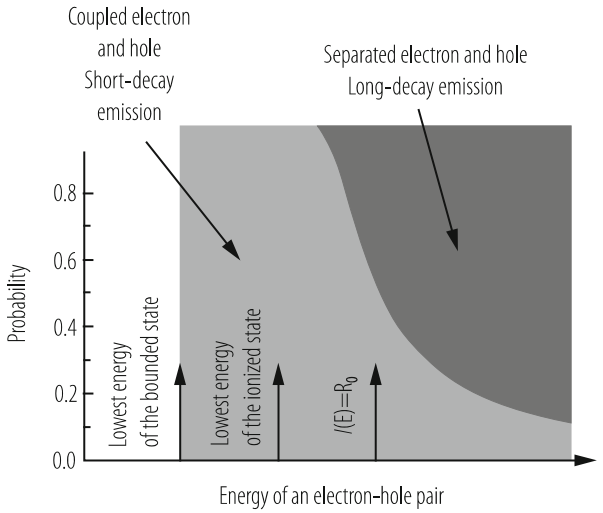


Fig. 3.11 Probability of binding or separation of an e-h pair as a function of energy (courtesy A. Vasiliev)

Two competing recombination processes can take place, both being intrinsically non-linear with energy as shown on Fig. 3.11, and inducing therefore a non-linear energy response of the scintillator. The first one is the self-trapping of the exciton in the vicinity of a luminescent centre which decreases rapidly with the e-h pair energy. The second one is the direct capture of the separate electron and/or hole by defects or luminescent centres and increases with the kinetic energy of the electron and hole.

The energy threshold between these two mechanisms is related to the correlation distance R_0 between the electron and hole, which is temperature dependant. As a result, the energy dependence of the scintillator response to thermalized e-h pars is strongly non-linear as shown on Fig. 3.10, which also shows the influence of the defects (crystal quality) on the excitation transfer efficiency to the luminescent centres.

The quantitative link between the low energy non-linearity of the scintillator response and the deviation of its energy resolution from the predicted counting statistics is far from being fully understood. It has however its seed in the fact that for the same total amount of deposited energy both photons and electrons release this energy in a number of quanta over a large energy range and that the light response for each of these quanta has different proportionality constants as a function of energy. The event-to-event variation of this cascade process induces a spread in the energy response, which deteriorates the energy resolution.

This is obvious in the case of Compton scattering. In a detector of a finite size, the events in the photopeak result from the sum of true photoelectric events and of events having undergone single or multiple Compton scattering interactions all

contained in the detector block. The total energy deposited in the detector block is the same whether it results from a single or multiple interactions. The light response may however differ due to the non-linear response of the scintillator. As a result, the event-to-event statistical variation of the energy deposition mechanism induces a broadening of the resolution.

As pointed out in ref. [22] one would expect an improvement of the resolution by reducing the detector size, as the fraction of fully contained Compton events decreases and consequently the proportion of true photoelectric events increases. This is actually not the case, because photoelectric events may result at the atomic scale from a complex cascade mechanism. Indeed, the photoelectric interaction of an X- or γ-ray produces a mono-energetic electron from one of the inner shells of the atoms of the absorber. However, this electron can be ejected not only from the K shell but also from a L or even a M shell (although the cross-section rapidly decreases for higher core levels) of the different atoms of the crystal. In the sequence of photon detection, recoil electrons with different energies are produced, each carrying the incident photon energy minus the binding energy of the shell, from which it has been ejected. Moreover, the deep hole produced in the inner shell will be filled by an electron from outer shells, which in turn will be replaced by electrons rom even lower bound shells through a cascade of relaxation events, each of them producing an X-Ray or an Auger electron converting in the crystal following the same mechanisms. Figure 3.12 depicts a part of this cascade process for LSO crystals, commonly used in medical imaging cameras.

Finally, the recoil electrons, as well as all charged particles detected in a scintillator, slow down through a sequence of energy transfers to the absorber with a progressively increasing ionization density.

The energy resolution of calorimeters used in high or medium energy physics is generally parametrized as a function of energy according to the following formula:

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{3.14}$$

where a is the statistical term, b the noise term and c a constant term, which takes into account all the systematics (intercalibration error, temperature effects, light yield non-uniformity in the crystal, shower leakage, etc.).

At high energy, the constant term is predominant and it requires a challenging engineering effort to reach the sub-percent level for large detector systems with tens of thousands of channels. This has been achieved for the LEP L3 BGO calorimeter (with 12,000 crystals) with a high energy resolution of 1% and in the LHC CMS PWO calorimeter (77,000 crystals) with a constant term of better than 0.5%.

At lower energy, the electronic noise plays an increasing role. The noise contribution, which is energy independent, contributes therefore to the *relative* energy resolution (3.14) as $1/E$.

An interesting example is given in Fig. 3.13 for two heavy Lutetium based crystals popular for medical imaging devices, LSO and LuYAP. In spite of a light

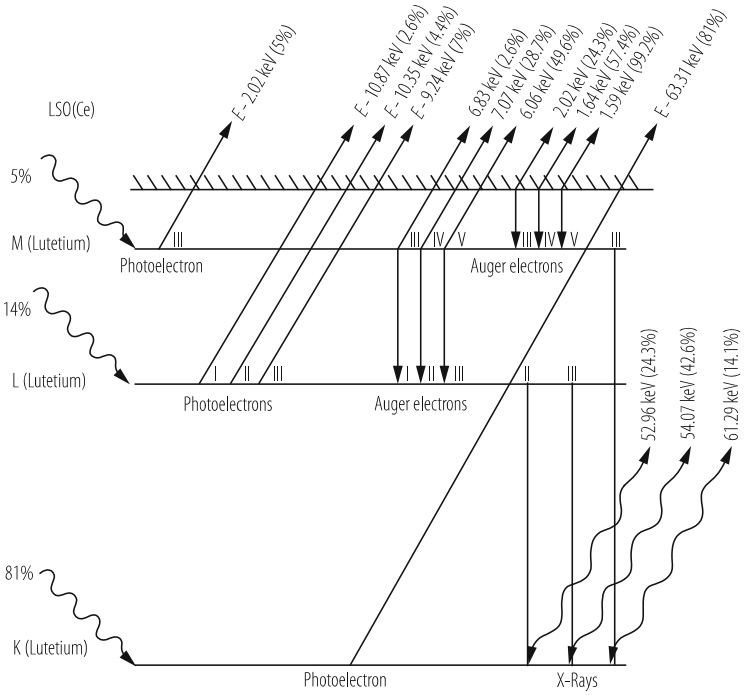


Fig. 3.12 Electron cascade following photoelectric absorption in LSO Crystal. *E* refers to the photoelectrically absorbed photon energy (ref. [22])

yield nearly three times lower LuYAP achieves a comparable energy resolution than LSO because of a much more linear behavior at low energy (see Fig. 3.4).

3.2.4 Scintillation Kinetics and Ultrafast Emission Mechanisms

Achieving ultimate time resolution on scintillator-based detectors requires a parallel effort on the light production mechanisms, light transport optimization to reduce the travel time spread of the photons on their way to the photodetector, on the photoconversion system as well as on the readout electronics.

As shown in Sect. 3.2.1 the radiative transition on the activator ion or on the intrinsic luminescent center only takes place after a complex relaxation mechanism of the primary electron-hole pairs that can last several nanoseconds. In this process large statistical fluctuations are therefore induced for the generation of the first scintillation photons, which influence the observed rise time. This presents an intrinsic limit to the achievable time resolution in a scintillator. It is related to the

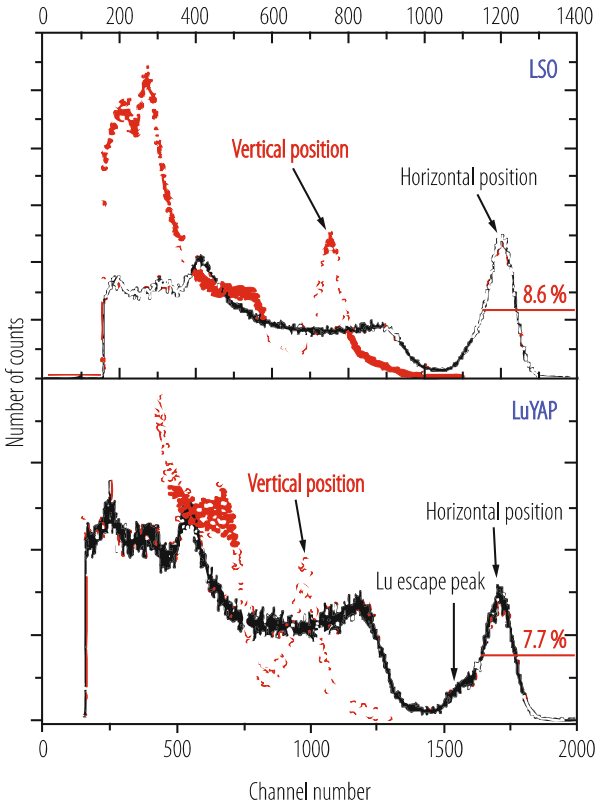


Fig. 3.13 Energy resolution for ^{137}Cs photons obtained with 2 × 2 × 10 mm^3 Ce doped LSO and LuYAP crystals measured in horizontal and vertical position. The electronic gain for LuYAP is three times higher to compensate for the lower light yield (6000 pe/MeV and 2000 pe/MeV for LSO and LuYAP, respectively, in horizontal position)

time fluctuations in the relaxation process that can be estimated to be of the order of 100 ps.

For sub-100 ps time resolution mechanisms involving the production of prompt photons need to be considered. Cherenkov emission and cross-luminescent materials can offer a solution. However, the number of Cherenkov photons from the recoil electrons resulting from a 511 KeV γ conversion is very small, of the order of 20 photons in crystals like LSO, LuAP and GSO. Moreover, these photons are preferentially emitted in the UV part of the spectrum, where the optical transmittance and the photodetector quantum efficiency are generally low. The same applies for cross-luminescent materials characterized by a reasonably fast emission (600 ps for BaF_2) which emit in the 100–250 nm spectral range. However, some transient phenomena in the relaxation process that can be possibly exploited for the generation of prompt photons. From this point of view, an interesting phase of the relaxation mechanism is the thermalization step when the hot electrons and

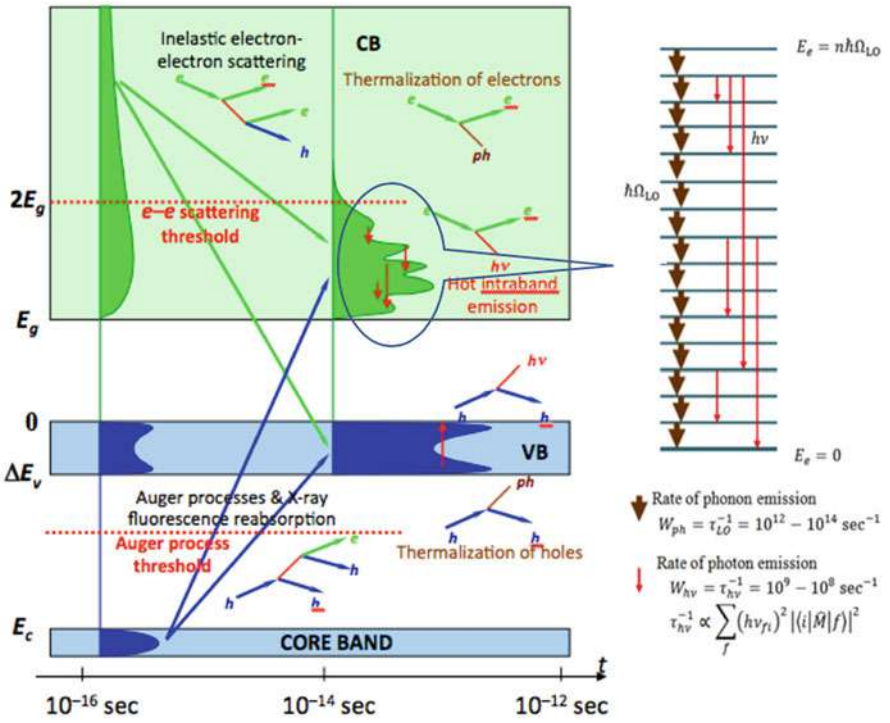


Fig. 3.14 Schematic description of the hot intraband luminescence, showing the competition of radiative and non-radiative (phonon-assisted) decay channels in the case of a non-uniform density of states in the conduction band. From Ref [23]

holes have passed the ionization threshold. The coupling to acoustic and optical phonons in the lattice is the source of hot intraband luminescence (HIBL) that could be exploited to obtain a time tag for the interaction of ionizing radiation with a precision in the picosecond range [23, 24]. This emission is rather weak but extremely fast (sub-ps) and is characterized by a flat spectrum in the visible for the electron-induced HIBL in the conduction band with an onset in the near infrared attributed to the hole HIBL in the valence band. Work is ongoing to engineer scintillators with a non-uniform density of states in the conduction and/or the valence band which may result in a more intense HIBL emission (Fig. 3.14). Already a few hundred prompt photons would suffice to significantly improve the time resolution of scintillators like LSO in the low energy (MeV) regime.

Hetero-structures based on a combination of standard scintillators (such as LSO or LYSO) and nanocrystals may be another way to produce prompt photons. Nanocrystals have gained considerable attention over the last two decades because of their excellent fluorescence properties. In such systems quantum confinement offers very attractive properties, among which a very high quantum efficiency and ultrafast decay time. Moreover, they have a broadband absorption and narrow emis-

sion, enhanced stability compared to organic dyes, and the fluorescence is tunable from the UV to the near-infrared spectral range (300–3000 nm) by nanocrystal size and material composition.

A novel route towards the realization of ultrafast timing resolution is possible with the use of colloidal CdSe nanosheets (CQwells) [24], a new class of two-dimensional materials. CQwells are solution-processed analogs to epitaxial quantum wells (Qwells). However, being synthesized in solution, they can be deposited on any substrate with arbitrary geometrical configuration. Further, a large dielectric mismatch between the inorganic CdSe CQwells and the surrounding organic environment results in much stronger quantum confinement than in epitaxial Qwells. This mismatch combined with very little dielectric screening due to the 1.5 nm CQwell thickness results in strongly enhanced exciton and biexciton binding energies of 132 and 30 meV, respectively, making both populations stable at room temperature.

The strong electron and hole confinement in one dimension and free motion in the plane has several important consequences, including strict momentum conservations rules (in contrast to quantum dots) and a giant oscillator strength transition. Momentum conservation in CQwells limits the available states for Auger transitions, reducing the recombination rate of this nonradiative channel. In addition to the enhanced exciton and biexciton binding energies, a giant oscillator transition results in radiative lifetimes that are significantly shorter than in bulk CdSe (~400 and ~100 ps, respectively). All of these properties contribute to the ultralow threshold stimulated emission (or superluminescence) with sub-ps decay time that has been observed with these CQwells (Fig. 3.15). Such systems could find

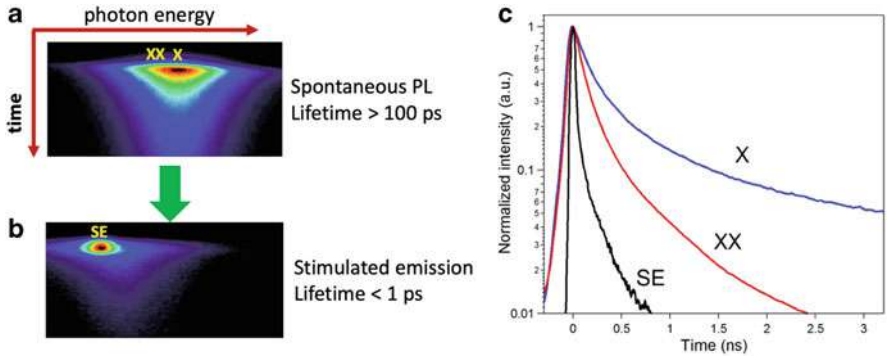


Fig. 3.15 Time-resolved spectral decay under femtosecond excitation (**a**) Streak image showing the spectral decay of exciton (X) and biexciton (XX) emission from CdSe CQwells. (**b**) Stimulated emission at an ultralow excitation fluence of $F_0 = 6\ \mu J/cm^2$, with characteristic spectral narrowing and lifetime shortening. From Ref [24]

interesting applications in ultrafast X-Ray imaging as well as providing a fast time tag in γ imaging if used in hetero-structures in combination with dense scintillators like LSO with a structuration dimension of the order of the recoil electron range, as suggested in Ref [25].

3.3 Role of Defects on Scintillation Properties and on Radiation Damage in Inorganic Scintillators

3.3.1 Structural Defects in a Crystal

The properties of a scintillator strongly depend on the structural quality of the crystal lattice. The presence of defects influences all stages of the scintillation process. They play also an important role in the light transport to the photodetector, as well as in the generation of optically active defects under radiation exposure. They continuously exchange charge carriers and phonons with the crystal lattice and are therefore in thermodynamic equilibrium with the medium. This can have a number of consequences such as reduced or enhanced scintillation efficiency if the charge carriers are channelled through these defects to non-radiative or radiative traps respectively, modification of the scintillation kinetics, afterglow, creation of perturbed emission centres, self-absorption, emission wavelength shift, radiation damage, radiation damage recovery. Depending on their size and physical nature, one can distinguish two main classes of structural defects, namely point size defects and impurities. Larger scale defects such as dislocations, twins, voids and other macroscopic defects also exist. They will not be described here, as their influence on the crystal properties is usually limited to the mechanical ruggedness and to a small extent to the optical homogeneity.

3.3.1.1 Point Size Defects

A perfect crystal is a virtual object that can only exist at absolute zero temperature. At higher temperature, a thermodynamic equilibrium is obtained by exchange of energy quanta (in the form of phonons) between the environment and the crystal lattice. Moreover, the finite dimensions of the crystal imposes conditions on the surface to compensate the electrostatic field unbalance for the atoms at the interface. This requires some level of plasticity of the lattice, which is generally achieved by a certain concentration of cation and anion vacancies. Thermodynamics imposes a relatively low concentration of such defects at room temperature, typically of the order of 10^{12} cm^{-3}. For comparison, the atomic density of the majority of known heavy scintillators is about 10^{23} cm^{-3}. In practise, the concentration of vacancies is determined by the crystal growth technology. The melt is a mixture of several chemical components, each of them with a different melting temperature

and vapour pressure, which leads to segregational evaporation of some components. Furthermore, close-to-surface vacancies can be partially compensated by absorption of ions or radicals from the surrounding atmosphere. Typically the concentration of such defects is at the level of 10^{18} cm^{-3} (10 ppm atomic) or even more. At such concentration, some collective effects can take place, leading to more complex molecular or cluster defects. Another typical point defect results from the displacement of an ion of the lattice to an interstitial position. The electrically neutral system behaves as a dipole and is called a Frenkel defect. In the case of Lead Tungstate an oxygen-based Frenkel defect is responsible for an absorption band at 360 nm and for an increased susceptibility to radiation damage.

3.3.1.2 Impurities

Impurities are ions of different nature than the constituents of the crystal lattice. They are generally introduced from imperfectly purified raw materials or by contamination, for instance from the crucible material, during the crystal growth process. Doping ions acting as luminescence activators, such as Ce^{3+} in LSO, LuAP and many other fast scintillators, can be considered as impurities with a positive role. Ions from the lattice, but in a different valence state than required by the electric charge balance, are another type of impurity. As an example, Ce^{4+} has been considered by some authors as a possible scintillation quencher in CeF_3 crystals. Two important parameters influence the way impurities can be introduced in a crystal: their electric charge and their ionic radius. If the ionic radius is close to the one of ions from the lattice, impurities can easily replace these ions, producing only a small distortion of the lattice. Isovalent ions will then easily produce a solid solution as is the case for LYSO or LuYAP when Y^{3+} ions substitute Lu^{3+} in LSO and LuAP crystals, producing locally a mixed compound of LSO-YSO and LuAP-YAP, respectively. If heterovalent impurities are introduced in the crystal their charge excess or deficit must be compensated by other impurities or by lattice ion vacancies. This mechanism can be used to suppress the detrimental role of some defects, which cannot be eliminated. A good example are the lead vacancies in PWO, which are efficient hole traps responsible for radiation damage and which can be compensated by substituting trivalent ions such as Y^{3+} or La^{3+} to neighbouring Pb^{2+} ions in the lattice.

Impurities with too large an ionic radius have generally little chance to be introduced in the lattice, whereas small ions can find interstitial positions and create strong local distortion of the crystal electronic configuration.

In practice it is difficult, or at least very expensive, to purify raw materials to the sub-ppm level. Most of the scintillators grown in good conditions have therefore an impurity concentration of about 10^{-17}–10^{-19} cm^{-3}, comparable to the concentration of point defects.

3.3.2 Impact of Defects on Optical Properties

Defects in a crystal influence its optical properties in a number of ways, affecting the charge carriers or the photon transport.

3.3.2.1 Charge Carrier Traps

Most point defects or impurities are electron or hole traps. They reduce therefore the transfer efficiency of charge carriers to the luminescent centres and therefore also the scintillation efficiency. For good quality crystals the density of defects (at 1–100 ppm level) is several orders of magnitude smaller than the density of luminescent centres, which is very high for intrinsic scintillators (about 10^{22} cm^{-3}) but also quite high for extrinsic scintillators, for which the activator concentration is typically at the atomic percent level. Under normal excitation conditions, it would look therefore rather unlikely that charge carriers are trapped by defects before they convert on luminescent centres. This does not take into account the charge carrier capture cross-section, which can vary by large factors for different kinds of traps. A typical example is given by the molybdenum molecular complex MoO_4^{2-}, which is a very efficient and stable electron trap with a radiative decay at 508 nm in PWO. At the level of only a few ppm it gives rise to a slow (500 ns) additional green component to the regular fast PWO emission band at 420 nm. As molybdenum is isomorphic to tungsten it can easily enter into the PWO lattice and locally produce a solid solution ($PbWO_4$-$PbMoO_4$). This slow green component is negligible if the molybdenum contamination of the tungsten oxide raw material is less than 1 ppm [26].

In some cases, the traps are non-radiative but have energy levels close enough to the valence or conduction bands so that the carriers can be released by thermal activation, eventually converting on the luminescent centres. If the trap is close to the radiative centre this thermally assisted transfer can take place directly between them without involving the valence or conduction bands. As a result, the regular emission will take place but with some delay associated with the transit of the carrier via the trap. This is the origin of the well-known afterglow or phosphorescence. When afterglow effects are undesirable, for instance for high X-ray counting rates in CT scanners, additional impurities can help opening some non-radiative relaxation channels for these traps. As an example, afterglow in $(Y,Gb)_2O_3$:Eu scintillators can be significantly reduced by the addition of heterovalent Pr^{3+} or Tb^{3+} ions to the lattice [27]. The Pr^{3+} and Tb^{3+} additives readily trap holes to form Pr^{4+} and Tb^{4+}, which compete with the intrinsic traps responsible for afterglow. This energy trapped in the Pr or Tb sites decays non-radiatively in the presence of the Eu^{3+} ion. As a consequence, afterglow emission is suppressed by one order of magnitude or more.

3.3.2.2 Defect Associated Absorption Bands

Defects have generally energy levels in the forbidden band, which reduce the optical transparency of the crystal. Small perturbations of the crystal lattice are energetically the most probable ones and give rise to a number of energy levels near the conduction and the valence bands. There is nearly a continuum of such levels, which reduces the optical transparency window of the crystal. For this reason, the shape of the optical transmission of a crystal near the band-edge is usually a good probe of its structural quality. Crystals with UV emission bands near the fundamental absorption edge are strongly affected by the optical transitions between these levels resulting in increased absorption.

Cross-luminescent crystals such as Barium Fluoride (BaF_2) are illustrative examples to demonstrate the role of impurities on the crystal properties. Their deep UV fast emission band (220 nm for BaF_2) requires a very good UV transmission to detect efficiently the light at the photodetector. Unfortunately, alkali earth fluorides are easily contaminated by oxygen and hydroxyl ions, causing strong absorption bands in the UV. A theoretical study of the charge state stability and electronic structure of O^0, O^- and O^{2-} centres in BaF_2 identified a large number of transitions from 2p to 3s and 5s states. In ref. [28] Hartree-Fock-Slatter local density discrete variation cluster calculations were made to obtain the energy levels of H_s^-, O_s^- and O_s^{2-} ions in BaF_2 crystals. Table 3.1 summarizes the optical absorption bands in the VUV and UV ranges.

As far as O^- and O^{2-} ions are concerned, the absorption bands are mainly the result of cross transitions between oxygen ions and Ba^{2+} or F^- ions, which significantly contribute to absorption around 200–240 nm.

These theoretical calculations are in good agreement with experimental results, confirming the existence of strong absorption bands overlapping the fast emission band in hydrolysed BaF_2 crystals, see Fig. 3.16.

Table 3.1 Calculated optical absorption band of H_s, O_s^- and O_s^{2-} -contaminated BaF_2 [28]

Impurities	λabs. [nm]	$h\nu$ [eV]	Cross transitions
H_S^-	209	5.9	H^- (1s) → H^- (2s)
O_S^-	230	5.4	F^- (2p) → O^- (2p,3p)
	175	7.2	F^- (2p) → O^- (3p)
	170 ≈ 175	7.0 ≈ 7.2	O^- (2p) → Ba^{2+} (5d)
O_S^{2-}	292	4.2	F^- (2p) → O^{2-} (3p)
	200	6.2	O^{2-} (2p) → Ba^{2+} (6s)
	130	9.5	O^{2-} (2p) → Ba^{2+} (5d)

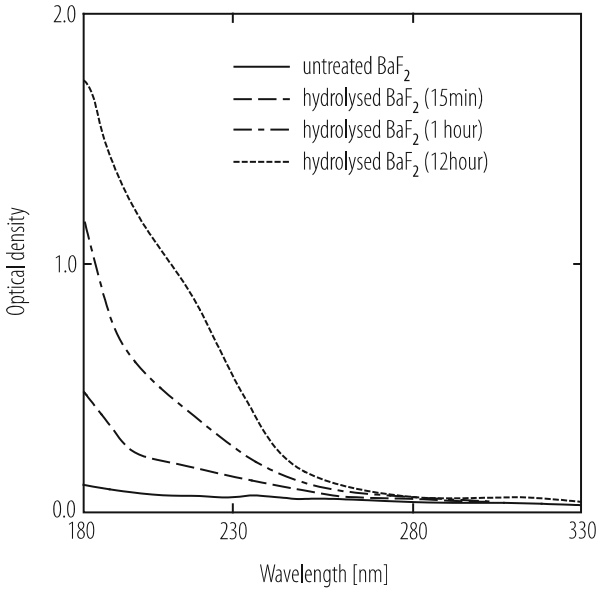


Fig. 3.16 Absorption spectra for different hydrolysed BaF_2 (ref. [28])

3.3.3 Radiation Damage

The exposure of crystals to ionizing or neutron radiation can induce a number of modifications of the crystal lattice with potential consequences for the scintillation efficiency and the light transport. These modifications can be related to pre-existing crystal defects, when exposed to a high density of charge carriers that are easily trapped producing colour centres with radiation-induced absorption bands. They can also be associated to the production of new defects by elastic or knock-on collisions of incident particles with the lattice ions resulting in a local modification of the lattice structure. Finally, heavy energetic charged particles or neutrons may produce dramatic events, such as heavily ionizing fission fragments. This last phenomenon is usually of little concern in the majority of applications, even for the new generation of high luminosity particle physics colliders, as it requires an enormous integral fluence (10^{17}–10^{18} cm^{-2}) to become significant. Indeed, it requires the formation of about 10^{17} cm^{-3} such defects to reach a 1 ppm contamination in the majority of scintillator materials. However, such defects are by nature irrecoverable and their progressive accumulation may affect parts of detectors highly exposed for very long periods of time.

The situation is different for the majority of other cases (charge trapping or ion displacement), for which relaxation processes play a fundamental role in the kinetics of damage build-up. These defects introduce a local perturbation in the crystal and do not change the main structure parameters and particularly the spatial

symmetry group. However, they locally modify the electronic configuration and affect the macroscopic crystal parameters, such as optical transmission, conductivity, thermo-luminescence properties, because these volume properties are sensitive to the microscopic structure modifications. In ionic crystals, containing anions and cations, five possible simple point defects of the crystalline structure have been observed: anion vacancy V_a, cation vacancy V_c, cation replacement by impurity ions, extrinsic atoms in inter-site positions and Frenkel type defects (anions and cations displaced to interstitial sites).

All these defects are efficient charge carrier traps and can be stabilized by capturing excess electrons or holes released by irradiation in the conduction or valence band respectively. In oxide compounds for instance, the oxygen vacancies are charge compensated by the capture of one or two electrons, which are in excess in the conduction band after irradiation. The resulting F^+: ($V_a + e^-$) and F: (V_a+2 e^-) electron centres play an important role in radiation damage effects. The captured electron or hole in these so-called recharged defects has generally a number of discrete energy levels available in the electrostatic environment of the defect and optical transitions to upper energy levels induce absorption bands in the crystal transparency window. These bands are the source of the crystal colouring under irradiation and justify the name of colour canters for these defects.

The main consequence of irradiating a crystal is to produce radiation induced absorption bands, which absorb a fraction of the scintillation light on its pathway to the photodetector. The light collected on the photodetector becomes therefore:

$$I_{\text{rad}} = \int_{\lambda} I_0\,(\lambda)\, \mathrm{e}^{-(\mu_0(\lambda)+\mu_{\text{rad}}(\lambda))L} \mathrm{d}\lambda \tag{3.15}$$

where I_{rad} is the intensity of the transmitted light after irradiation, $I_0(\lambda)$ is the intensity of transmitted light at the wavelength λ before irradiation, $\mu_0(\lambda)$ and $\mu_{\text{rad}}(\lambda)$ are, respectively, the intrinsic and radiation induced absorption coefficient at the wavelength λ and L is the mean path-length of optical photons from the emission point to the crystal exit surface. Dense and small radiation length crystals have an obvious advantage as for the same stopping power the path-length L is reduced as compared to lighter materials. Moreover, non-uniformities introduced by different path-lengths as a function of the position of the scintillation emission point are also reduced. Figure 3.17 shows the radiation induced absorption coefficient spectrum for PWO crystals as a function of the accumulated ^{60}Co dose.

At radiation levels currently experienced in particle physics detectors and in X-ray imaging devices the radiation damage only affects the optical transparency of the majority of known scintillators, but not the scintillation mechanism. One exception is CsI(Tl), characterized by an overlap of the radiation induced hole centres absorption maximum in CsI with the excitation spectrum of the Tl^+ ions. The presence of stable hole centres causes a fraction of excitations to be trapped rather than transferred to TI^+ thereby causing non-radiative losses. As a result, the efficiency of energy transfer to luminescence centres drops, decreasing the scintillation efficiency. Similarly, radiation-induced charge transfer processes can

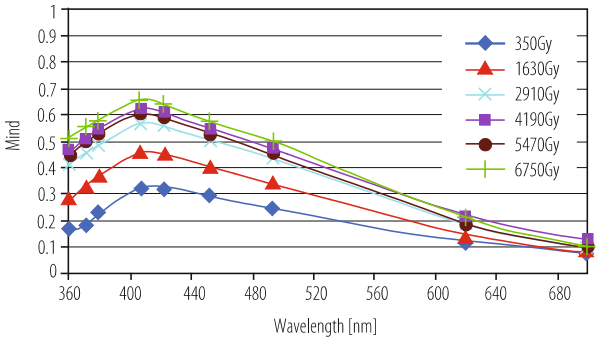


Fig. 3.17 Wavelength dependent absorption coefficient of PWO crystals as a function of the absorbed ^{60}Co dose (courtesy CMS collaboration)

modify the charge state of activator ions. This is seen for instance in some Ce^{3+} doped scintillators, such as YAP and LuYAP, when grown in vacuum or inert atmosphere, where up to several percent of the scintillating Ce^{3+} ions can be reduced under irradiation to the Ce^{2+} non-scintillating state, decreasing by the same amount the scintillation efficiency. Annealing the crystals under oxygen atmosphere restores the scintillation efficiency by re-oxidizing the Ce^{2+} ions. Ref. [11] provides more details.

The kinetics of the radiation damage build-up and recovery is determined by the depth of the traps at the origin of colour centres. Very shallow traps induce transient absorption bands, which recover so quickly that the monitoring of the crystal transparency becomes very difficult. Much attention has been paid when optimizing PWO crystals for the CMS calorimeter at LHC to suppress as much as possible such defects or to compensate their effect by specific doping [8, 29]. On the other hand, deep traps are generally very stable and are characterized by a continuous increase of the corresponding absorption bands, even at low dose rate, until they are completely saturated. The monitoring of the crystal transparency allows correcting for light yield variations but the concentration of such defects must be maintained small enough to minimize the loss in light yield. For most of the known scintillators a concentration of such defects at the ppm level can produce a radiation induced absorption coefficient limited to about 1 m^{-1}.

At room temperature a large fraction of the radiation induced defects are metastable. Temperature dependant relaxation processes take place in the crystal lattice so that these defects, once produced, are ionized at a rate, which depends on their energy depth and the temperature following the Boltzmann law. As a consequence, the transmission damage reaches a saturation level, which is dose-rate-dependent up to the point where the rate of trapping of the charge carriers induced by radiation is exactly balanced with the rate of spontaneous relaxation at this working temperature. For a uniform distribution of defects of type *i* in the crystal

and in the absence of an interaction between them the kinetics of the concentration of damaged centres of type i is described by the following differential equation:

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = -\omega_i N_i + \frac{S}{d_i}\left(N_i^* - N_i\right) \tag{3.16}$$

where N_i is the amount of damaged centres of type i at time t, ω_i is their recovery rate, S is the dose rate, N_{i*} is the amount of pre-existing defects of type i and d_i is a damage constant, which depends on the capture cross-section of free carriers by the centres of type i. The induced absorption coefficient μ produced by irradiation is proportional to the concentration of absorbing centres N through $\mu = \sigma N$, where σ is the cross-section of the absorbing centre. The solution of this equation gives the kinetics of the induced absorption build-up:

$$\mu = \mu_{\mathrm{sat}} \frac{S}{S + \omega d}\left\{1 - \exp\left[-\left(\omega + \frac{S}{d}\right)t\right]\right\} \tag{3.17}$$

where $\mu_{sat} = N_* \sigma$ corresponds to the maximum possible saturation when all centres are damaged. The recovery of the transmission after the end of the irradiation at time t_0 is described by:

$$\mu = \mu_{\mathrm{sat}} \frac{S}{S + \omega d}\left\{1 - \exp\left[-\left(\omega + \frac{S}{d}\right)t_0\right]\right\} \exp\left(-\omega\left(t - t_0\right)\right) \tag{3.18}$$

Figure 3.18 illustrates the impact of this behaviour on the light output of a 23 cm long PWO crystal exposed to a cycle of several irradiations separated by periods of recovery.

There are two ways to increase the radiation hardness of scintillating crystals. The first one is to make every effort to reduce the density of point charge defects

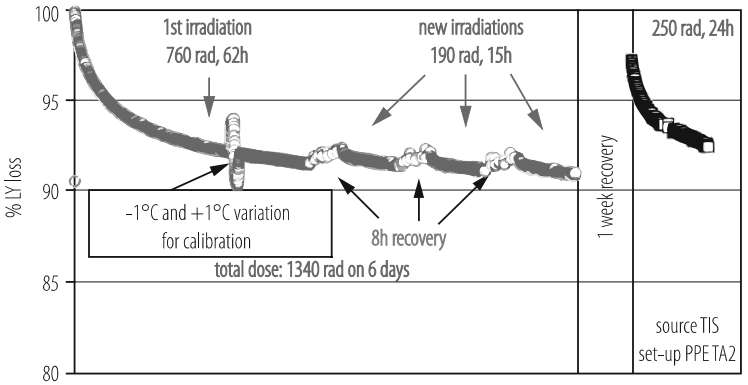


Fig. 3.18 Variations of light out-put for a PWO crystal exposed to a cycle of several irradiations separated by periods of recovery at 18°C (courtesy CMS collaboration)

related to structural defects, impurities and anion or cation vacancies induced by differential evaporation of the chemical components during the crystal growth. This can be achieved for the majority of crystals, through different cycles of purification of the raw materials, multiple crystal growth and annealing of the crystals in specific atmosphere and temperature conditions. This approach is however costly and limited to defect concentration levels in the ppm range. For some applications, such as in high luminosity collider experiments, this is sometimes not enough to guarantee the optical stability of the crystals over long periods.

In another approach additional well selected defects are produced in the crystal, which compete with the uncontrollable defects and reduce their influence. This so-called co-doping strategy has been the result of improved understanding of the mechanisms of light production and charge carrier transport and trapping, opening the way to a defect engineering of the crystals. It has been shown for instance that divalent doping with Ca^{2+} or Mg^{2+} in some Ce^{3+} activated crystals (in particular in ortho-silicates and aluminium garnets), not only increases the light yield, but also suppresses slow scintillation components and improves the radiation hardness [30]. This is the result of easier charge carrier transport to the luminescent centres through the energy levels of these impurities and easier delocalization of trapped carriers due to the smaller energy gap between these traps and the conduction band, which may even absorbed in the conduction band.

3.4 Crystal Engineering. Impact of New Technologies

The conditions of synthesis of the chemical components of a crystal are governed by thermodynamic relations between composition, temperature and pressure of the mixture. At a given pressure, the composition-temperature equilibrium for both the liquid and solid phases is represented by a phase diagram. The phase diagram shows the domains of stability of a given chemical composition and the influence of deviations from stoichiometry (composition of the mixture), unwanted impurities or specific doping. An example of such a phase diagram is shown in Fig. 3.19 for PWO crystals.

Two stable compositions can be grown from a PbO-WO_3 mixture, namely $PbWO_4$ (PWO) and Pb_2WO_5. The $PbWO_4$ melts congruently, i.e. without decomposition of the compound, at 1123°C. The analysis of this phase diagram helps to define some practical parameters for the $PbWO_4$ crystals. First of all the melting temperature restricts the choice of the crucibles to metals with melting points much higher than 1123°C, such as platinum, iridium and their various alloys. Moreover, such crucibles must be chemically inert with melts of similar oxides like $PbMoO_4$, $CaMoO_4$, $ZnWO_4$, as Mo, Ca and Zn are impurities likely to be present in the raw materials. Secondly, the possibility to deviate from the perfect stoichiometric composition of the raw material with some excess of either WO_3, or PbO is of great importance to compensate for a strong differential evaporation of the different components of the melt during the growth process. An initial deviation

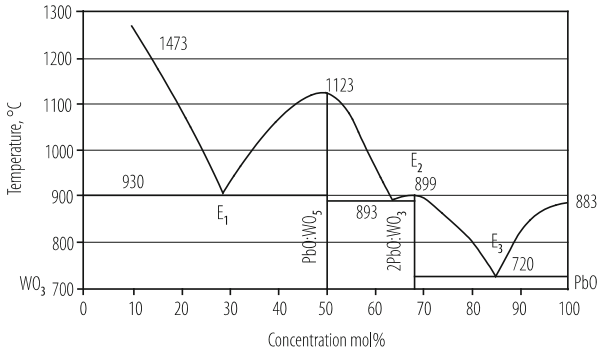


Fig. 3.19 Phase diagram of the PbO-WO_3 system

from the perfect stoichiometry can compensate non-stoichiometry defects. Some restrictions can appear because of segregation processes of additional doping ions. The segregation coefficient k defines how the concentration of doping ions or impurities will vary along the crystal according to the formula:

$$C_s = \frac{kC_0}{1-(1-k)\,g} \tag{3.19}$$

where g is the fraction of the melt already crystallized, C_S is the impurity concentration in the melt at some point, C_0 is the initial impurity concentration in the melt, k is the segregation coefficient. If the segregation coefficient k is too different from 1, as a result of too small or too large ionic radii or different valence states as compared to the ions of the crystal lattice, the doping ion will be pumped in or repelled from the crystal during the growth process.

The majority of crystal growth methods are based on the principle of oriented crystallization. An oriented seed (a small piece of the same crystal or of different composition but similar lattice parameters) is introduced in contact with the melt to initiate the growth process. A temperature gradient is applied so that heat transfer is used as the driving force of crystallization. Several crystallization methods have been developed, which differ in the way the heat transfer and the hydrodynamic conditions are applied:

- Establishing a temperature gradient between the crystal and the melt by heat transfer from the seed. Such heat transfer methods occurred in nature to form crystals and are still used for cheap crystal production, when the requirements on quality are not too high.
- Floating temperature gradient through the melt (Bridgeman and Stockbarger methods). The raw material is placed in a closed crucible, at the end of which a seed has been placed. The crucible is moved through a thermal gradient zone, where the temperature is lowered below the melting point. This is the area where the crystallization takes place. The volume of the melt will therefore decrease

continuously and the growing crystal starts substituting for the melt. This method is relatively inexpensive and multiple crystal pulling is possible by moving several crucibles together through the temperature gradient zone of a single oven [31]. If the simplicity and reliability of the Bridgeman and Stockbarger methods make them particularly attractive for many applications, these methods suffer from several drawbacks, such as large variations of the temperature field parameters during the crystal growth and strong non-uniformities in the distribution of doping ions, impurities and defects in the crystal.

- Establishing a temperature gradient between the crystal and the melt in an open crucible by progressive cooling of the melt after seeding or extracting the growing crystal from melt (Kyropoulos and Czochralski methods, respectively). In the classical Kyropolos method [32] the entire crystallization process starts with the seeding and propagates through the melt as a result of a continuous temperature decrease applied during the process. There is no relative movement of the seed and the crucible. In the Czochralski method [33] the crystal is pulled from the melt. The seed is attached to a Platinum rod and put in contact with the melt in the crucible. The rod or the crucible (sometimes both) are rotating at a few rpm to maintain a good homogeneity of the melt in contact with the crystallized phase. The rod is simultaneously pulled up at a speed of typically 1–10 mm/h depending on the crystal. This method is the most widely used for growing oxide scintillators and several other types of scintillators because of its potential to grow high quality crystals by concentrating impurities and defects in the bottom part of the crucible.

More details about crystal engineering techniques are given in ref. [11].

Technologies for the production of crystals are rapidly evolving. The impressive progress in nanotechnologies in particular open new perspectives for the production of pre-reacted raw materials of excellent quality with a high uniformity of the grain sizes. With these new materials, transparent ceramics of heavy scintillators can be produced (Fig. 3.20), with the advantage over standard crystal growth techniques to be much more cost effective: not only the scintillator can be produced to its final shape, saving on the cost of mechanical processing, but also the temperature for sintering is usually much lower than for standard crystal growth.

The recently developed pulling-down technology from a shape-controlled capillary die gives the possibility to produce elongated crystals with dimensions that are not accessible using traditional cutting and polishing of bulk crystals grown by the more standard Czochralski or Bridgeman methods (Fig. 3.21). This approach has important advantages, such as growing the crystal in the final shape (round, oval, square, rectangular, hexagonal), very rapidly (several millimeters per minute instead of millimeters per hour), simultaneous multifibre pulling, increased activator doping concentration, etc. Excellent quality BGO, YAG and LSO fibers have been grown with a length of up to 2 m and a diameter between 0.3 and 3 mm. Some other materials are being studied, in particular from the very interesting perovskite family: YAP and LuAP [34].

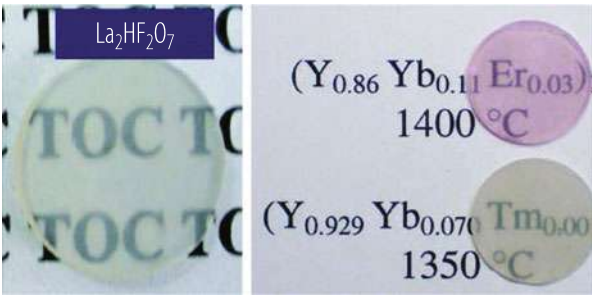


Fig. 3.20 Transparent ceramics of different heavy scintillators prepared with pre-reacted nanopowders

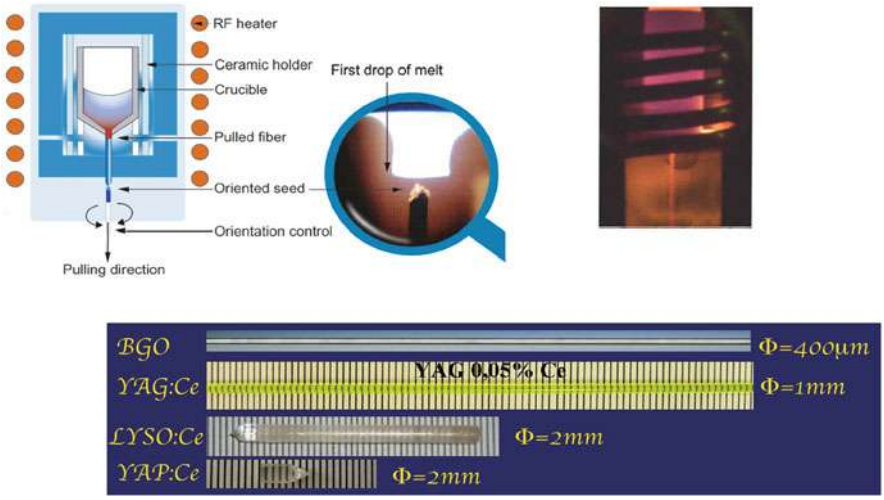


Fig. 3.21 The micro-pulling down crystal growth technology (courtesy Fibercryst)

3.5 Table of Commonly Used Scintillators

Inorganic scintillators generally considered for a majority of applications, and in particular, for particle physics detectors and medical imaging cameras are listed in Table 3.2 with their most important physico-chemical and optical properties. A much more exhaustive list of scintillators classified according to their chemical structure is presented in ref. [11].

Table 3.2 Most commonly used scintillators with their main physico-chemical and optical parameters

Scintillator	Simplified name	Density [g/cm^3]	Light Yield [ph/MeV]	Emission wavelength [nm]	Decay time [ns]	Hygro-scopic	Main application
NaI :Tl		3.67	38,000	415	230	Yes	Medical imaging, industrial γ camera, homeland security
CsI :Tl		4.51	54,000	550	1000	Lightly	Physics detectors
$CdWO_4$	CWO	7.9	28,000	470/540	20,000/5000	No	X-Ray scanner
$(Y,Gd)_2O_3$:Eu	YGO	5.9	19,000	610	1000	No	X-Ray scanner
Gd_2O_2S:Pr,Ce,F	GOS	7.34	21,000	520	3000	No	X-Ray scanner
$Bi_4Ge_3O_{12}$	BGO	7.13	9000	480	300	No	Physics detectors medical imaging
Gd_2SiO_5	GSO	6.7	12,500	440	60	No	Medical imaging
Lu_2SiO_5	LSO	7.4	27,000	420	40	No	Medical imaging
Lu_2AlO_3	LuAP	8.34	10,000	365	17	No	Medical imaging
$LaBr_3$:Ce	BriLanCe™	5.29	61,000	358	35	Very	Medical imaging
BaF_2		4.89	2000/8000	220/310	0.7/620	No	Physics detectors
CeF_3		6.16	2400	310/340	30	No	Physics detectors
$PbWO_4$	PWO	8.28	200	420	5/15	No	Physics detectors

References

1. B. Rossi, High Energy Particles, Prentice-Hall, Inc. Englewood Cliffs, NY, 1952.
2. U. Fano, Annu. Rev. Nucl. Sci. 13 (1963) 1.
3. J.D. Jackson, Classical Electrodynamics, 2nd ed., Wiley, New York, 1975, chap. 13.
4. R.D. Evans, The Atomic Nucleus, Krieger, New York, 1982.
5. A. Lempicki et al., *Fundamental limits of scintillator performance*, Nucl. Instrum. Meth. A 333 (1993) 304-311.
6. P. Lecoq et al., *Lead Tungstate ($PbWO_4$) scintillators for LHC EM calorimetry*, Nucl. Instrum. Meth. A 365 (1995) 291-298.
7. Scintillation Detectors, Catalog, Saint Gobain, Ceramiques Industrielles, March 1992.
8. A. Annenkov et al., *Suppression of the radiation damage in Lead Tungstate scintillation crystal*, Nucl. Instrum. Meth. A 426 (1999) 486-490.
9. The CMS Collaboration (2015) Technical proposal for the phase II upgrade of the compact muon solenoid, CERN-LHCC-2015-010/LHCC-P-008
10. Gundacker S. et al, (2013), Time of Flight positron emission tomography towards 100ps resolution with L(Y)SO: an experimental and theoretical analysis, JINST 8:P07014
11. P. Lecoq, A. Annenkov, A. Gektin, M. Korzhik, C. Pedrini, Inorganic Scintillators for Detector Systems, Springer-Verlag, 2006. Second edition 2017, ISBN 978-3-319-45521-1 doi 10.1007/978-3-319-45522-8
12. A.J. Dean, *Imaging in high-energy astronomy*, in: *Heavy Scintillators for scientific and industrial applications*, Proc. Int. Workshop Cristal 2000, Sep 22-26, 1992, Chamonix, France, F. De Notaristefani, P. Lecoq, M. Schneegans (eds.), Editions Frontières France, 1992, pp. 53-64.
13. C. Kuntner et al., *Intrinsic energy resolution and light output of the Lu0.7Y0.3AP:Ce scintillator*, Nucl. Instrum. Meth. A 493 (2002) 131-136.
14. M.C. Abreu et al., *ClearPEM: A PET imaging system dedicated to breast cancer diagnostics*, Nucl. Instrum. Meth. A 571, (2007) p. 81-84.
15. J.B. Birks, The Theory and Practice of Scintillation Counting, Pergamon, London, 1964.
16. A. Vasiliev, *Relaxation of hot electronic excitations in scintillators: account for scattering, track effects, complicated electronic structure*, Proc. Fifth Int. Conf. Inorganic Scintillators and their Applications, SCINT99, V.V. Mikhailin (ed.), Moscow State University, Moscow, 2000, pp. 43-52.
17. C. Pedrini et al., *Time-resolved luminescence of CeF_3crystals excited by X-ray synchrotron radiation*, Chem. Phys. Lett. 206 (1993) 470–474.
18. M. Kronberger, E. Auffray, P. Lecoq, *Probing the concept of Photonics Crystals on Scintillating Materials*, Proc. 9th Int. Conf. Inorganic Scintillators and their Applications, Winston-Salem, NC, June 4-8, 2007, IEEE Trans. Nucl. Sci. 55(3) (2008) 1102-1106.
19. A. Knapitsch, P. Lecoq, Review on photonic crystal coatings for scintillators, International Journal of Modern Physics A, Vol. 29 (2014) 1430070 (31 pages) https://doi.org/10.1142/S0217751X14300701
20. R.B. Murray, A. Meyer, *Scintillation response of activated inorganic crystals to various charged particles*, Phys. Rev. 122 (1961) 815-826.
21. W.W. Moses, S.A. Payne, W.S. Choong, *Scintillator Non-Proportionality: Present Understanding and Future Challenges*, IEEE Trans. Nucl. Sci. 55(3) (2008) 1049-1053.
22. B.D. Rooney, J.D. Valentine, *Scintillator Light Yield Non-proportionality: Calculating Photon Response Using Measured Electron Response*, IEEE Trans. Nucl. Sci. 44(3) (1997) 509-516.
23. P. Lecoq, M. Korzik, A. Vasiliev, *Can transient phenomena help improving time resolution in scintillators?*, IEEE Trans. Nucl. Sciences, Vol. 61, NO. 1, February 2014
24. J.Q. Grim et al., *Continuous-Wave Biexciton Lasing at Room Temperature Using Solution-Processed Quantum Wells,* Nat. Nanotechnol. **9,** 891–895 (2014)
25. P. Lecoq, *Metamaterials for novel X- or γ-ray detector designs*, 2008 IEEE Nuclear Science Symposium Conference Record, N07-1, pp.680-684

26. A. Annekov, M. Korzhik, P. Lecoq, et al., *Slow components and afterglow in PWO crystal scintillation*, Nucl. Instrum. Meth. A 403 (1998) 302-312.
27. US Patent 5521387
28. Chen Lingyan, Du Jie, Wang Liming, Xiang Kaihua, *An investigation of radiation damage induced by hydroxyl and oxygen impurities in BaF2 crystal*, Scintillator and Phospor Materials MRS Proceeding 348 (1994) 447-454.
29. A. Annenkov, M. Korzhik, P. Lecoq, *Lead tungstate scintillation material*, Nucl. Instrum. Meth. A 490 (2002) 30–50.
30. M. Nikl et al., *Defect Engineering in Ce-Doped Aluminum Garnet Single Crystal Scintillators,* Cryst. Growth Des. 2014, 14, 4827−4833, dx.doi.org/10.1021/cg501005s
31. Z.W. Yin, Q. Deng, D.S. Yan, *Research and Development Works on PbWO4 Crystals in Shanghai Institute of Ceramics*, Proc. 5th Int. Conf. Inorganic Scintillators and their Applications, Moscow, Apr 16-20, 1999, pp. 206-211.
32. S. Kyropolos, F. Zeits, *Ein Verfahren zur Herstellung grosser Kristalle*. Anorg. Allgem. Chem. 154 (1926) 308–313.
33. J. Czochralski, *Ein neues Verfahren zur Messung der Kristallisationsgeschwindigkeit der Metalle*, Z. Phys. Chem. 92 (1918) 219-224.
34. B. Hautefeuille et al., J. Crystal Growth 289 (2006) 172.

Chapter 4
Gaseous Detectors

H. J. Hilke and W. Riegler

4.1 Introduction

All gaseous detectors signal the passage of charged particles by gathering the electrons from the ion pairs produced in the gas, usually after some amplification. The history of the gas detectors starts with the counter described by Rutherford and Geiger in 1908 [1]. It consisted of a cylindrical metallic tube filled with air or other simple gases at some 5 Torr and with a 0.45 mm diameter wire along its axis. The negative high voltage on the tube with respect to the wire was adjusted to below the discharge limit. With a gas gain of a few 10^3, only α-particles could be detected as current pulses with an electrometer. This counter was the first electronic counter, following the optical counting of light flashes in the study of radioactive substances with scintillating crystals. A major step was taken when Geiger found that by replacing the anode wire by a needle with a fme pin, electrons could also be detected [2]. These *needle counters* became the main particle counter for years. Already in 1924, Greinacher started using electronic tubes to amplify the signals [3].

The *Geiger-Mueller-counter* was first described in 1928 [4]: it produced strong signals independent of the primary ionization. Used with rare gases, these counters required load resistances of $10^8 - 10^9$ Ohm to avoid continuous discharges, resulting in dead times of $10^{-3} - 10^{-4}$ s. Later, external circuits were introduced to shorten the dead time. The real progress, however, brought the discovery in 1935 by Trost [5] that the addition of alcohol quenches the gas discharges internally,

The author H. J. Hilke is deceased at the time of publication.

H. J. Hilke · W. Riegler (✉)
CERN, Geneva, Switzerland
e-mail: Werner.Riegler@cern.ch

C. W. Fabjan, H. Schopper (eds.), *Particle Physics Reference Library*,
https://doi.org/10.1007/978-3-030-35318-6_4

permitting low load resistances and thus high counting rates. Cosmic ray physics in particular profited from systems of such counters used with electron tube coincidence circuits. It took a number of years to understand the basic processes in different gases and under various operation conditions.

Proportional counters regained interest, when the development of more sensitive readout electronics permitted energy determination. In the second half of the 1940s, however, the demand for faster counters with longer lifetime and higher sensitivity initiated a move towards scintillation techniques, which saw a rapid development, especially after the introduction of the photomultiplier, soon providing fast response and time resolutions below 10^{-8} s. On the gas detector side, only the novel technique *of parallel plate counters* [6] could compete, with time resolutions down to 10^{-10} s, however with lower rate capability. A detailed account of the developments up to the 1950s can be found in [7].

The field of gas detectors was revived with the introduction of the *multiwire proportional chamber* by Charpak in 1968 [8] and shortly afterwards with the extension by two groups to *drift chambers* with different geometries [9, 10]. The following decades saw a rapid development of the techniques, especially in high energy physics but also for nuclear physics and other fields. An additional major R&D effort was triggered in the 1990s by the requirements for the LHC: extreme particle rates and radiation hardness. Solutions demanded very careful choice of gas fillings as well as of construction materials and methods. Gas detectors were and are still used mainly for tracking but also in calorimeters, Cherenkov counters and the detection of transition radiation. Only in the layers closest to the interaction points in accelerator experiments and in other applications where spatial resolution is the prime requirement, finely grained silicon detectors have taken over as first choice. Most of the detector developments were made possible only by the extremely rapid progress in the field of electronics, with respect to miniaturization, integration density, cost and radiation hardness.

Powerful simulation programs have been developed in the past decades and have been widely used in the development and optimization of gas detectors. The program *Garfield* [11] calculates electric fields, electron and ion trajectories and induced signals. The program *Heed* [12] describes primary ionization produced by fast particles in gases and the program *Magboltz* [13] electron transport properties in gas mixtures. The agreement of simulation and measurement has become impressive.

We shall at several occasions refer to designs and studies from the LHC experiments. Recent detailed reports them may be found in [14–17]. The development of the last years can well be followed in the Proceedings of the Vienna Conference on Instrumentation initiated in 1977 as Wire Chamber Conference on a tri-annual basis [18] and of the annual IEEE Nuclear Science Symposia.

The following sections will start with a description of the basic processes in gaseous detectors: ionization of the gas by charged particles (Sect. 4.2.1), transport of electrons and ions in electric and magnetic fields (Sect. 4.2.2), avalanche amplification in high electric fields (Sect. 4.2.3), formation of the readout signals (Sect. 4.2.4) and ‘ageing’ of detectors under irradiation (Sect. 4.2.6). A discussion

of major directions of detector design and performance follows in Sect. 4.3: Single-wire tubes (Sect. 4.3.1), Multi-Wire-Proportional Chambers (Sect. 4.3.2), Drift Chambers (Sect. 4.3.3), Resistive Plate Chambers (Sect. 4.3.4) and Micropattern Devices (Sect. 4.3.5).

4.2 Basic Processes

As most processes depend on the velocity of a particle, we shall often state numerical values for minimum ionizing particles (*mip*), i.e. for $\gamma = 3 - 4$.

4.2.1 Gas Ionization by Charged Particles

The passage of charged particles through a gas is signaled by the production of electron/ion pairs along its path. The electrons are attracted by electrodes on positive potential, in the vicinity of which they are usually amplified in a avalanche process. We give a short summary of the various aspects of the ionization processes, following to some extent [19].

4.2.1.1 Primary Clusters

The ionizing collisions of the particle are occurring randomly with a *mean distance* λ, related to the ionization cross-section per electron σ_{I} and the electron density N_{e} of the gas:

$$1/\lambda = N_{\mathrm{e}}\, \sigma_{\mathrm{I}}. \tag{4.1}$$

The number k of ionizing collisions on a path length L thus follows a Poisson distribution with mean L/λ:

$$\mathrm{P}\,(k|L,\ \ \lambda) = \left((L/\lambda)^k/k!\right) \exp\left(-L/\lambda\right). \tag{4.2}$$

The probability to have no ionization in L is

$$\mathrm{P}\,(0|L,\ \ \lambda) = \exp\left(-L/\lambda\right). \tag{4.3}$$

This relation is used to determine λ and defines the *inefficiency* of a counter measuring a track length L, if it is sensitive to a single primary electron.

The probability distribution f(l)dl for a free flight pass l between two ionizing collisions—i.e. the probability of no ionization in l and one in dl-is an exponential,

$$\mathrm{f}(l)\mathrm{d}l = (\mathrm{d}l/\lambda)\exp(-l/\lambda), \tag{4.4}$$

i.e. short distances are favoured.

An electron ejected in a primary collision on atom A may have enough energy to ionize one or more other atoms. Thus *clusters* of two or more electrons are formed by *secondary ionization*. These clusters are mostly rather localized, as the ejection energy is usually low and results in a short range. High ejection energies for so-called *δ-electrons* are rare, their average number per cm is approximately inversely proportional to energy:

$$P(E > E_0) - y/\left(\beta^2 \mathrm{E}_0\right)/\mathrm{cm}, \tag{4.5}$$

with E_0 in keV, and $y = 0.114$ for Ar, and $y = 0.064$ for Ne [20]; β= particle velocity/speed of light in vacuum. Thus, in Ar, for β~1 and $E_0 = 10\,\mathrm{keV}$, $P = 0.011$/cm, i.e., on average one collision with $E > 10\,\mathrm{keV}$ will occur on a track length of 90 cm. The range of a 10 keV electron is about 1.4 mm. The range decreases very rapidly with decreasing energy and is only about 30 μm for a 1 keV electron.

It tums out that, although the majority of the 'clusters' consist of a single electron, clusters of size >1 contribute significantly to the *mean total number n_T of electrons produced per cm*, so that n_T is significantly larger than n_p, the *mean number of primary clusters per cm*. Table 4.1 gives experimental values for some of the common detector gases. In Ar at NTP one finds on average $n_\mathrm{p} = 26$ and $n_\mathrm{T} = 100$ electrons/cm for a minimum ionizing particle, where n_T depends on the volume around the track taken into account. The most probable value for the total ionization is $n_\mathrm{mp} = 42$ electrons/cm. The big difference between n_T and n_mp is an indication of the long tail of the distribution.

Table 4.1 Properties of gases at 20°C, 1 atm

Gas	n_p	n_T	w [eV]	E_I [eV]	E_x [eV]	p [mg/cm^3]
He	4.8	7.8	45	24.5	19.8	0.166
Ne	13	50	30	21.6	16.7	0.84
Ar	25	100	26	15.7	11.6	1.66
Xe	41	312	22	12.1	8.4	5.50
CH_4	37	54	30	12.6	8.8	0.67
CO_2	35	100	34	13.8	7.0	1.84
i-butane	90	220	26	10.6	6.5	2.49
CF_4	63	120	54	16.0	10.0	3.78

n_p, n_T mean primary and total number of electron-ion pairs per cm; w: average energy dissipated per ion pair; E_I, E_x: lowest ionization and excitation energy [21]

4.2.1.2 Cluster Size Distribution

The space resolution in gaseous detectors is influenced not only by the Poisson distribution of the primary clusters along the track but also by the *cluster size distribution*, i.e., by the number of electrons in each cluster and their spatial extent. Little was known experimentally (except for some measurements in cloud chambers) until the first detailed theoretical study [22] for Ar at 1 atm and 20°C. Based on the experimental cross-sections for photo absorption, the oscillator strengths and the complex dielectric constants are calculated and from this the distribution of energy transfers larger than the ionization energy (15.7 eV). Finally, a detailed list is obtained for the distribution of cluster sizes for $\gamma = 4$ and $\gamma = 1000$, to estimate the relativistic rise. A cut of 15 keV was applied to the maximum energy transfer, thus concentrating on the local energy deposition. The mean number of clusters is found to be $n_p = 26.6$/cm at $\gamma = 4$, and 35/cm at $\gamma = 1000$. For a MIP, 80.2% of the clusters are found to contain 1 electron, 7.7% two electrons, 2% three electrons, and 1.4% more than 20 electrons.

Several years later, a detailed experimental study of several gases is reported in [23]. For Ar, 66/15/6 and 1.1% of the clusters are found to contain 1/2/3 and $\geq$20 electrons, respectively. The values for low cluster sizes are quite different from the calculated values mentioned above and the calculated bump around 10 electrons is not seen in the measurement, see Fig. 4.1. The authors suggest as a possible explanation that one assumption made in the simulation may not be appropriate, namely that the absorption of virtual photons can be treated like that of real photons, which also leads to the bump at the L-absorption edge. A simpler model starting from measured spectra of electrons ejected in ionizing collisions gives good agreement with the measurements, in particular for the probability of small cluster sizes.

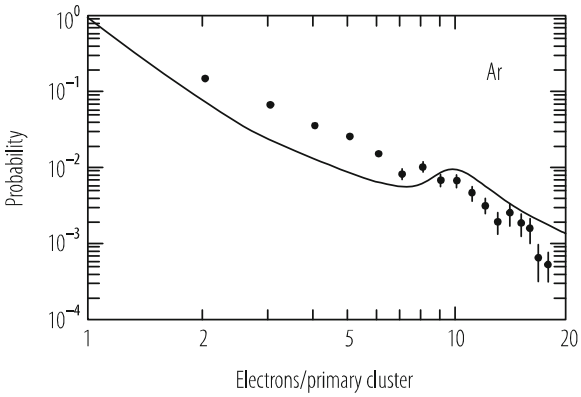


Fig. 4.1 Cluster size distribution: simulation for Ar (continuous line) [22] and measurements in Ar/CH_4 (90/10%) [23]

Space resolution in drift chambers is influenced by the clustering in several ways. The arrival time of the first n electrons, where n times gas amplification is the threshold for the electronics, depends both on the spatial distribution of the clusters and the cluster size. For large clusters, δ-electrons, ionization may extend far off the trajectory.

4.2.1.3 Total Number of Ion Pairs

The detector response is related to the cluster statistics but also to the total ionization n_T, e.g., in energy measurements. A quantity W has been introduced to denote the average energy lost by the ionizing particle for the creation of one ion pair:

$$W = E_i/n_E, \tag{4.6}$$

where E_i is the initial kinetic energy and n_E the average total number of ion pairs *after full dissipation of* E_j.

Measurements of W by total absorption of low energy particles show that it is practically independent of energy above a few keV for electrons and above a few MeV for α-particles. For that reason the differential value w, defined by

$$w = x < dE/dx > / < n_T > \tag{4.7}$$

may be used alternatively, as is usually done in Particle Physics, to relate the average total number of ion pairs n_T, created in the track segment of length x, to the average energy lost by the ionizing particle. For relativistic particles, dE/dx can not be obtained directly from the difference of initial and final energy (about 270 keV/m for $\gamma = 4$ in Ar), as it is below the measurement resolution. Therefore, w has to be extrapolated from measurements of lower energy particles. For the rare gases one finds $w/I = 1.7 - 1.8$ and for common molecular gases $w/I = 2.1 - 2.5$, where I is the *ionization potential*, indicating the significant fraction of dE/dx spent on excitation. Values for photons and electrons are the same, also for α particles in rare gases; in some organic vapours they may be up to 15% higher for α-particles. At low energy, close to the ionization potential, W increases.

In gas mixtures, where an excitation level of component A is higher than I of component B, excited molecules of A often produce a substantial increase in ionization, as has e.g. been observed even with minute impurities in He and Ne: adding 0.13% of Ar to He changed W from 41.3 to 29.7 eV per ion pair. This energy transfer is called *Jesse effect* or *Penning effect*, if metastable states are involved. It is also possible that more than one electron is ejected from a single atom, e.g., by Auger effect following inner shell ionization.

The distribution of n_T in small gas segments is very broad, see an example in Fig. 4.2 [24]. To describe the measurement result, it is thus appropriate to use the *most probable value* instead of the *mean*, since the mean of a small number of measurements will depend strongly on some events from the long tail

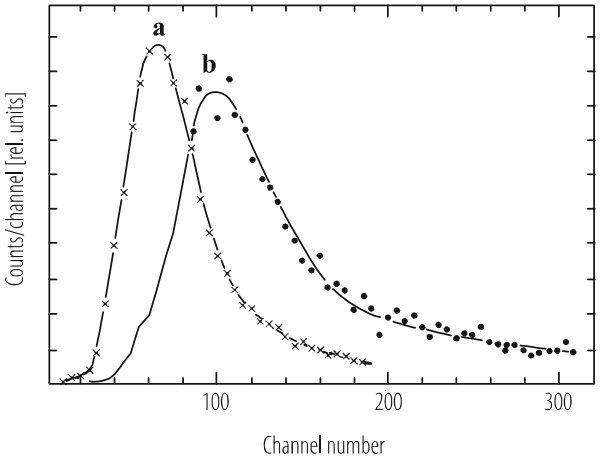


Fig. 4.2 Measured pulse height distribution for 2.3 cm in Ar/CH_4 at 1 atm: (**a**) protons 3 GeV/c, (**b**) electrons 2 GeV/c [24]

of the distribution. The measured pulse height spectrum contains some additional broadening from the fluctuations of the avalanche process. For a mixture of Ar and 5% CH_4, a most probable value of $n_{\mathrm{mp}} = 48$ ion pairs/cm was found for minimum ionizing particles [25].

4.2.1.4 Dependence of Energy Deposit on Particle Velocity

As mentioned above, for position detectors one is interested in the ionization deposited close to the particle trajectory. The Bethe-Bloch formula for dE/dx describes instead the average total energy loss from the incoming particle, including the energy spent on the ejection of energetic δ-electrons which deposit ionization far from the trajectory. To describe the local energy deposit, it is sensible to exclude the contribution from these energetic δ-electrons. This is done by replacing the maximum possible energy transfer T_{max} by a cut-off energy $T_{\mathrm{cut}} << T_{\mathrm{max}}$. This energy cut-off will depend on the experimental conditions and may lie between 30 keV and 1 MeV (in a magnetic field) [19] One then obtains the modified Bethe-Bloch formula for the *mean restricted energy deposit* [20, 21] (see also Chap. 2)

$$\mathrm{d}E/\mathrm{d}x_{\mathrm{restricted}} = K z^2 (Z/A) \left(1/\beta^2\right) \left[0.5 \ln \left(2 m_{\mathrm{e}} c^2 \beta^2 \gamma^2 T_{\mathrm{cut}}/I^2\right) - \beta^2/2 - \delta/2\right], \tag{4.8}$$

with $K = 4\pi N_{\mathrm{A}} r_{\mathrm{e}}^2 m_{\mathrm{e}} c^2$, N_{A} = Avogadro constant, m_{e}, r_{e}= mass and classical radius of the electron.

Due to the cut-off, this relation applies not only to heavy particles but also to ionization by electrons [19]. The minimum $\mathrm{d}E/\mathrm{d}x$ deposited by a *minimum ionizing particle* (mip) still lies around $\gamma = 3 - 4$, with $\delta = 0$. For $\beta \to 1$, the *density correction* δ approaches

$$\delta \to 2 \ln \left(h\nu_\mathrm{p} \gamma / I \right) - 1, \tag{4.9}$$

$h\nu_\mathrm{p}$ being the quantum energy of the plasma oscillation of the medium. The restricted energy deposit then reaches a constant value, the *Fermiplateau*, the δ-term compensating the $\ln\gamma$ term:

$$\mathrm{d}E/\mathrm{d}x_\mathrm{restricted} \to \mathrm{P}^2 \, (Z/A) \, 0.5 \ln \left[2mc^2 T_\mathrm{cut} / \left(h\nu_\mathrm{p} \right)^2 \right]. \tag{4.10}$$

In Ar one obtains for the ratio R of energy deposit on the Fermi plateau to the minimum deposit $R = 1.60$, 1.54, and 1.48 for a cut-off $T_\mathrm{cut} = 30$,150 and 1000 keV, respectively [19]. A precise determination of R requires a good estimate of T_cut.

To use the β-dependence of $\mathrm{d}E/\mathrm{d}x$ for particle identification, one has to measure many samples and take their *truncated mean*, e.g., the mean of the lowest 50% pulse heights, to be insensitive to the long tail and to obtain an approximation to the most probable value. See Chap. 2 for details.

4.2.2 Transport of Electrons and Ions

4.2.2.1 Drift Velocities

On the microscopic scale, electrons or ions drifting through a gas are scattered on the gas molecules. In a homogenous electric field E they will acquire a constant *drift velocity* u in the E field direction or, in the presence of an additional magnetic field B, in a direction determined by both fields. Their drift velocities are much smaller than their *instantaneous velocities* c between collisions. Electrons and ions will behave quite differently because of their mass difference.

In the chapters on drift velocities and diffusion we shall follow the argumentation developed in [19]. A relatively simple derivation brings out the main characteristics and does describe a number of experimental results with good approximation. The main approximation of the simple models is to take a single velocity c to represent the motion between collisions. In reality, these velocities c are distributed around a mean value. The shape of the distribution depends on the variation of cross-section and energy loss with the collision velocity. The rigorous theory takes these distributions into account. For lowest velocities there is only elastic scattering, for higher energies various inelastic processes contribute. The elastic and the inelastic spectrum may be described by a single *effective cross-section* $\sigma(c)$ combining the various processes, sometimes called *momentum transfer cross-section*, and by the *average fractional energy loss* $\Delta(c)$ per collision.

Collision cross-sections σ have in some cases been measured directly. Often, however, σ as well as $\Delta(c)$ have to be deduced from measurements of $u(E)$, the dependence of on E, and of diffusion, based on some assumptions on the excitation functions. The consistency of the methods, when applied to other gas mixtures, has improved over the years and is presently very good in a number of practical cases, in particular for the Magboltz simulation [13]; for a comparison of experiments with various models see e.g. [26].

Drift of Electrons

Because of their small mass, electrons will scatter isotropically in a collision and forget any preferential direction. They will acquire a *drift velocity* u given by the product of the acceleration eE/m and the average time τ between collisions

$$u = eE\tau/m. \tag{4.11}$$

Instead of, the notion of *mobility* μ is often used, with μ defined by

$$u = \mu E \rightarrow \mu = e\tau/m. \tag{4.12}$$

Over a drift distance x there will be a balance between the *collision loss* $\Delta\varepsilon_E$ and the energy picked up:

$$(x/u)\,(1/\tau)\,\Delta\varepsilon_E = eEx. \tag{4.13}$$

Here ε_E is the energy gained between collisions, Δ the average fraction of the energy lost in a collision, and $(x/u)(1/\tau)$ the number of collisions on a distance x.

For an *instantaneous velocity* c, the mean time τ between collisions is related to the collision cross- section σ and the number density N of gas molecules by

$$1/\tau = N\sigma c. \tag{4.14}$$

The total energy ε of the electron is given by

$$(m/2)\,c^2 = \epsilon = \epsilon_E + (3/2)\,kT, \tag{4.15}$$

including the thermal energy.

In the approximation $e >> (3/2)kT$, which is often fulfilled for drift of electrons in particle detectors, one obtains

$$\begin{aligned} u^2 &= (eE/mN\sigma)\,\sqrt{(\Delta/2)}\,, \text{ and} \\ c^2 &= (eE/mN\sigma)\,\sqrt{(2/\Delta)} \quad \text{for } \varepsilon - \varepsilon_E >> (3/2)\,kT. \end{aligned} \tag{4.16}$$

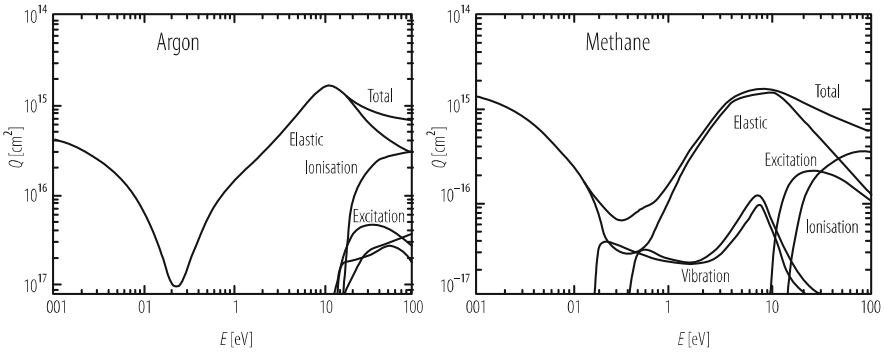


Fig. 4.3 Electron collision cross-sections for Argon and Methane used in Magboltz [13, 27]

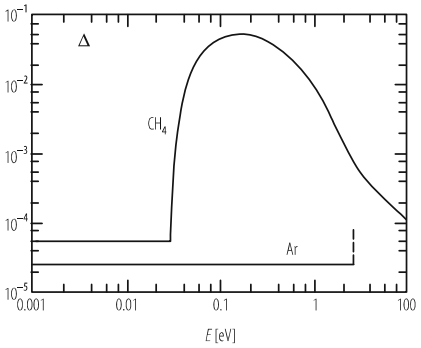


Fig. 4.4 The fraction Δ of energy lost per collision as function of mean energy ε of the electron [28]

The rigorous theory assuming a Dryvestem distribution for the random velocities c adds a multiplication factor of 0.85 to the right sides.

It is important to note that E and N only appear as E/N, the reduced electric field, for which often a special unit is used: one *Townsend* (*Td*) with 1 Td $=10^{-17}$ Vcm^2.

The important role of σ and Δ is obvious; both depend on ε. Below the first excitation level the scattering is elastic and $\Delta - 2m/M - 10^{-4}$ for electrons scattered on gas molecules with mass M. For a high drift speed a small σ is required. Figure 4.3 shows the cross-sections σ for Ar and CH_4. A pronounced minimum, the so-called 'Ramsauer dip' is clearly visible. It leads to high drift velocities in Ar - CH_4 mixtures at low E-values. TPCs take advantage of this.

10^{-17} Vcm^2 =250 V/cm atm at 20°C.

From precise measurements of drift velocity u (to 1%) and longitudinal diffusion D/μ (to 3–5%), σ and Δ have been deduced for some gases [28]. The consistency of the calculated values with measurements of u and D/μ in various other gas *mixtures* gives confidence in the method. Figure 4.4 presents calculated values for Δ as function of ε. Figure 4.5 shows $\varepsilon_k = (2/3)\varepsilon$ derived in the same way in another

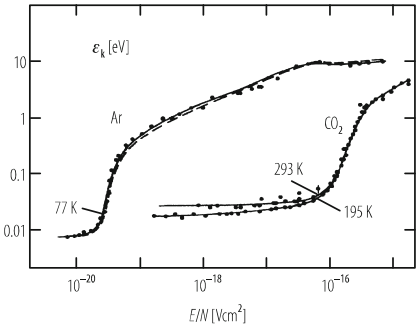


Fig. 4.5 Values for the electron energy ε derived from diffusion measurements as function of the reduced electrical field [29]

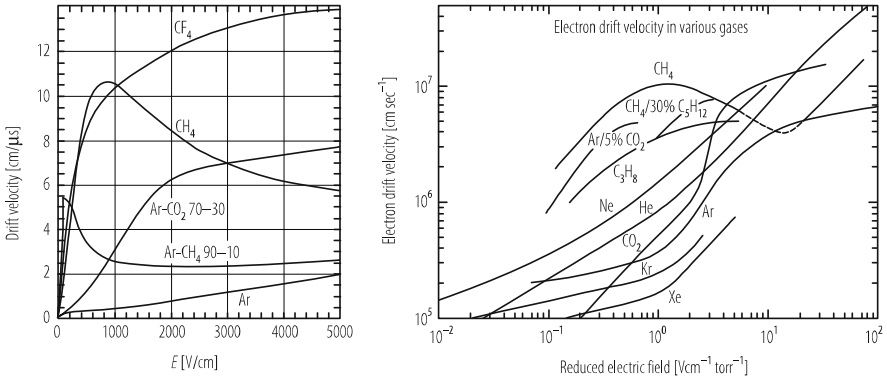


Fig. 4.6 Some examples measured electron drift velocities. (*left*) [30], (*rights*) [31]

study for two extremes, *cold* CO_2 and *hot* Ar [29]. Gases are denoted as *cold*, if ε stays close to the thermal energy $(3/2)kT$ in the fields under consideration. This is the case for gases with vibrational and rotational energy levels, the excitation of which causes inelastic energy losses to the drifting electrons. Cold gases are of interest since they exhibit the smallest possible diffusion.

For *gas mixtures* with number densities $n_i (N = \Sigma\, n_i)$, the effective σ and Δ are given by

$$\sigma = \Sigma n_i \sigma_1 / N, \text{ and}$$

$$\Delta\sigma = \Sigma n_i \Delta_i \sigma_1 / N. \tag{4.17}$$

At low E, drift velocities rise with electric field. Some (e.g. CH_4 and Ar Ar - CH_4 mixtures) go through a maximum, decrease and may rise again. Drift velocities are shown in Fig. 4.6 for some gases [30, 31].

Drift of Ions

Ions of mass m_i acquire the same amount of energy between two collisions as electrons but they lose a large fraction of it in the next collision and their random energy thus remains close to thermal energy. On the other hand the direction of their motion is largely maintained. The result is a much smaller diffusion compared to electrons and constant mobility up to high fields (to ~20 kV/cm atm for A^+ ions in A). In the *approximation for low E field*, the random velocity is considered thermal, i.e. the relative velocity c_rel between the ion and the gas molecules of mass M, which determines τ, is

$$c_\mathrm{rel}^2 = c_\mathrm{ion}^2 + c_\mathrm{gas}^2 = 3kT\left(m_\mathrm{i}^{-1} + M^{-1}\right) \tag{4.18}$$

An argumentation similar to the one followed for electrons [19] leads to

$$u = \left(m_\mathrm{i}^{-1} + M^{-1}\right)^{1/2} (1/3kT)^{1/2} eE/\left(N\sigma\right) \tag{4.19}$$

The ion drift velocity at *low fields* is thus proportional to the electric field. Typical values at 1 atm are around $u = 4$ m/s for $E = 200$ V/cm, to be compared with a thermal velocity around 500 m/s.

In the other extreme of *very high fields*, where thermal motion can be neglected, one finds the drift velocity being proportional to the square root of E. Measurements on noble gas ions [32] in their own gas clearly show both limits with a transition between them at about $15 - 50$ kV/cm atm; see Fig. 4.7. As typical drift fields in drift chambers are a few hundred V/(cm atm), the 'low field approximation' is usually applicable, except in the amplification region.

In a *gas mixture* it is expected that the component with the lowest ionization energy will rapidly become the drifting ion, independently of which atom was ionized in the first place. The charge transfer cross-section is in fact of similar magnitude as the other ion molecule scattering cross-sections. Even impurities rather low concentration might thus participate in the ion migration.

Magnetic Field Effects

A simple macroscopic argumentation introduced by Langevin produces results which are a good approximation in many practical cases.

The motion of a charged particle is described by

$$m\mathrm{d}\mathbf{u}/\mathrm{d}t = e\mathbf{E} + e\,[\mathbf{u}\ \mathrm{x}\ \mathbf{B}] - k\ \mathbf{u}, \tag{4.20}$$

where m, e and $\mathbf{u}$ are the particle's mass, charge and velocity vector, respectively; $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic field vectors; k describes a frictional force proportional to $-\mathbf{u}$.

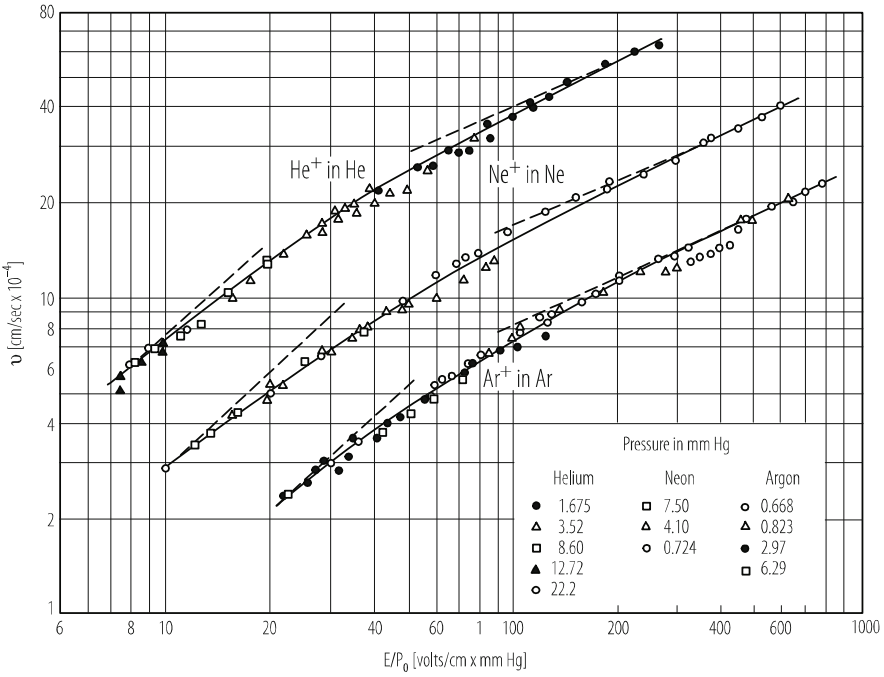


Fig. 4.7 Drift velocities of singly charged ions of noble gases [32]

In the steady state d**u**/dt = 0 and

$$\mathbf{u}/\tau - (e/m)\,[\mathbf{u}\ \mathrm{x}\ \mathbf{B}] = (e/m)\,\mathbf{E}, \tag{4.21}$$

with τ = m/k. The solution for **u** is

$$\mathbf{u} = (e/m)\,\tau \mid \mathbf{E} \mid \left(1/\left(1+\omega^2\tau^2\right)\right)\left\{\mathbf{E}^* + \omega\tau\left[\mathbf{E}^*\ \mathrm{x}\ \mathbf{B}^*\right] + \omega^2\tau^2\,(\mathbf{E}^*\mathbf{B}^*)\,\mathbf{B}^*\right\}, \tag{4.22}$$

where $\omega = (e/m) \mid \mathbf{B}\mid$, and ω carries the sign of e and $\mathbf{E}^*$ and $\mathbf{B}^*$ are unit vectors.

For ions, $\omega\tau \approx 10^{-10}$ Therefore, magnetic fields have negligible effect on ion drift.

For electrons, **u** is along **E**, if $B = 0$, with

$$\mathbf{u} = (e/m)\,\tau\mathbf{E}. \tag{4.23}$$

This is the same relation as the one derived from the microscopic picture (4.11), which provides the interpretation of τ as the *mean time between collisions*.

For large $\omega\tau$, **u** tends to be along **B**, but if **EB** = 0, large $\omega\tau$ tums u in the direction of **ExB**.

Two *special cases* are of practical interest for *electron drift*:

E orthogonal to B

With $\mathbf{EB} = 0$ and choosing $\mathbf{E} = (E_x,\ 0, 0)$ and $\mathbf{B} = (\mathrm{O},\ 0,\ B_z)$, we get

$$\begin{aligned} u_x &= (e/m)\,\tau \mid \mathbf{E} \mid / \left(1 + \omega^2\tau^2\right)\ , \\ u_y &= -\,(e/m)\,\tau\omega\tau \mid \mathbf{E} \mid / \left(1 + \omega^2\tau^2\right), \\ u_z &= 0, \end{aligned} \tag{4.24}$$

and

$$\mathrm{tg}\psi = u_y/u_x = -\omega\tau. \tag{4.25}$$

The latter relation is used to determine $\omega\tau$, i.e. τ, from a measurement of the drift angle ψ, the so-called *Lorentz angle*. In detectors, this angle increases the spread of arrival times and sometimes also the spatial spread. A small $\omega\tau$ would, therefore, be an advantage but good momentum resolution requires usually a strong B.

The absolute value of $\mathbf{u}$ is

$$\mid \mathbf{u} \mid = (e/m)\,\tau \mid \mathbf{E} \mid \left(1 + \omega^2\tau^2\right)^{-1/2} = (e/m)\,\tau \mid \mathbf{E} \mid \cos\psi. \tag{4.26}$$

This means that, independent of the drift direction, the component of $\mathbf{E}$ along $\mathbf{u}$ determines the drift velocity (Tonks' theorem). This is well verified experimentally.

E Nearly Parallel to B

This is the case in the Time Projection Chamber (TPC). Assuming E along z and the components $|B_X|$ and $|B_y| << |B_z|$, one finds in first order

$$\begin{aligned} u_x/u_z &= \left(-\omega\tau\, B_y/B_z + \omega^2\tau^2 \mathrm{B_x/B_z}\right) / \left(1 + \omega^2\tau^2\right), \text{ and} \\ u_y/u_z &= \left(\omega x \mathrm{B_x/B_z} + \omega^2\tau^2 \mathrm{B_y/B_z}\right) / \left(1 + \omega^2\tau^2\right). \end{aligned} \tag{4.27}$$

In a TPC this will produce a displacement after a drift length L of $\delta_x = Lu_x/u_z$ and $\delta_y = Lu_y/u_z$ From measurements with both field polarities and different fields, B_X, B_y and τ can be determined.

If B_x and B_y can be neglected with respect to B_z, u_z remains unaffected by B.

4.2.2.2 Diffusion

Due to the random nature of the collisions, the individual drift velocity of an electron or ion deviates from the average. In the simplest case of isotropic deviations, a point-like cloud starting its drift at $t = 0$ from the origin in the z direction will at time t assume a Gaussian density distribution

$$N = (4\pi Dt)^{-3/2} \exp\left(-r^2/(4Dt)\right), \tag{4.28}$$

with $r^2 = x^2 + y^2 + (z - ut)^2$, D being the *diffusion coefficient*. In any direction from the cloud centre, the mean squared deviation of the electrons is

$$\sigma_{\mathrm{I}} = (2Dt)^{1/2} = (2Dz/u)^{1/2} = D^* z^{1/2}. \tag{4.29}$$

with D^* called *diffusion constant*. In terms of the microscopic picture, D is given by

$$D = \lambda^2/(3\tau) = c\lambda/3 = c^2\tau/3 = (2/3)(\varepsilon/m)\tau, \tag{4.30}$$

with λ being the *mean free path*, $\lambda = c\tau$, and ε the *mean energy*.

With the *mobility* μ defined by

$$\mu = (e/m)\tau, \tag{4.31}$$

the mean energy ε can be determined by a measurement of the ratio D/μ:

$$\varepsilon = (3/2)(D/\mu)e. \tag{4.32}$$

Instead of ε, the *characteristic energy* $\varepsilon_k = (2/3)\varepsilon$ is often used.

The *diffusion width* σ_x of an initially point-like electron cloud having drifted a distance L is determined by the electron energy ε:

$$\sigma_X^2 = 2Dt \quad = 2DL/(\mu E) = (4/3)\varepsilon L/(eE) \tag{4.33}$$

This relation is used for the *determination* of D *and* ε.

For a good spatial resolution in drift chambers, a low electron energy and high electric fields are required. The lower limit for ε is the thermal energy $\varepsilon_{\mathrm{th}} = (3/2)kT$. In this limit, the relationship known as *Einstein or Nernst-Townsend formula* follows:

$$D/\mu = kT/e. \tag{4.34}$$

The *minimum diffusion width* is thus

$$\sigma_{x,\mathrm{mm}}{}^2 = (kT/e)(2L/E). \tag{4.35}$$

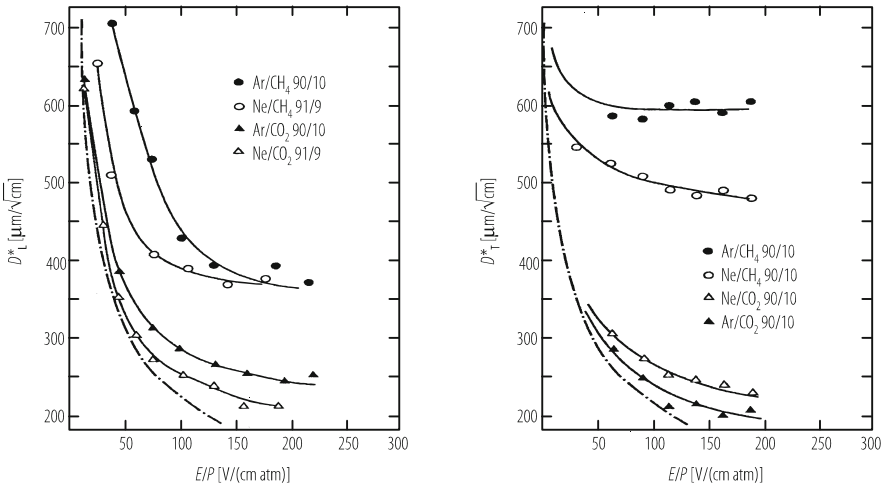


Fig. 4.8 Longitudinal and transverse diffusion constants for low electric fields [33]. The dash-dotted line denotes the thermal limit

As can be seen in Fig. 4.8, this minimum is approached for 'cold gases' like Ar/CO_2 up to E~150 V/cm at 1 atm, for 'hot gases' like Ar/CH_4 only for much lower fields.

Anisotropic Diffusion

So far, we have assumed isotropic diffusion. In 1967 it was found experimentally [34], that the *longitudinal diffusion* D_L along E can be different from the *transversal diffusion* D_T Subsequently it has been established that this is usually the case.

For ions this anisotropy occurs only at high E. As in a collision ions retain their direction to a large extent, the instantaneous velocity has a preferential direction along E. This causes diffusion to be larger longitudinally. However, this high field region is beyond the drift fields used in practical detectors.

For electrons a semi-quantitative treatment [35], restricted to energy loss by elastic collisions, shows that

$$D_L/D_T = (1+\gamma)\,/\,(1+2\gamma) \;\text{ with }\; \gamma = (\varepsilon_0/v_0)\,(\delta v/\delta\epsilon)\,. \tag{4.36}$$

It follows that longitudinal and transversal diffusion will be different, if the collision rate v depends on the electron energy ε.

Figures 4.8, 4.9, 4.10 show measured diffusion for a drift of 1 cm for some common gas mixtures [30, 33, 36]. Simulated diffusion curves are compiled in [33].

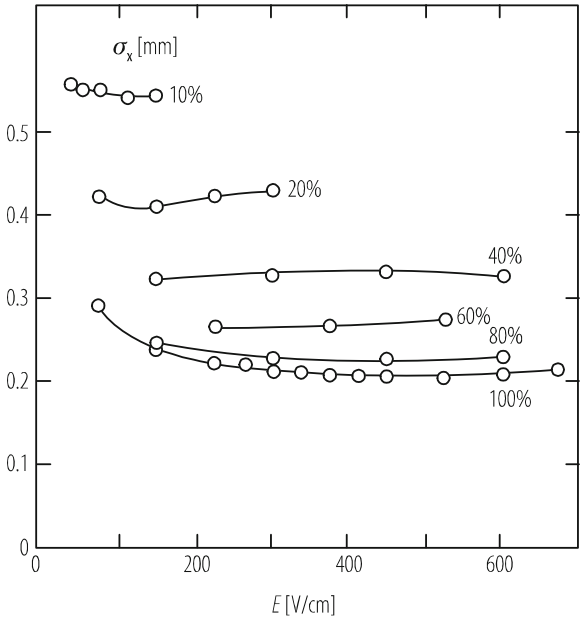


Fig. 4.9 Transverse diffusion for 1 cm drift in Ar/CH_4 mixtures; CH_4 % is indicated [36].

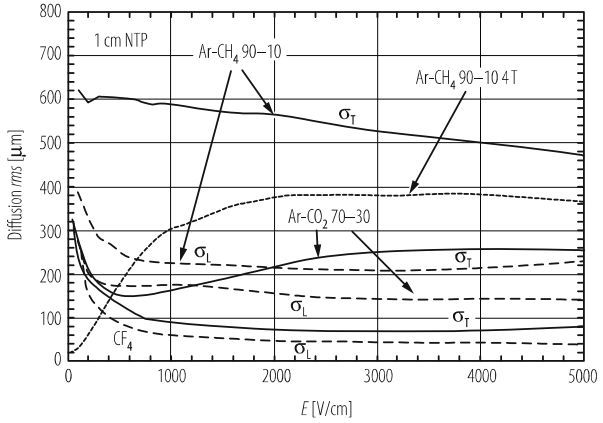


Fig. 4.10 Transverse and longitudinal diffusion for 1 cm drift up to higher E fields [30]

A *magnetic field B* along z will cause electrons to move in circles in the x-y projections in between collisions. The random propagation is diminished and the transverse diffusion will be reduced:

$$D_{\mathrm{T}}(\omega)/D_{\mathrm{T}}(0) = 1/\left(1+\omega^2\tau^2\right). \tag{4.37}$$

This reduction is essential for most TPCs with their long drift distances.

A more rigorous treatment of averages [19] shows that different ratios apply to low and high B:

$$\begin{aligned} D(0)/D(B) &= 1 + \omega^2\tau_1^2 \quad \text{for low } B, \text{ and} \\ D(0)/D(B) &= C + \omega^2\tau_2^2 \quad \text{for high } B. \end{aligned} \tag{4.38}$$

This behaviour was indeed verified [37], by measuring $D(B)$ over a wide range of B. In an Ar/CH_4 (91/9%) mixture the data could be fitted with $\tau_1 = 40$ ps, $\tau_2 = 27$ ps and $C = 2.8$. The high field behaviour is approached above about 3 kg, close to $\omega\tau = 1$.

The longitudinal diffusion remains unchanged: $D_L(\omega) = D_L(0)$.

The effects of E and B combine if both fields are present.

4.2.2.3 Electron Attachment

In the presence of electronegative components or impurities in the gas mixture, the drifting electrons may be absorbed by the formation of negative ions. Halogenides (e.g. CF_4) and oxygen have particularly strong electron affinities. Two-body and three-body attachment processes are distinguished [38].

In the *two-body process*, the molecule may or may not be broken up:

$$\begin{aligned} &e^- + AX \rightarrow Ax^- * \rightarrow A \ (\text{or } A^*) + X^- \left(\text{or } X^- *\right) \text{ or} \\ &e^- + AX \rightarrow Ax^- * \rightarrow AX^- + \text{energy}. \end{aligned} \tag{4.39}$$

The attachment rate R is proportional to the density N:

$$R = c\sigma N \tag{4.40}$$

for an electron velocity c and attachment cross-section σ The rate constants of freons and many other halogen-containing compounds are known [39].

The best known three-body process is the Bloch-Bradbury process [40]. In this process, an electron is attached to a molecule through the stabilizing action of another molecule. It is important for the attachment of electrons with energy below 1 eV to O_2, forming an excited unstable state with a lifetime τ of the order of 10^{-10} s. A stable ion will be formed only if the excitation energy is carried away during τ by another molecule. The attachment rate is proportional to the square of the gas pressure, as it depends on the product of the concentrations of oxygen and of the stabilizing molecules [19]:

$$R = \tau c_e c_2 \sigma_1 \sigma_2 N\,(O_2)\, N\,(X)\,. \tag{4.41}$$

Here c_e is the electron velocity, c_2 the relative thermal velocity between O_2 and X In an Ar/CH_4 (80/20%) mixture at 8.5 atm with an O_2 contamination of 1 ppm, an absorption of 3%/m was measured at a drift speed of 6 cm/μs.

4.2.3 Avalanche Amplification

4.2.3.1 Operation Modes

Gas detectors generally use gas amplification in the homogeneous field of a parallel plate geometry or, more frequently, in the inhomogeneous field around a thin wire. We shall start with the discussion of the second case.

Near a wire with a charge q_s per cm, the electric field at a distance r from its centre is

$$E = q_s / (2\pi\varepsilon_0 r) . \tag{4.42}$$

When raising the field beyond the *ionization chamber regime*, in which all primary charges are collected *without* any *amplification*, at some distance from the wire a field is reached, in which an electron can gain enough energy to ionize the gas and to start an avalanche. The avalanche will grow until all electrons have arrived on the anode wire. For a *gas amplification A* of $1000 \sim 2^{10}$, some 10 ionization generations are required. As the mean free path between collisions is of the order of microns, the field to start an avalanche has to be several times 10^4 V/cm. This is usually achieved by applying a voltage of a few kV to a thin wire, with a diameter in the $20 - 50$ μm range.

Besides ionization, *excitation* will always occur and with it photon emission. A fraction of these photons may be energetic enough to produce further ionization in the gas or on the cathode. Only those photons which ionize outside the radius r_{av} of the moving electron avalanche may be harmful, as their avalanches will arrive later. If γ called the *second Townsend coefficient*, is the probability per ion pair in the first avalanche to produce one new electron, and if A denotes the amplification of the first avalanche, *breakdown* will occur for

$$A\gamma > 1. \tag{4.43}$$

In this case the first avalanche will be followed by a bigger one, this by an even bigger and so on, until the current is limited by external means. If $A\gamma << 1$, $A\gamma$ gives the probability for producing an after-discharge. If a photoionization takes place inside r_{av}, the effect will be an increase of A.

The resulting need to suppress far-traveling photons produced in the rare gases is the reason for the use of 'quench gases' like Methane, Ethane, CO_2, etc., which have large absorption coefficients for UV photons.

The positive ions produced in the avalanche have too little energy to contribute to the ionization in the avalanche. They will move slowly to the cathode(s), where they get neutralized but where rare gas ions may also liberate additional electrons. The addition of the quencher reduces this risk significantly, as its recombination energy can be dissipated in other ways, e.g. by disintegration. This explains why more complex molecules provide higher protection.

Up to a certain value A_p, one has a *proportional regime*: the signal produced will on average be proportional to the number of primary electrons. The amplification will rise approximately exponentially with voltage. The azimuthal extension of the avalanche around the wire will grow with amplification and eventually the avalanche will surround the wire.

When the field is raised above this proportional regime, *space charge effects* will set in. The space charge of the positive ions—moving only very slowly compared to the electrons—will reduce the field at the head of the avalanche and the amplification will rise more slowly with voltage and will no longer be proportional to the primary ionization. In addition, space charge effects will depend on the track angle with respect to the wire and on the density of the primary ionization. This is the so-called *limited proportionality regime*.

Increasing the field further, the positive space charge may produce additional effects. Near the avalanche tail the electric field is increased. If the absorption of UV photons in the quench gas is high, the photons may ionize this high field region and start a *limited streamer* moving backwards by starting avalanches further and further away from the sense wire. As the electric field at large radius weakens, this development will stop after typically 1–3 mm. The total charge is almost independent of the primary charge starting the streamer. The process depends quite strongly on experimental conditions. An example is presented in Fig. 4.11, which shows a steep step from the proportional regime [41]. In the narrow transition zone one finds a rapid change of the ratio of streamer/proportional signal rates. In other experimental conditions a smoother transition has been observed.

If the absorption of the UV photons is weak, photons travel further and avalanches may be started over the full length of the wire, leading to the *Geiger mode*, if the discharge is limited by external means.

4.2.3.2 Gas Gain

With multiplication, the number n of electrons will grow on a path ds by

$$\mathrm{d}n = n\ \alpha\ \mathrm{d}s, \tag{4.44}$$

where α is the *first Townsend coefficient*. Ionization growth is obviously proportional to the gas density p and depends on the ionization cross-sections, which are a function of the instantaneous energy ε of the electrons. This energy itself is a function of E/p. The relationship between α and E is, therefore, given in the form α/p as function of E/p or for a specific temperature as $\alpha/p(E/p)$. Figure 4.12 gives some

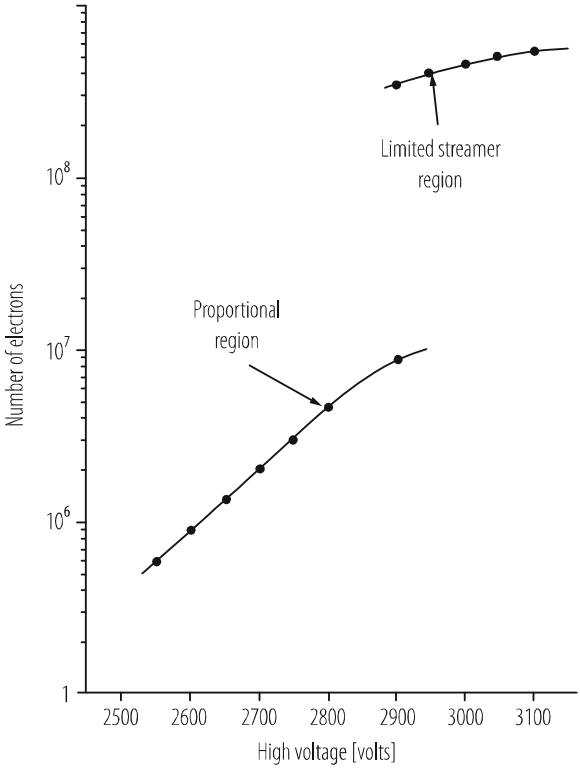


Fig. 4.11 Pulse-height transition from limited proportionality to limited streamer mode [41]

examples of measurements [42]; it shows the strong increase of α with electric field in the region of interest to gas detectors, up to about 250 kV/cm. No simple relation exists for α as function of electric field E, but Monte Carlo simulation has been used to evaluate α. Figure 4.13 shows an example [27]. For the lower field values there is reasonable agreement with measurement. The discrepancy at the highest fields is attributed to photo-and Penning-ionization not being included.

The amplification A in the detector is obtained by integration

$$A = n/n_0 = \exp \int \alpha(s)\, \mathrm{d}s = \exp \int \alpha(E)/\left(\mathrm{d}E/\mathrm{d}s\right) \mathrm{d}E, \tag{4.45}$$

from $E_{\min}$, the minimum field to start the avalanche, to the field $E(a)$ on the wire. $E_{\min} e$ is equal to the ionization energy of the gas molecules divided by the mean free path between collisions. Near the wire and far from other electrodes, the field is

$$E(r) = q_{\mathrm{s}}/\left(2\pi r \varepsilon_0\right), \tag{4.46}$$

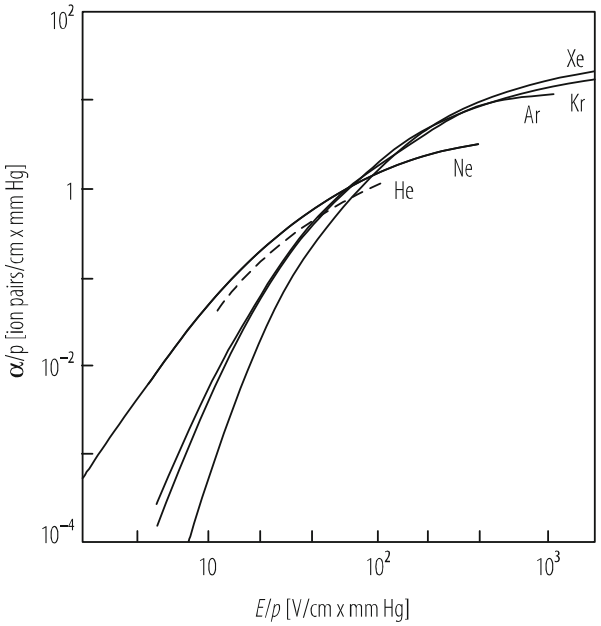


Fig. 4.12 Examples of measured 'First Townsend coefficient' α in rare gases [42]

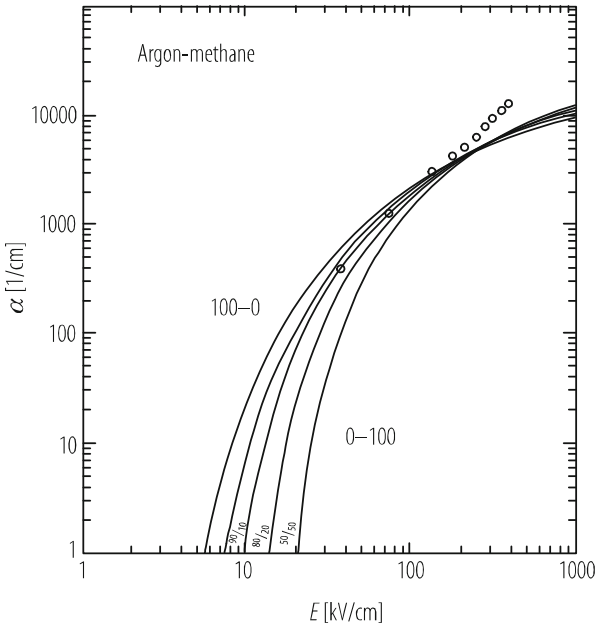


Fig. 4.13 Simulated 'First Townsend coefficient' α in Ar/CH_4 mixtures at 1 atm (100–0 means 100%Ar). Measured values are indicated as circles [27]

where q_s is the charge per cm. Therefore,

$$A = \exp \int q_s \, \alpha(E) \, \mathrm{d}E / \left(2\pi \varepsilon_0 E^2\right). \tag{4.47}$$

Two approximations in particular have been used to describe practical cases.
The early *Korff model* [43] uses the parameterization

$$\alpha / p = A \exp\left(-Bp/E\right), \tag{4.48}$$

with empirical constants A and B depending on the gas.

In the *Diethorn approximation* [44], α is assumed to be proportional to E. One then obtains for a proportional tube with wire radius a and tube radius b

$$\ln A = (\ln 2 / \ln(b/a)) \, (V/\Delta V) \ln\left(V / (\ln(b/a) \, a E_{\min}\right), \tag{4.49}$$

where the two parameters E_{mm} and ΔV are obtained from measurements of$A$ at various voltages and gas pressures. E_{mm} is the minimum E field to start the avalanche and $e\Delta V$ the average energy required to produce one more electron. E_{mm} is defined for a density p_0 at STP. For another density $E_{\mathrm{mm}}(p) = E_{\mathrm{mm}}(p_0)(p/p_0)$. A list for E_{mm} and ΔV for various gases is given, e.g., in [19]. Reasonable agreement with the experimental data is obtained; discrepancies show up at high A.

4.2.3.3 Dependence of Amplification on Various Factors

The gas amplification depends on many operational and geometrical parameters. Some examples are:

Gas Density

The Diethom approximation gives

$$\mathrm{d}A/A = -(\ln 2 / \ln(b/a)) \, (V/\Delta\nabla) \, (\mathrm{d}p/p) \rightarrow = (5\text{-}8) \;\; \mathrm{d}p/p \tag{4.50}$$

typically.

Geometrical Imperfections

The effects will obviously depend on the geometry and the operation details. An early publication [45] gives analytic estimates of the effects of wire displacements and variations in wire diameter. In a typical geometry $\mathrm{d}A/A \sim 2.5 \, \mathrm{d}r/r$, where r is the wire radius; $\mathrm{d}A/A \sim 9 \Delta gap/gap$.

Edge Effects

Near edges, the electric field is reduced over distances similar to the gap between the electrode planes. It can be recovered largely by additional field shaping lines on the edges [46].

Space Charge

Due to the low velocity of the positive ions (falling off as $1/r$ from >1mm/μs at $r = a$), space charge will build up at high particle fluxes and lower the avalanche amplification. In drift tubes the voltage drop due to the space charge from a given particle flux is proportional to the third power of the tube radius. A smaller radius thus improves the rate capability drastically.

4.2.3.4 Statistical Fluctuations of the Amplification

In the proportional regime, the amplification A is simply defined by $A = n/n_{\mathrm{T}}$ and one assumes that each of the n_{T} initial electrons produces on average the same A ion pairs. We define $P(n)$ as the probability to produce n electrons in the individual avalanche with mean A and variance σ^2 If $n_{\mathrm{T}} > > 1$ and if all avalanches develop independently, it follows from the central limit theorem that the distribution function $F(n)$ for the sum of the n_{T} avalanches approaches a Gaussian with mean $n = n_{\mathrm{T}}A$ and variance $S^2 = n_{\mathrm{T}}\sigma^2$, independent of the actual $P(n)$.

On the other hand, for detection of single or a few electrons, knowledge of the individual $P(n)$ is required.

For a parallel plate geometry calculations [47] agree well with measurements [48]. The distributions found theoretically [49] and experimentally [50] for the strong inhomogeneous field around a thin wire also look similar and approach Polya distributions (Fig. 4.14).

For these distributions

$$(\sigma_A / < A >)^2 = f, \text{with } f \leq 1. \tag{4.51}$$

The limiting case $f = 1$ is an exponential distribution (Yule-Furry law)

$$P(A) = (1/ < A >) \exp(-A/ < A >). \tag{4.52}$$

Experimental results point to $f = 0.6 - 1.0$. Measurements with laser tracks [19], indicate that the r.m.s. width σ_A of a single-electron avalanche is close to the mean, as it is for the Polya distribution with $f = 1$. An approximately exponential distribution for single-electron avalanches is also reported in [28, 48].

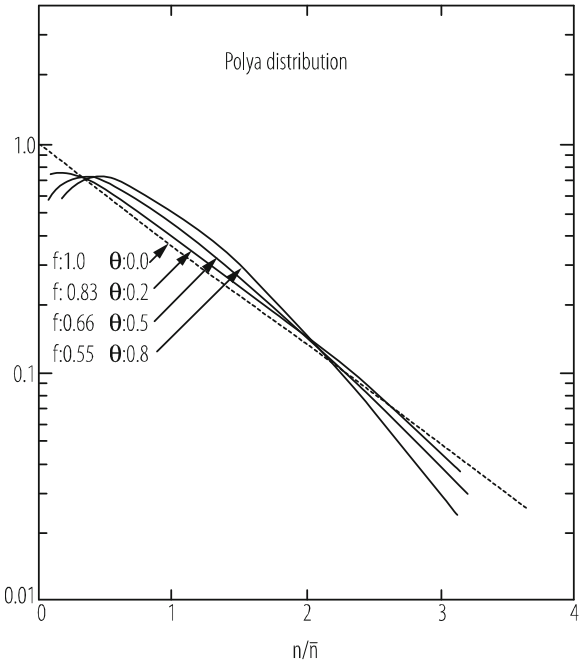


Fig. 4.14 Polya distributions [22]

4.2.4 Signal Formation

In wire chambers, signal formation is very similar to the one in the simplest geometry of a cylindrical tube with a coaxial wire, because most of the useful signal is produced in the immediate vicinity of the sense wire and the electric field around the sense wires in a MWPC can be considered as radial up to a radius equal to about one tenth of the distance between sense wires [45].

Signals are *always* produced *by induction* from the moving charges.

Ramo [51] and Shockley [52] have shown that in general the current I_R induced on the readout electrode R is given by

$$I_R = -q\ \mathbf{E}_w \mathbf{v}, \tag{4.53}$$

where q is the signed charge moving with the vectorial velocity $\mathbf{v}$, and $\mathbf{E}_w$ is a vectorial *weighting field*, a conceptual field defined by applying + 1V on R and 0 V on all other electrodes. The unit of E_w is 1/cm. The actual v is calculated by applying the normal operation voltages, including possibly a B field.

In the special case of a two electrode system like the wire tube, $\mathbf{E}_w = \mathbf{E}_{op}/V$, where $\mathbf{E}_{op}$ is the actual operating field obtained with the voltage V on R (the anode wire) and zero V on the cathode.

For the proportional tube with wire radius a and cathode radius b, E_{op} and E_{w} are obviously radial with

$$E_{\mathrm{op}} = V/\left[r \ln\left(b/a\right)\right]. \tag{4.54}$$

We assume constant mobility μ for the positive ions. Therefore

$$v^{+}\left(\mathrm{t}\right) = \mu V/\left[r(t) \ln\left(b/a\right)\right]. \tag{4.55}$$

For an ion starting at $t = 0$ from $r = r_1$,

$$r(t) = r_1(1 + (t/t_0))^{1/2} \text{ with } t_0 = r_1^2 \ln\left(b/a\right)/\left(2\mu V\right). \tag{4.56}$$

The maximum time for an ion to drift from a to b is

$$T^{+}{}_{\max} = (b/a)^2 t_0, \text{ as } (b/a)^2 >> 1. \tag{4.57}$$

The induced current I^{+} is

$$I^{+} = -q\ \mathbf{E}_{\mathrm{w}}\mathbf{v}^{+} < 0, \tag{4.58}$$

as $\mathbf{v}^{+}$ is parallel to $\mathbf{E}_{\mathrm{w}}$.

For the integrated charge Q, one obtains

$$Q^{+}(t) = \int I\ \mathrm{d}t = \int I\left(1/v^{+}\right)\ \mathrm{d}r = \int -q\,E_{\mathrm{w}}\ \mathrm{d}r. \tag{4.59}$$

Integration from r_1 to r_2 gives

$$Q^{+}{}_{1\to 2} = -q \ln\left(r_2/r_1\right)/\ln\left(b/a\right), \text{ with } q > 0 \text{ and } r_2 > r_1, \tag{4.60}$$

For an electron one obtains

$$Q^{-}{}_{1\to 2} = +q\ |\ln\left(r_2/r_1\right)|\ /\ln\left(b/a\right), \text{ with } q < 0 \text{ and } r_2 < r_1, \tag{4.61}$$

as $\mathbf{v}$ is antiparallel to $\mathbf{E}_{\mathrm{w}}$.

We shall give numbers for a typical proportional tube with $a = 10\ \mu\mathrm{m}$, $b = 2.5$ mm, $E_{\mathrm{op}}(r = a) = 200$ kV/cm, $\mu^{+} = 1.9$ atm cm^2/(Vs), $\mathrm{v}^{-} \approx 5 \cdot 10^6$ cm/s and—to estimate the gas amplification A—the Diethom parametrization $\alpha = (\ln 2/\Delta\nabla)E$ and $E_{\min} = V/(r_{\mathrm{mm}} \ln\left(b/a\right))$, taking for an Ar/$CH_4$(90/10) mixture $\Delta V = 23.6$ V and $\mathrm{E}_{\min} = 48$ kV/cm [19], p. 136. Here r_{mm} is the starting radius for the avalanche and $E_{\min}$ the minimum field permitting multiplication. We obtain: $t_0 = 1.3$ ns, $T^{+}_{\max} = 82\ \mu\mathrm{s}$, $r_{\min} = 42\ \mu\mathrm{m}$, $A = 4400$.

The last electron will be collected in a very short time of about 0.6 ns, the vast majority even faster. Half of the electrons move only about 2 μm, the next 25%

some 4 μm and so on. A rough estimate of the induced electron charge signal is, therefore,

$$Q^{-}{}_{\text{total}} = q \ln(14/10) / \ln(2500/10) = 0.06\, q. \tag{4.62}$$

Only about 6% of the total induced signal is due to the movement of the electrons, the rest from the ions, if one integrates over the full ion collection time.

In practice, however, one mostly uses much faster integration. The long tail in the signal caused by the very slow ion movement has to be corrected for by electronic pulse shaping to avoid pile-up at high rates (see Sect. 4.69). If one uses fast pulse shaping, say 20 ns integration, only a fraction of the ion charge will be seen: an ion starting at $r_1 = a$, reaches $r_2 = 40\,\mu\text{m}$ in 20 ns and induces about 25% of its charge. That means: with 20 ns pulse shaping, one may expect to see an *effective charge* of about 30% of the total charge produced, of which one fifth is due to the electrons.

4.2.5 Limits to Space Resolution

The space resolution σ_X obtained from a single measurement of the anode wire signals in a multi-wire proportional chamber is given by the separation s of the wires: $\sigma_x = s/\sqrt{12}$. The minimum practical s for small chambers is 1 mm. The best resolution is thus about 300 μm.

Significantly better resolution may be obtained either from 'centre of gravity' determination or from the electron drift times in drift chambers.

4.2.5.1 'Centre of Gravity' Method

In this method one uses the signals induced on cathode strips or pads, see Fig. 4.19. The rms width of the induced charge distribution is comparable to the anode-cathode gap d. If one chooses a strip width of (1–2)d, one obtains signals above threshold on typically 3–5 strips. Depending on the signal to noise ratio, a resolution of typically (1–5)% of the strip width is achieved, i.e. about $40 - 100\,\mu\text{m}$. This method is used for the read out of TPCs, as well as for high precision cathode strip chambers, see e.g. [15, 16].

4.2.5.2 Drift Time Measurement

The main contributions to the error of the drift time determination come from electronics noise, electron clustering, δ-rays, diffusion. This assumes that additional effects on the space-time correlation including magnetic field corrections, gain variations, gas contamination and others are kept small by careful construction and calibration. Figures 4.16 and 4.24 (*right*) show typical results. Electronic noise

contributes a constant error. Near the anode wire, the effects of the clustering of the primary charges adds a significant error. At large distances from the anode, the contribution from diffusion grows as square root of distance. Resolutions achieved are typically $50 - 200\,\mu\text{m}$.

A detailed discussion of limits to space resolution is presented in [19] and for the particular case of proportional tubes in [15].

4.2.6 Ageing of Wire Chambers

Deterioration of performance with time has been observed since the early days of gas detectors but has gained importance with the ever increasing radiation loads due to the demand for higher detection rates over long periods. Typical effects of *ageing* are: pulse height decrease, a broadening of the energy resolution and increase in dark current, in the extreme also electrical breakdown or broken wires. An enormous number of studies has been carried out. They are well documented in the proceedings of workshops [54] and several reviews [55].

Upon opening of damaged chambers, deposits have been observed on anode wires and/or on cathodes. On the wire they can take any form from smooth layers to long thin whiskers [56], see Fig. 4.15. On the cathodes, deposits usually consist in spots of thin insulator. Defects of this latter kind can often be correlated with a discharge pattern, which may be interpreted as Malter effect [57]: under irradiation, charges build up on the insulator until the electric field is strong enough to extract electrons from the cathode through the layer into the gas where they initiate new avalanches. The facts that the buildup time decreases with higher ionization rate and that the discharges take some time to decay after irradiation is timed off, give support to this explanation, as does the observation that addition of water vapour is reducing the discharges, probably introducing some conductivity.

Analysis of the layers and whiskers on the anode wires often indicate carbon compounds, more surprisingly also often silicon, sometimes other elements: Cl, O, S.

The aging results are often characterized by a drop in pulse height ΔPH as function of integrated charge deposition in Coulomb per cm wire, although it was found in some cases that the rate of the charge deposition has an influence. Typical values with classical gas mixtures containing hydrocarbons are

$$\begin{aligned} \Delta PH/PH &\sim 0.01 - 0.1\%/0/\text{mC/cm} \quad \text{for small detectors,} \\ \Delta PH/PH &\approx 0.1 - 1\%/\text{mC/cm} \qquad\quad \text{for large detectors.} \end{aligned} \tag{4.63}$$

It is obvious that the control of ageing is one of the major challenges for the LHC experiments, possibly even the major one.

Unfortunately, however, it has not been possible to establish a common fundamental theory, which could predict lifetimes of a new system. On the other hand, the

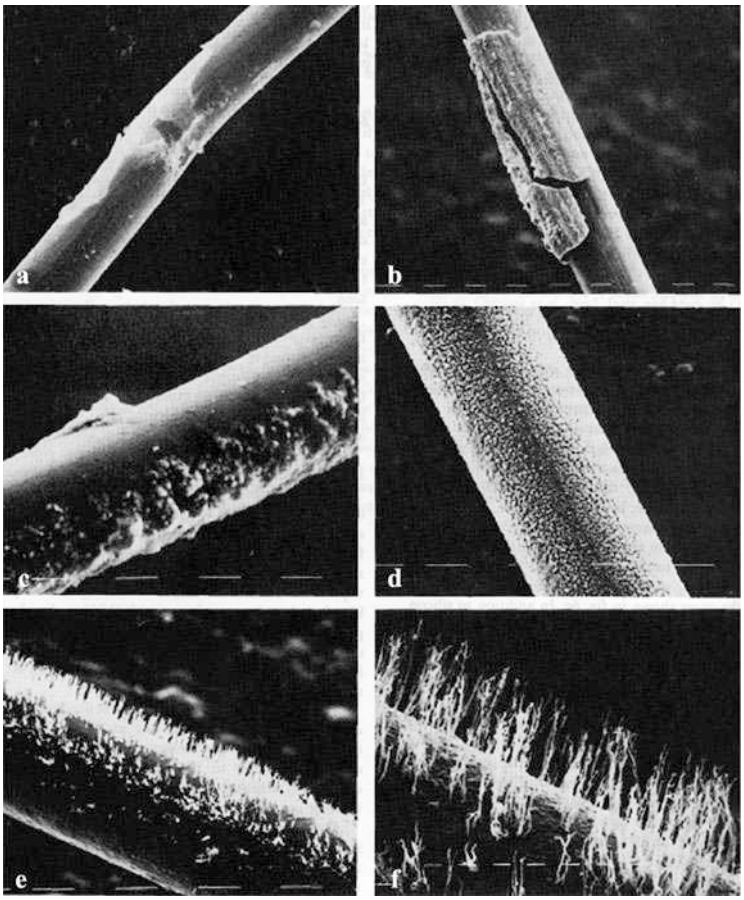


Fig. 4.15 Examples of deposits on 20 μm anode wires after strong irradiation [56]

reasons for ageing in particular circumstances have been elucidated and the studies permit to establish *some general rules* on how for improving the chances for a longer lifetime:

- *Many materials have to be avoided*, in the gas system, in the detector and during the construction: Si compounds, e.g. in bubbler oils, adhesives, vacuum grease or protection foils, PVC tubing, soft plastics in general, certain glues and many more. The workshop proceedings and reviews mentioned present details, also on materials found acceptable.
- For the highest radiation loads, up to a few C/cm, gas components most frequently used in the past, namely *hydrocarbons* like Methane, Ethane, etc., *should be avoided*. Indeed, the LHC experiments make use of them only exceptionally. There remains only a very restricted choice of acceptable gases: mixtures of rare gases and CO_2 and possibly N_2, CF_4 or DME. CF_4 is offering high electron drift

velocities and has proven to be capable under certain conditions to avoid or even to etch away deposits, in particular in the presence of minute Si impurities. But its aggressive radicals may also etch away chamber components, especially glass [15]. In any case the water content has to be carefully controlled to stay below 0.1%, if CF_4 is used, to avoid etching even of gold-plated wires. Also DME, offering low diffusion, has in some cases provided long lifetime. It has, however, shown to attack Kapton and to be very sensitive to traces of halogen pollutants at the ppb level.

- During production, *high cleanliness* has to be observed, e.g. to avoid resistive spots on the cathodes. The sense wire has to be continuously checked during wiring to assure the required quality of its geometrical tolerances and of the gold plating.
- The gas amplification should be kept as low as possible.
- In any case, a final detector module with the final gas system components should be extensively tested under irradiation. As an accelerated test is usually required for practical reasons, to obtain the full integrated charge for some 10 years of operation in a 1 year test, an uncertainty un- fortunately will remain, because a rate dependence of the ageing cannot be excluded.

4.3 Detector Designs and Performance

4.3.1 Single Wire Proportional Tubes

Despite the revolution started with the multiwire proportional chambers (MWPC), single wire tubes are still widely used, mostly as drift tubes. They have circular or quadratic cross-section and offer independence of the cells, important, e.g., in case wire rupture. We present three examples.

ATLAS has chosen for the muon system *large diameter* (*3 cm*) *aluminum tubes* operated with Ar/CO_2 (93/7%) at 3 atm, with the addition of about 300 ppm of water to improve HV stability. A pair of 3 or 4 staggered tube layers, separated by a support frame, form a module. The disadvantages of the gas mixture, a non-linear space-drift time relation and relatively long maximum drift time, had to be accepted in order to obtain a high radiation tolerance. The spatial resolution for a single tube under strong γ- irradiation producing space charge is shown in Fig. 4.16. An average resolution per tube of 80 μm is expected with a maximum background rate of 150 hits/cm^2 s. With a relative positioning of the wires during construction to 20 μm, an adjustment of the tube curvature to the gravitational wire sag and a relative alignment and continuous monitoring of the pair of layers inside a chamber, a combined resolution for the 6–8 1ayers of ~35 μm is aimed at. These chambers provide only one coordinate, the other one being measured in other subdetectors of the experiment.

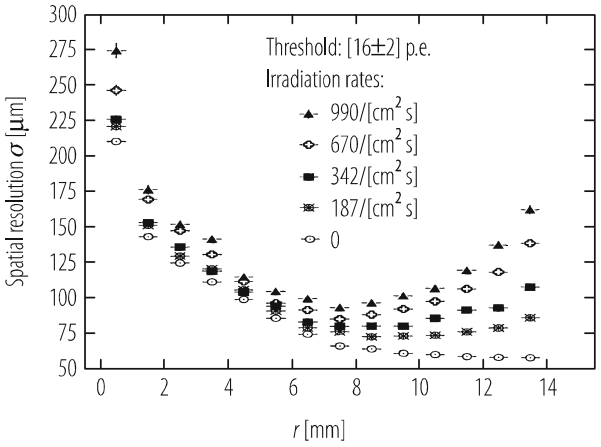


Fig. 4.16 ATLAS MDT drift tubes: space resolution as function of impact radius and background rates. Expected rates are $\leq$150 hits/cm^2 s [15]

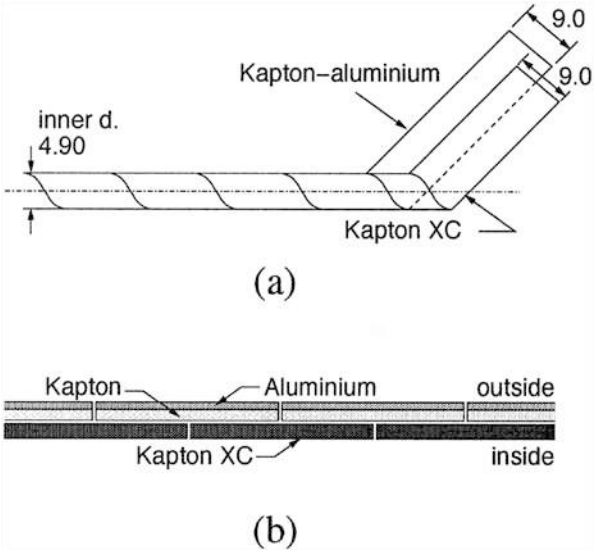


Fig. 4.17 LHCb straw tubes: (**a**) Winding scheme. (**b**) Details of the double foil. Kapton XC is on the inside [17]

A second type of tube design, *straw tubes*, has become very popular since a number of years. Straw tubes offer high rate capability due to small diameters and relatively little material in the particle path. In LHCb, where local rates (near one end of the wire) up to 100 kHz/cm have to be handled, an internal diameter of 4.9 mm was chosen [17], with a construction shown in Fig. 4.17. Two strips of thin foils are wound together with overlap. The inner foil, acting as cathode, is made

of 40 μm carbon doped polyimide (Kapton-XC), the outer is a laminate of 25 μm polyimide, to enhance the gas tightness, and 12.5 μm aluminum, to ensure fast signal transmission and good shielding. The tubes are up to 2.5 m long and have the 25 μm wire supported every 80 cm.

Staggered double layers tubes are glued to light support panels to form modules up to 5 m long. An average spatial resolution of a double layer below 200 μm was measured with Ar/CO_2 (70/30%). In a station, 4 double layers are aligned along $0, +5, -5, 0^\circ$, thus providing a crude second coordinate measurement.

In another design for ATLAS [15], mechanical strengthening of 4 mm straws is achieved with carbon fibers wound around the tubes and straightness by supporting them every 25 cm with alignment planes vertical to the straws. This construction reduces the material along the radial tracks, which is essential for the role of the straws to detect transition radiation originating in fibers stacked in between the tube layers. This role also determines the need for Xe in a mixture of $Xe/CO_2/O_2$ (70/27/3%), the oxygen addition increasing the safety margin against breakdown.

Another quasi-single wire design is that of *plastic streamer tubes* usually called *Iarocci tubes*. Because they are easy to construct in large size and cheap, they have been widely used, especially as readout planes in hadron calorimeters.

A plastic extrusion with an open profile with typically 8 cells of lxl cm^2 and a PVC top plate, is coated on the inside with graphite with a minimum resistivity of 200 kΩ/square. All this is slid into a plastic box, which serves also as gas container. For stability, wires are held by plastic spacers every 50 cm. Using a thick wire of 100 μm, self-quenching streamers are initiated in a gas containing a strong quencher, typically isobutane in addition to Ar. Electrodes of any shape placed on one or both external surfaces pick up the rather strong signals. The dead time is long but only locally, so that particle rates up to 100 Hz/cm^2 can be handled.

4.3.2 Multiwire Proportional Chambers (MWPC)

Already 1 year after the invention of the multiwire proportional counter (MWPC) by Charpak in 1968, a system of small chambers was used in an experiment [58], another year later a large chamber 2 m × 0.5 m had been tested with Ar/Isobutane [59]. A number of developments like bi-dimensional readout were discussed [60]. In 1973 already, a large system of MWPC containing 50,000 sense wires had been constructed for an spectrometer at the ISR, the *Split-Field Magnet* (SFM) [61]. All these chambers had a geometry for the sense wires and gap size similar to the *original design*, shown in Fig. 4.18.

The SFM chambers of 2 × 1 m^2 contained three light support panels forming two amplification gaps of 2 × 8 mm each, one with vertical, the other with horizontal wires of 20 μm diameter. The cathodes on the panels were sprayed with silver paint to provide readout strips 5.5 cm wide and at angles of $\pm 30^\circ$, to resolve ambiguities. Special emphasis was put on high precision with a light construction, including frames of only 5 mm thickness. A total thickness of 1.7% of a radiation length

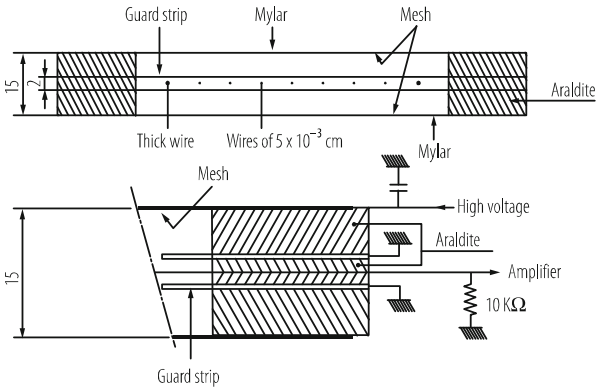


Fig. 4.18 Design of the first multiwire proportional chamber [8]

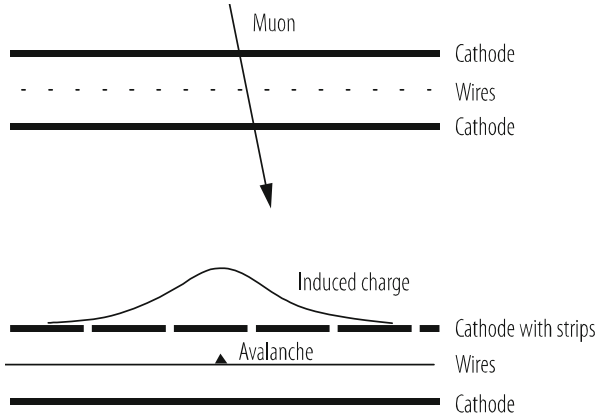


Fig. 4.19 Principle of 'centre of gravity' cathode measurement of cathode strip signals

per chamber was achieved. The stringent quality demands can be inferred from the definition of the *efficiency plateau*: the 'beginning of plateau' was defined as efficiency ε 99.98% and the end by 10 times the 'normal noise', corresponding to cosmics rate With this tight definition, the measured plateau length for a chamber was $50 - 100$ V with the magic gas mixture of Ar/isobutane/freon/methylal (67.6/25/0.4/7%).

In the following decades, drift chambers imposed themselves more and more, but MWPCs remained valid options, especially when speed was more important than high spatial resolution. Thus three of the four LHC experiments employ MWPCs, for triggering and momentum measurements. All of them make use of the signals induced on the cathode strips. In ATLAS and CMS, who call their chambers cathode strip chambers (CSC), a high precision measurement is obtained in the bending coordinate by determining the 'centre of gravity' of the strip signals, see Fig. 4.19. In ATLAS, each third strip is read out (pitch ~5.5 mm), leaving two strips floating,

and a resolution of 60 μm is obtained [15, p. 178]; in CMS, 75 μm resolution is achieved with each strip read out at a minimum pitch of 8.4 mm [16, p. 197]. In LHCb, spatial resolution is secondary to fast timing and high efficiency for a five-fold coincidence trigger. Adjustment to the requirements on spatial resolution, which change strongly with radius, is achieved by forming readout pads of variable size ($0.5 \times 2.5 - 16 \times 20$ cm^2) on the cathodes and by grouping sense wires [17, p. 130].

4.3.3 Drift Chambers

Already in the very first publications, the basic two types of drift chambers were described: (i) with the drift volume, through which the particles pass, separated from the amplification volume [9] and (ii) a geometry, in which the particles pass directly through the volume containing the anode wires alternating with field wires to improve the drift field [10], see Fig. 4.20.

The first design finally evolved into the TPC, the second into a large number of different designs. One can differentiate between planar and cylindrical geometries.

4.3.3.1 Planar Geometries

Most planar geometries are rather similar to each other. To obtain a more homogeneous drift field, additional field shaping electrodes are introduced, see Fig. 4.21. Also shown is a recent example, one element of a layer for the Barrel Muon system of CMS. The space resolution per layer is about 250 μm. One muon station consists of 2×4 layers of such tubes fixed to an aluminum honeycomb plate. The other coordinate is provided by a third set of 4 layers oriented at 90°.

The central detector of UAl used a special arrangement, see Fig. 4.22. In a horizontal magnetic field, at right angle to the beam, a cylinder 6 m long and 2.2 m in diameter is filled with planar subelements. In the central part, vertical anode planes

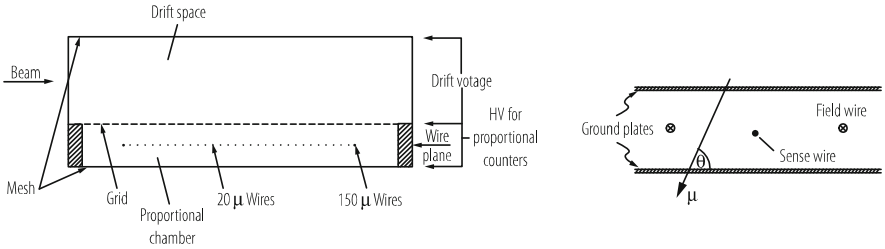


Fig. 4.20 First two drift chamber designs. *Left*: separate drift and amplification gaps [9]. *Right*: Common drift and amplification volume. The additional field wires improve the drift field [10]

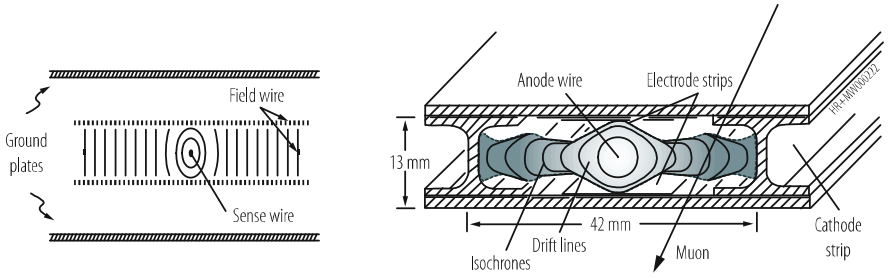


Fig. 4.21 Planar arrangement with field shaping electrodes. *Right*: cross-section of large drift tubes for CMS [16]

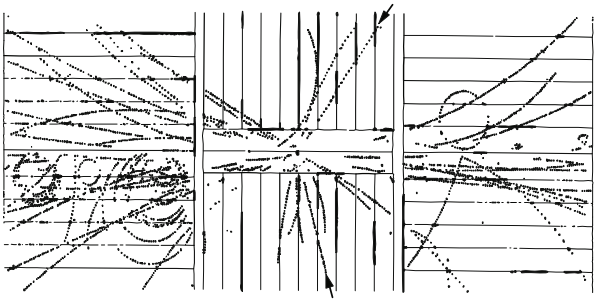


Fig. 4.22 Horizontal view of a reconstructed event in the central detector of UAl. The first Z° decay observed [19]

ultimate with cathode planes, leaving 18 cm drift spaces. These planes are horizontal in the two ends. Charge division is used for the coordinate along the wire. The average point accuracy along the drift direction was 350 μm.

4.3.3.2 Cylindrical Geometries

Many different arrangements have been worked out. Figure 4.23 shows an example of a wire arrangement and electron drift lines following the electric field in the absence of a magnetic field. The change of the electron drift in a magnetic field in a similar cell is indicated to the right.

Figure 4.24 (*left*) shows the conceptual design of the drift chamber for OPAL [53], a wire arrangement called *Jet Chamber*. The left-right ambiguity is solved by staggering the sense wires alliteratively by ±100 μm. The measured space resolution in $r\phi$ for a single wire is presented in Fig. 4.24 (*right*). The figure shows the typical dependence on the distance r from the sense wire: for small r, the primary ion statistics dominates, at large r diffusion. In addition, there is a constant contribution from the noise of the electronics. The coordinate along the wires is

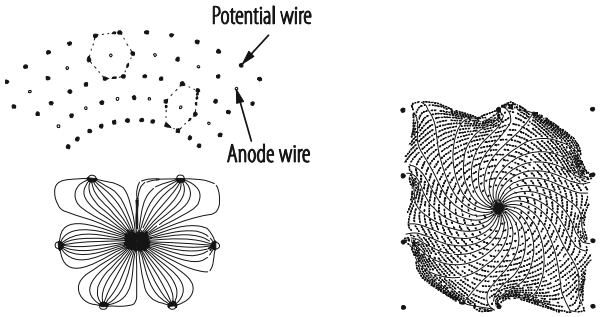


Fig. 4.23 Two multi-layer wire arrangements and electron drift lines without (*left*) and with (*right*) magnetic field

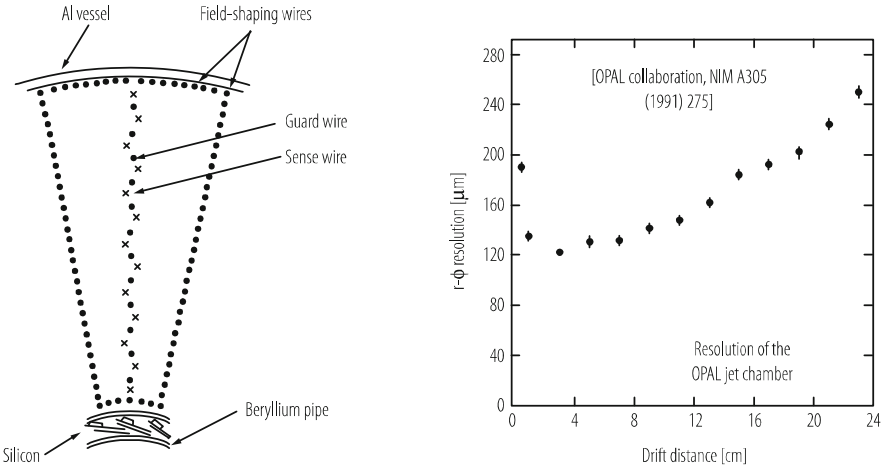


Fig. 4.24 Jet Chamber. *Left*: conceptual design with staggered sense wires. *Right*: Space resolution obtained in OPAL with 4 atm [53]

obtained from *charge division* by using resistive sense wires and read-out on both ends of the wires: a resolution of about 1% of the wire length is reached.

In other designs the *second coordinate* is obtained from orienting successive layers in stereo angles. Sometimes relative timing with read-out of both ends of the wire is used, providing again a resolution of about 1% of the wire length.

4.3.3.3 Time Projection Chambers (TPC)

The TPC concept proposed by Nygren [62] in 1974 for the PEP4 experiment [63] offered powerful pattern recognition with *many unambiguous 3-D points* along a track and particle identification by combining dE/dx information from many samples with momentum measurement. Originally proposed to resolve jets at a low

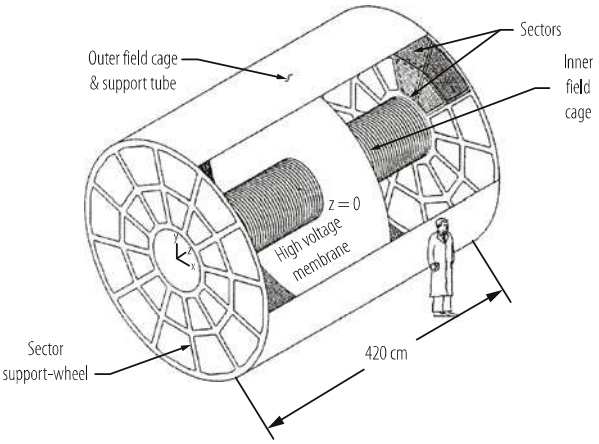


Fig. 4.25 Conceptual design of the STAR TPC operating at RIC [64]

energy e^+e^- collider, the TPC design has proven years later to be the most powerful tracker to study central heavy ion collisions with up to several thousand particles in an event, at more than 100 events per second.

The basic design elements have hardly changed over the years. Cylindrical field cages provide a homogeneous electric field between the central electrode and the planar wire chambers at both ends; see Fig. 4.25 for the conceptual design of the latest TPC in operation, the STAR TPC at RHIC [64]. The typical gas mixture is Ar/CH_4, which offers high drift velocity at low electric field and low electron attachment. The electrons from the track ionization drift to one of the two endcaps. They traverse a gating grid and a cathode grid before being amplified on 20 μm anode wires, separated with field wires. The avalanche position along the wires is obtained from measuring the centre of gravity of pulse heights from pads of the segmented cathode beneath. Figure 4.26 shows the electric field lines for a closed and an open gating grid. *Gating* is essential for the TPCs with their long drift length, to reduce space charge build-up. The gate is only opened on a trigger.

All TPCs except PEP4 and TOPAZ operated at latm and profit from a strong reduction of lateral diffusion due to the factor $\omega\tau$~5 in the strong magnetic field B oriented parallel to the electric field E. Higher pressure is rather neutral: $\omega\tau$ decreases, but more primary electrons reduce relative fluctuations and thus **ExB** and track angle effects. Typical point resolutions in $r\Phi$ range from 150 to 200 μm at the e^+e^- colliders [65]. Figure 4.27 shows a reconstructed Pb–Pb interaction observed in STAR.

All TPCs except STAR and ALICE use the signals from the anode wires for dE/dx information. In STAR and ALICE, all information is taken from the pads, some 560,000 in ALICE [14]. Pressure improves dE/dx and the PEP4 TPC operating at 8.5 atm produced the best dE/dx resolution despite a smaller radius [65].

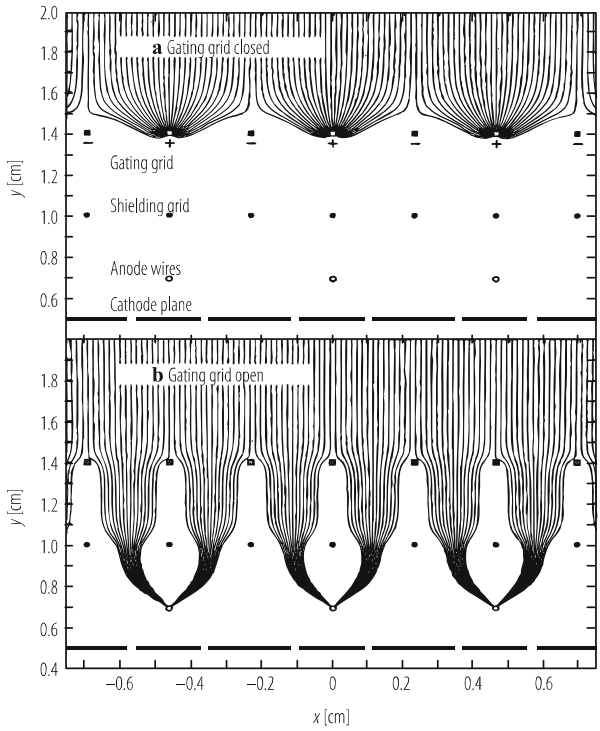


Fig. 4.26 TPC wire chamber: electric field lines for a closed (**a**) and open *x* [cm] gating grid **(b)**

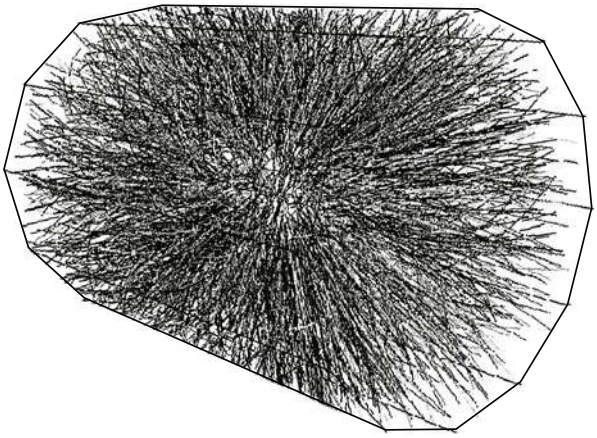


Fig. 4.27 A reconstructed Pb–Pb interaction observed in the STAR TPC [64]

4.3.4 Parallel Plate Geometries, Resistive Plate Chambers (RPCs)

Parallel plate devices offer fast response, as there is no drift delay and the avalanche amplification starts immediately.

Keuffel's spark counter [6] featured two metal electrodes at millimeter distance in a gaseous atmosphere, where the primary electrons deposited in the gap provoke a fast discharge and therefore a detectable signal. By the end of the 1960s the spark counters had arrived at time resolutions around 100 ps, the rates however were limited to 1 kHz and small areas of about 30 cm^2, since after each discharge the entire counter was insensitive during the recharge time of typically a few hundred microseconds. Parallel Plate Chambers (PPCs) use the same geometry but operate below the discharge voltage. The avalanche therefore induces a signal but does not create a discharge, allowing a rate capability of 100 kHz for a 80 cm^2 detector [6]. Still, the fact that the detector mechanics and especially the detector boundaries have very carefully controlled to ensure stability, limits this detector to a rather small surface.

The Pestov spark counter [66] uses the same parallel plate geometry, with one electrode made from resistive material having a volume resistivity of $\rho = 10^9 - 10^{10}$ Ωcm. The charge deposited locally on this resistive layer takes a time of $\tau \approx \rho\varepsilon$ to be removed, where ε is the permittivity of the resistive plate. This time is very long compared to the timescale of the avalanche process, the electric field is therefore reduced at the location of the avalanche, avoiding a discharge of the entire counter. This allows stable operation of the detector at very high fields and particle rates. A counter with a size of 600 cm^2 and a gas gap of 1 mm, operated at atmospheric pressure achieved a time resolution of <0.5 ns and efficiency of 98%. By decreasing the size of the gas gap to 0.1 mm and operating the detector at 12 bar pressure, a time resolution of 27 ps was achieved with this detector [67].

Resistive Plate Chambers (RPCs) [68] are building on this same principle and they are widely used as trigger detectors and for time-of-flight measurements, as they allow relatively cheap large area construction. Large detector systems of several hundred m^2 surface have been built with Bakelite plates ($\rho = 10^{10} - 10^{12}$ Ωcm) or window glass ($\rho = 10^{12} - 10^{13}\,\Omega$ cm). Tetrafluorethane ($C_2F_4H_2$) is nowadays widely used as the main component of the RPC gas mixture due to the large number of primary ionization clusters (8–10/mm) leading to large detection efficiency and due to it's electronegativity that reduces the probability for the formation of streamers. Small additions of SF_6 are further reducing this streamer probability.

The time resolution for RPCs is given by $\sigma_t \approx 1.28/\alpha v$, where α is the effective Townsend coefficient of the gas mixture and v is the drift-velocity of the electrons. RPCs with a single gas gap of 2 mm at a field of 50 kV/cm ($\alpha \approx 10$/mm, $v \approx 130\,\mu$m/ns) provide a time resolution of $\approx$1 ns and efficiency close to 100%. Figure 4.28a shows the geometry as used for the muon system of the ATLAS experiment. In addition to collider experiments, similar geometries are used as

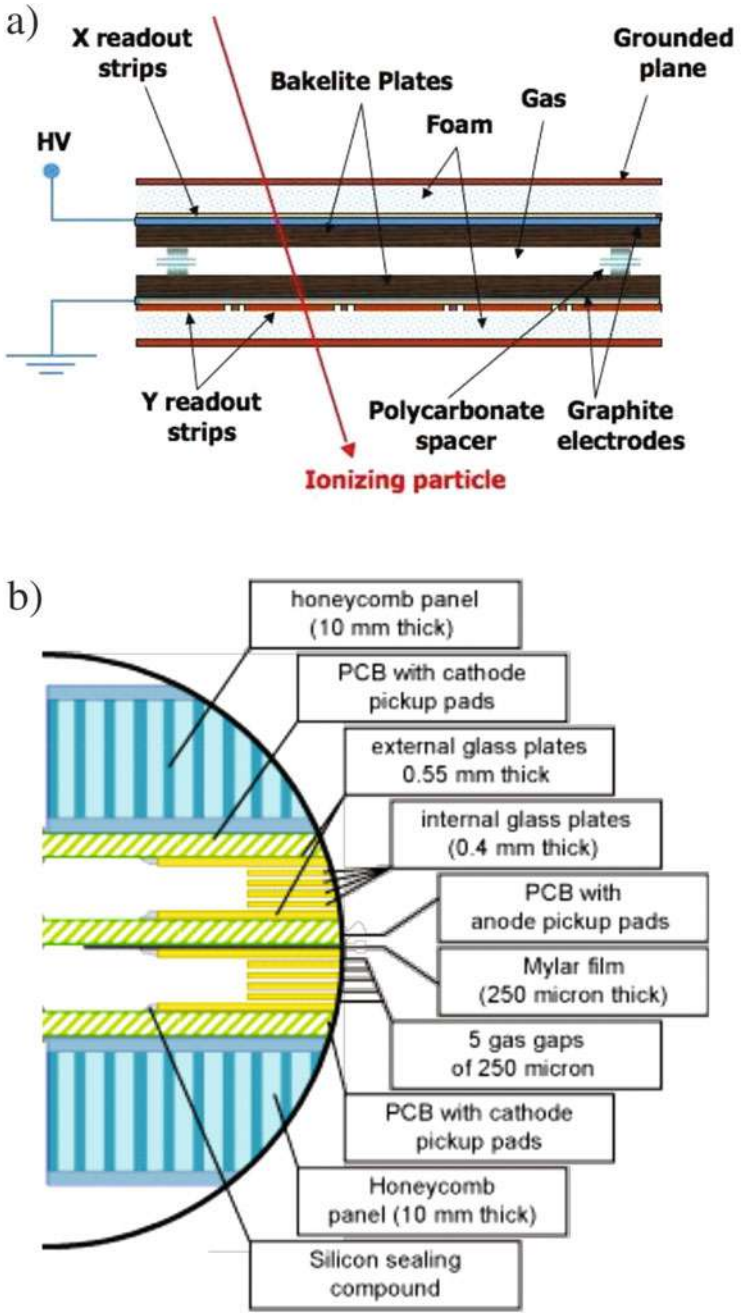


Fig. 4.28 (**a**) Single gap RPC as used by the ATLAS experiment for the muon trigger system. (**b**) Multigap RPC as used by the ALICE experiment for the time of flight system

trigger or veto detectors in neutrino experiments like OPERA [69] and Daya Bay [70] and as large area cosmic ray detectors like ARGO [71].

Using a small gas gap of 0.25–0.3 mm with a field around 100 kV/cm ($\alpha \approx 113$/mm, $v \approx 210\,\mu$m/ns) results in a time resolution of $\approx$50 ps, making the detector well suited for time-of-flight measurements. The reduced efficiency due to the narrow gas gap is overcome by using a multi-gap structure [83]. Figure 4.28b shows the geometry as used for the time of flight system of the ALICE experiment. The avalanche process in RPCs is significantly affected by spacecharge effects. After the initial exponential increase of the electron number, the ions produced in the avalanche are significantly reducing the electric field and therefore resulting in strong slow down of the avalanche growth. This results in moderate signal charges in the pC range even for very large Townsend coefficients [72].

The rate capability of RPCs is defined by the thickness d of the resistive plates and their volume resistivity ρ. The current I produced per unit area inside the gas gap is flowing through these plates, which results in an effective voltage drop of $\Delta V = I\rho d$ across a single plate. The rate limit of the RPC is reached at the point where the effective voltage across the gas gap moves outside the efficiency plateau. For the values and geometries quoted above, this limit is in the range of 10–1000 Hz/cm^2.

4.3.5 Micropattern Devices

The constantly increasing particle rates and track densities in modern day experiments exceed the capabilities of standard gaseous detectors. Semiconductor technology dominates this regime. On the other hand, numerous novel designs of gaseous detectors have been studied. Two have emerged and attract much attention, the so-called GEM and Micromegas devices. Offering small ExB track distortions and low ion feedback, they are also being used for the readout of TPCs.

4.3.6 Gas Electron Multiplier (GEM)

In a thin metal-coated polymer foil, holes are chemically etched at high special density [73], see Fig. 4.29. A voltage applied to the metal layers produces gas amplification in the holes. Typical parameters are: Foil thickness = 50 μm, inner hole diameter = 70 μm, hole pitch = 140 μm, voltage = 400 V. To achieve a practical gas gain of the order of $10^4 - 10^5$ with an acceptable low discharge probability, usually three GEMs are put in series. In COMPASS [74], a system of 20 triple-GEMs with an active area of 31 $\times$ 31 cm^2 each was operated in a very high intensity muon beam. With 2-D readout via superposed orthogonal strips, a space resolution of 70 μm was achieved at rates up to 2.5 MHz/cm^2. The efficiency was 99% with 50 ns pulse shaping at an effective gain of 8000.

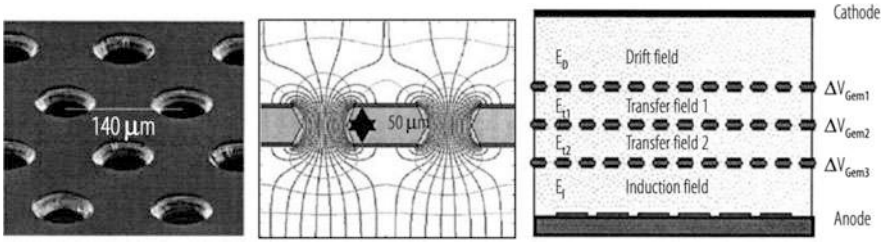


Fig. 4.29 GEM. Left: hole structure. Centre: electric field lines. Right: Conceptual design of tripleGEM [73]

In LHCb [75], with a much shorter peaking time of 10 ns, an efficiency of $\geq$96% was reached for two triple-GEMs in OR and a gain of 6000 with $Ar/CO_2/CF_4$ (45/15/40). The time resolution with this chamber was $\leq$3 ns with Ar/CO_2 (70/30).

Charge build-up is observed on the insulating holes in the GEM foils, which varies with the particle flux and is accompanied by some change in gain, well described by simulation [76].

Several large scale GEM systems are under construction. The wire chambers of the ALICE TPC are replaced by an a quadruple-GEM arrangement that is optimized for low ion backflow. This allows the TPC to operate in continuous mode without the need for gating. The top and bottom GEMs have a hole separation of 140 μm as described above, while the two middle GEMs have twice this distance. This results in an ion feedback below 1% and energy resolution of <12% for 5.9 keV photons at an effective gain of $\approx$2000 [77].

The muon system of the CMS experiment is being equipped with standard triple-GEM detectors over a surface of $\approx$220 m^2 providing spatial resolution of $200 - 400\,\mu$m and a time resolution of 8 ns [78].

4.3.7 Micromegas

In a Micromegas detector [79], the ionization produced in the drift gap is channeled through an extremely fine mesh into the amplification gap, terminated by an anode plane segmented into readout strips or pads as seen in Fig. 4.30. The 'micromesh' is woven from $\approx 15\,\mu$m wires leaving holes of about 50 μm^2. The amplification gap is only 50–150 μm thick and behaves on average like a parallel counter. The mesh is supported by pillars every $\approx$2.5 mm. Due to the high amplification field in comparison to the drift field, the electrons are moving to the anode only inside a very thin funnel, see Fig. 4.30.

In the COMPASS experiment, 12 chambers 40×40 cm^2 have operated in fluxes up to 25 MHz/mm^2 and obtained resolutions of 70–90 μm and 9 ns. The near detector of the T2K experiment [75] uses a TPC with 9 m^2 of MICROMEGAs detectors with readout pads of 10×7 mm^2. The muon system of the ATLAS detector is

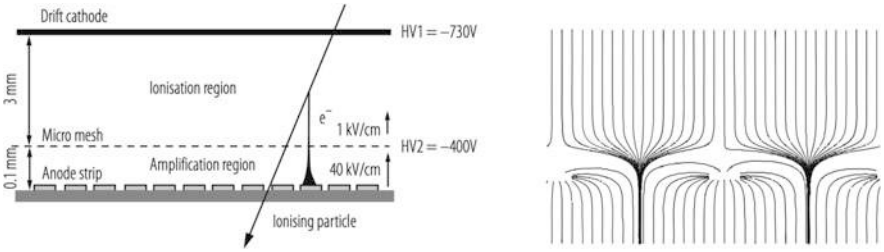


Fig. 4.30 Micromegas. Left: conceptual design. Right: Electrical field lines

implementing two 'wheels' of 8 m diameter with 4 layers of MICROMEGAs [80]. The readout readout strips of 300 μm width achieve a position resolution around 100 μm. In order to increase the stability against discharges for these very large surfaces, resistive strips are placed on top of the readout strips at a distance of 64 μm. The resistance value of 10–20 MΩ/cm ensures that the rate capability is sufficient for the application.

4.4 Outlook

The availability of large area silicon sensors has allowed most of the recent detector setups to use silicon trackers for vertexing and momentum spectroscopy in the detector volume upstream of the calorimeter systems. Muon systems do however have surfaces of up to several thousands of m^2 with particle rates and resolution requirements that make the application of gas detectors still the most viable solution. The TPC is still a very appealing detector for setups where very low material budget as well as PID capabilities are important requirements. Experiments such as NEXT [81] for the search of neutrinoless double beta decay are building on the unique features of gas detectors like low density of the detection medium and the related possibility for tracking of very low energy particles. Gas detector will therefore continue to be essential elements of particle physics instrumentation.

The last two sections on Resistive Plate Chambers and Micropattern Devices were updated in this new edition, while the remainder of this chapter is in its original form by H.J. Hilke.

References

1. E. Rutherford, H. Geiger, Proc. Roy. Soc. London A 81 (1908) 141.
2. H. Geiger, Verh. D. Phys. Ges. 15 (1913) 534.
3. H. Greinacher, Z. Phys. 23 (1924) 261.
4. H. Geiger, W. Mueller, Phys. Z. 29 (1928) 839.

5. A. Trost, Z. Phys. 105 (1937) 399.
6. (a) J.W. Keuffel, Rev. Sci. Instrum. 20 (1949) 202; (b) J. Christiansen, Z. Angew. Phys. 4 (1952) 326.
7. E. Fuenfer, H. Neuert, Zählrohre und Szintillationszähler, G. Braun, Karlsruhe (1959).
8. G. Charpak et al., *The Use of Multiwire Proportional Counters to select and localize Charged Particles*, Nucl. Instrum. Meth. 62 (1968) 262.
9. T. Bressani, G. Charpak, D. Rahm, C. Zupancic, *Track Localization by Means of a Drift Chamber*, Proc. Int. Seminar Filmless Spark and Streamer Chambers, Dubna, USSR, 1969, p. 275.
10. A.H. Walenta, J. Heintze and B. Schürlein, *The Multiwire Drift Chamber-A Novel Type of Proportional Wire Chamber*, Nucl. Instrum. Meth. 92 (1971) 373.
11. R. Veenhof, *Garfield-simulation of gaseous detectors*, http://consult.cem.ch/writeup/garfield/
12. I. Smimov, *Heed*: *Interactions ofparticles with gases*, http://consult.cem.ch/writeup/heed/
13. S. Biagi, *Magboltz*: *Transport of electrons in gas mixtures*, http://consult.cem.ch/writeup/magboltz/
14. The ALICE experiment at the CERN LHC, J. Instrum. 3 (2008) S08001.
15. The ATLAS Experiment at the CERN LHC, J. Instrum. 3 (2008) S08002.
16. The CMS experiment at the CERN LHC, J. Instrum. 3 (2008) S08003.
17. The LHCb Detector at the LHC, J. Instmm. 3 (2008) S08004.
18. Special volumes of Nucl. Instrum. Methods: the last three Proceedings are found in Nucl. Instrum. Meth. A 581 (2007) 1; 535 (2004) 1; 478 (2002) 1.
19. W. Blum, W. Riegler, L. Rolandi, Particle Detection with Drift Chambers, 2nd ed., Springer-Verlag (2008).
20. H. Bichsel, (a) Nucl. Instrum. Meth. A 562 (2006) 154; (b) Phys. Lett. B667 (2008) 267.
21. Review Particle Physics, Phys. Lett. B667(1–5) (2008) 292.
22. F. Lapique, F. Piuz, Nucl. Instrum. Methods 175 (1980) 297
23. H. Fischle, J. Heintze, B. Schmidt, Nucl. Instrum. Methods 301 (1991) 202
24. H. Walenta, Nucl. Instrum. Methods 161 (1979) 45.
25. I. Lehraus et al., Nucl. Instrum. Methods 153 (1978) 347
26. J. Va'vra et al. SLAC-PUB-5728, (1992).
27. S.F. Biagi, Nucl. Instrum. Meth. A 283 (1989) 716; ref. to J.C. Armitage, Nucl. Instrum. Meth. A 271 (1988) 588.
28. B. Schmidt, Doctoral thesis, Univ. Heidelberg, (1986).
29. G. Schultz, Doctoral thesis, Univ. Louis-Pasteur, Strasbourg, (1979).
30. Review Particle Physics, Physics Letters. B667(1–5) (2008) 1.
31. Compilation by A. Jeavons et al., Nucl. Instrum. Meth. 176 (1980) 89–97.
32. J.H. Hombeck, Phys. Rev. 84 (1951) 615.
33. (a) RD-32, CERN/DRDC 94-10 (1994); (b) A. Ishikawa, http://www-hep.phys.saga-u.ac.jp/ILC-TPC/gas/ps/
34. E.B. Wagner, F.J. Davies, F.J. Hurst, J. Chem. Phys. 47 (1967) 3138.
35. J.H. Parker, J.J. Lowke, Phys. Rev. 181 (1969) 290 and 302
36. PEP-4 Proposal, SLAC-Pub-5012 (1976).
37. S.R. Amendolia et al., Nucl. Instrum. Meth. 244 (1986) 516.
38. H.S.W. Massey, E.H.S. Burhop, H.B. Gilbody, Electronic and ionic impact Phenomena, Clarendon, Oxford (1969).
39. D.L. McCorkle, L.G. Christophorou, S.R. Hunter, in: Proc. 2nd Int. Swarm Seminar, Oak Ridge, USA, Pergamon, New York (1981) p. 21.
40. F. Bloch, N.E. Bradbury, Phys. Rev. 48 (1935) 689.
41. M. Atac, A.V. Tollestrup, D. Potter, Nucl. Instrum. Meth. 200 (1982) 345.
42. Collected by A. von Engel, Handbuch der Physik 21 (1956), Springer, Berlin.
43. M.E. Rose, S.A. Korff, Phys. Rev. 59 (1941) 850.
44. W. Diethom, USAEC Report NY 6628 (1956).
45. G.A. Erskine, Nucl. Instrum. Meth. 105 (1972) 565.
46. C. Brand et al. Nucl. Instrum. Meth. A 237 (1985) 501.

47. W. Legler, Z. Naturforschung A 16 (1961) 253.
48. H. Schlumbohm, Z. Phys. 151 (1958) 563.
49. G.D. Alkahazov, Nucl. Instrum. Meth. 89 (1970) 155.
50. S.C. Curran, A.L. Cockroft, J. Angus, Philos. Mag. 40(1949) 929.
51. S. Ramo, Proc. I.R.E. 27(1939) 584.
52. W. Shockley, J. Appl. Phys. 9 (1938) 635.
53. OPAL Collaboration, *The OPAL Detector at LEP*, Nucl. Instrum. Meth. A 305 (1991) 275.
54. (a) Proc. Workshop Radiation Damage to Wire Chambers, Berkeley (org. J. Kadyk), LBL 21170 (1986); (b) Proc. Int. Workshop Aging Phenomena in Gaseous Detectors, Hamburg, 2001, Nucl. Instrum. Meth. A 515 (2003); (c) NASA, a very large database on outgassing of materials can be reached via the NASA home- page.
55. (a) J. Va'vra, Nucl. Instrum. Meth. A 515 (2003) 263, and N. Tesch, IEEE Trans. Nucl. Sci. 49 (2002) 1609; (b) M. Capeans, Nucl. Instrum. Meth. A 515 (2003) 73; (c) R. Bouclier et al., Nucl. Instrum. Meth. A 350 (1994) 464.
56. J. Adam et al., Nucl. Instrum. Meth. 217 (1983) 291.
57. L. Malter, Phys. Rev. 50 (1936) 48.
58. C. Bemporad, Nucl. Instrum. Meth. 80 (1969) 205.
59. P. Schilly et al., Nucl. Instrum. Meth. 91 (1970) 221.
60. G. Charpak, D. Rahm, H. Steiner, Nucl. Instrum. Meth. 80 (1970) 13.
61. R. Bouclier et al., Nucl. Instrum. Meth. 115 (1973) 235
62. D.R. Nygren, Proposal to investigate a novel concept in particle detection, LBL intemal report, Berkeley, February 1974.
63. Proposal for a PEP Facility based on the Time Projection Chamber, PEP 4, Dec. 1976.
64. M. Anderson et al., *The STAR Time Projection Chamber*, Nucl. Instrum. Meth. A 499 (2003) 659.
65. (a) A. Shirahashi et al., *TOPAZ Time Projection Chamber*, IEEE Trans. NS-55 (1988) 414; (b) W.B. Atwood et al., *ALEPH Time Projection Chamber*, Nucl. Instrum. Meth. A 306 (1991) 446; (c) P. Abreu et al., *DELPHI Detector*, Nucl. Instrum. Meth. A 178 (1996) 57; (d) J. Alme et al., *The ALICE TPC*, Nucl. Instrum. Meth. A 622 (2010) 316-367; (e) a review on TPCs: H.J. Hilke, Rep. Prog. Phys. 73 (2010) 116201.
66. V.V. Parkhomchuk, Y.N. Pestov and N.V. Petrovykh, A spark counter with large area, Nucl. Instr. and Meth. A 93 (1971) 269–270.
67. Yu.N. Pestov, Status and future developments of spark counters with localized discharge, Nucl. Instr. and Meth. A 196 (1982) 45–47.
68. R. Santonico, R. Cardarelli, Development of Resistive Plate Counters, Nucl. Instr. and Meth. A 187 (1981) 377.
69. A. Bertolin et al., The RPC system of the OPERA experiment, Nucl. Instr. and Meth. A 602 (2009) 631–634.
70. J. Cao and Kam-Biu Luk, An overview of the Daya Nay reactor neutrino experiment, Nuclear Physics B 908 (2016) 62–73.
71. G. Aielli et al., Layout and performance of the RPCs used in the Argo-YBJ experiment, Nucl. Instr. and Meth. A 562 (2006) 92–96.
72. W. Riegler et al., Detector physics and simulation of Resistive Plate Chambers, NIMA 500 (2003) 144–162.
73. F. Sauli, Nucl. Instrum. Meth. A 386 (1997) 531.
74. B. Ketzer et al., Nucl. Instrum. Meth. A 535 (2004) 314.
75. N. Abgrall et al., Time projection chambers for the T2K near detectors, NIMA 637 (2011) 25–46.
76. V. Tikhonov, R. Veenhof, Nucl. Instrum. Meth. A 478 (2002) 452.
77. M.M. Aggarwal et al. NIM A 903 (2018) 215.
78. D. Abbaneo et al., Upgrade of the CMS muon system with tripple-GEM detectors, JINST 9 C10036.
79. Y. Giomataris, Nucl. Instrum. Meth. A 419 (1998) 239.

80. F. Kuger et al., Performance studies of the resistive Micromegas detectors for the upgrade of the ATLAS Muon spectrometer, NIMA 845 (2017) 248–252.
81. J.J. Gomez-Cadenas, The next experiment, NPPP 273-275 (2016) 1732–1739.
82. IAEA-TECDOC-799 (1995) 560, Atomic and Molecular Data for Radiotherapy.
83. A. Akindinov et al., The multigap resistive plate chamber as a time-of-flight detector, Nucl. Instr. and Meth. A 456 (2000) 16.

Chapter 5
Solid State Detectors

G. Lutz and R. Klanner

5.1 Introduction

Semiconductor detectors, and in particular silicon detectors, are very well suited for detection and measurement of light and of ionizing radiation caused by interaction with charged particles and (X-ray) photons. Precise position, time and energy measurement can be combined when use is made of the excellent intrinsic material properties in well thought out detector concepts.

Development and large scale use of silicon detectors has been initiated by particle physics. The discovery of the rare and short lived charmed particles lead to the desire to use their decay topology as signature for identification and separation from non-charm background. Detectors were required that combined very good position measurement (in the range of several μm) with high rate capability (few hundred kHz), a task not achievable with available detectors at that time.

Semiconductor detectors, in particular silicon and germanium detectors were used for quite some time, but not too frequently, in Nuclear Physics for the

In the updated version a number of detector developments which took place after the publication of the original version have been taken into account. These are in particular new sections on Radiation Damage, 3-D Detectors, MAPS (Monolithic Active Pixel Sensors), SiPMs (Silicon Photomultipliers) and Ultrafast Tracking Detectors (LGAD = Low Gain Avalanche Detectors). In addition, the section Summary and Outlook has been updated.

The author G. Lutz is deceased at the time of publication.

G. Lutz
PNSensor GmbH and MPI-Halbleiterlabor, Munich, Germany

R. Klanner (✉)
Department of Physics, University of Hamburg, Hamburg, Germany
e-mail: robert.klanner@desy.de

C. W. Fabjan, H. Schopper (eds.), *Particle Physics Reference Library*,
https://doi.org/10.1007/978-3-030-35318-6_5

purpose of measuring particle and X-ray photon energies, not however for position measurement. This task was left mostly to gas detectors and to scintillation hodoscopes, both of them not able to provide the required position measurement resolution.

It was realized rather soon that semiconductors offer in principle the required capabilities and silicon strip detectors were developed and used for the detection and investigation of charmed particles. This development rapidly increased in speed and scope so that today it is rare to find particle physics experiments that do not rely heavily on silicon strip detectors for particle tracking and identification. Strip detectors have also entered many other fields of science. Important features of this development were the introduction of more sophisticated detector concepts and the development of multi-channel low noise-low power integrated readout electronics adapted to the requirements of strip detectors.

A further challenge in particle tracking poses the ambiguities occurring in case of high particle densities. This problem is alleviated considerably when replacing the strip geometry by pixels. Hybrid pixel detectors became possible with the enormous progress in miniaturization of electronics. Each pixel has its own readout channel. Detector and electronics with matched geometry are connected face to face by bump bonding. Recently Monolithic Active Pixel Sensors (MAPS), pixel detectors in which sensor and readout electronics are integrated on the same silicon chip, are reaching maturity.

Although in the initial phase of this rapid development position measurement was in the focus of interest, energy resolution with high readout speed came back to its right, sometimes in combination with position resolution. This development opened the door of semiconductor detectors in X-ray astronomy, synchrotron radiation experiments and in many other fields.

A major step on this way was the invention by E. Gatti and P. Rehak of the semiconductor drift chamber [1]. This concept also became the basis for further new concepts as are the pnCCD [2], the silicon drift diode [3] and the DEPFET [3] that forms the basis for several types of pixel detectors with rather unique properties.

In the last decade, a major progress in the field of silicon photo-detectors took place: Multi-pixel avalanche photo diodes operating in the Geiger mode, frequently called silicon photo-multipliers, SiPM, have been developed and found many applications in research, medicine and industry.

In the following, detection principles and properties of the various detector types will be described and some applications will be sketched. Emphasis is on detector physics and concepts while it is impossible to cover all important activities in the field. In addition, a short summary of radiation damage, which presents a major challenge for the use of silicon detectors in the harsh radiation environment at colliders, like at the CERN Large Hadron Collider, LHC, will be presented.

5.2 Basic Detection Process of Single Photons in Semiconductors

The simplest detector is a reverse-biased planar diode (Fig. 5.1). Photons interacting in silicon will, dependent on their energy, produce one or more electron-hole pairs close to their points of interaction. Charged particles will generate pairs along their path within the semiconductor. An average energy of 3.6 eV is needed for creation of a pair in silicon with a band gap of 1.12 eV at room temperature. This should be compared with the ionization energy of gases which is more than an order of magnitude higher. Electrons and holes will be separated by the electric field within the space charge region and collected at the electrodes on opposite sides of the diode.

The small band gap and the corresponding large signal charge generated in the photon absorption process is the principal cause for the excellent properties of semiconductor radiation detectors manifesting themselves in particular in very good spectroscopic resolution down to low energies. Further reasons are the high density and corresponding low range of delta-electrons which makes very precise position measurement possible. High charge carrier mobilities combined with small detector volume leads to short charge collection time and makes the use of detectors in high rate environment possible. The excellent mechanical rigidity makes the use of gas containment foils superfluous and allows operation in the vacuum. Therefore very thin entrance windows can be constructed and high quantum efficiency can be reached down to low photon energies. Position dependent doping of semiconductors allows construction of detectors with sophisticated electric field configurations and intrinsically new properties.

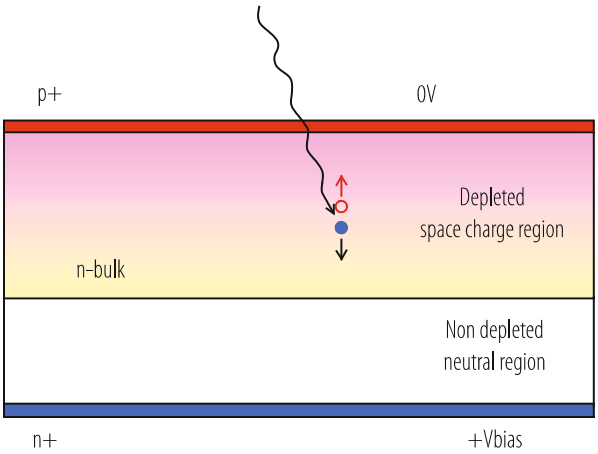


Fig. 5.1 Schematic structure of a reverse-biased semiconductor diode used as photon detector. The region heavily doped with acceptors is denoted p^+, and n-bulk and n^+ the regions lightly and heavily doped with donors, respectively

Reaching all these good detector properties requires a readout electronics which is well matched to the detectors. Here we notice a point specific to silicon which is also the basic material of most of present day electronics. For that reason it is natural to integrate the sensitive front-end part of electronics into the detector. This is the case, for example, in CCDs and drift diodes [3] with very high spectroscopic resolution. A further device (DEPFET) [3] combines the function of detector and amplifier in the basic structure. In MAPS (Monolithic Active Pixel Sensors) sophisticated readout electronics is directly integrated on the silicon chip of the sensor.

Dependent on the field of application different aspects of semiconductors are in the focus of interest. In particle physics tracking requires high position resolution and often high speed capabilities while energy resolution is of less importance. Recently at the CERN LHC also a timing accuracy of a few tens of picoseconds in combination with precision tracking became a requirement. In X-ray spectroscopy and imaging, as well as in X-ray astronomy both energy and position resolution are of importance. For light detection, high photon-detection efficiency and resolving single photons are typically more important than position accuracy.

5.3 Basics of Semiconductor Physics

After these introductory remarks on semiconductor detectors we will look into the underlying mechanisms in a little more detail.

Most commonly used semiconductors are single crystals with diamond (Si and Ge) or zinc blende (GaAs and other compound semiconductors) lattice. Each atom in the crystal shares their outermost (valence) electrons with the four closest neighbours. At very low temperature all electrons are bound to their respective locations and the material is an insulator. At elevated temperature thermal vibrations will sometimes break a bond and both the freed electron and the hole (the empty place left behind to be filled by a neighbouring electron) are available for electrical conduction. The density of free electrons/holes is called intrinsic carrier density n_i. For silicon its value at room temperature is about 10^{10} cm^{-3}, resulting in an intrinsic resistivity of about 350 kΩ·cm.

Creation of electron–hole pairs can also be accomplished by electromagnetic radiation or by the passage of charged particles knocking out of their covalent bond some of the valence electrons. This is the mechanism used in the detection process. These free charge carriers will then be moved by an applied electrical field (drift) and redistribute due to concentration variations (diffusion) until finally reaching an external electrode connected to the readout electronics.

So far we have only dealt with intrinsic semiconductors, perfect crystals without foreign atoms. One may, however, replace a small fraction of atoms with some having either one more, called donors (e.g. P in Si) or one less, called acceptors (e.g. B in Si) valence electron. The additional electron or the missing electron (hole) is only weakly bound, resulting in states in the silicon band gap located about 40 meV

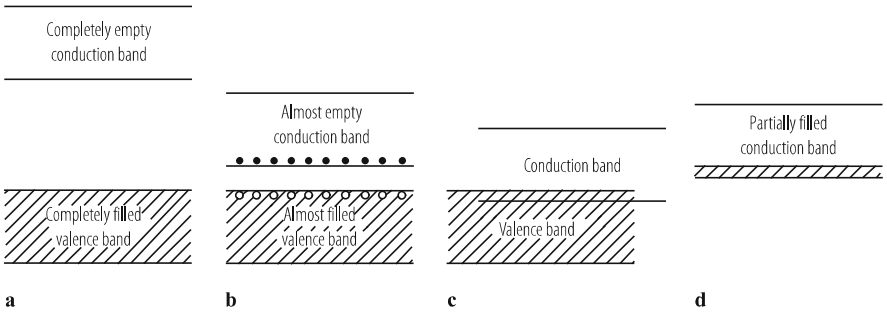


Fig. 5.2 Energy band structure of insulators (**a**), semiconductors (**b**), and conductors (**c**, **d**)

from the conduction or valence band, respectively. silicon doped with donor atoms is called *n*-type, and *p*-type for acceptors. These states are already ionized well below room temperature, and the electrons or holes can move freely in the silicon lattice, resulting in a decrease of the resistivity. For silicon detectors crystals with a typical doping density of 10^{12} cm^{-3} are used, which results in a similar density of free charge carriers and a significantly reduced resistivity of a few kΩ·cm. Applying an external electric field the free charge carriers can be removed and a space charge region due to the surplus charge of the doping atoms is created.

The discussion so far has used the simple bond picture. A more sophisticated treatment that allows also quantitative calculations requires the quantum mechanical band model. While single atoms possess discrete energy levels, in crystals these are transformed into energy bands.

Figure 5.2 shows the (almost) fully occupied valence band and the lowest laying (almost) empty conduction band for insulators, semiconductors and conductors. In insulators (a) valence and conduction band are separated by a big band gap so that electrons cannot be thermally excited from the valence to the conduction band. Conductors have overlapping bands (c) or a partially filled conduction band (d) and are therefore electrically conducting.

In intrinsic (undoped) semiconductors only a small fraction of the electrons in the valence band are thermally excited into the conduction band. Extrinsic (doped) semiconductors have additional localized energy states within the band-gap. Donor states close to the conduction band (e.g. P in Si) emit their electrons into the conduction band and are (almost) completely ionized (positively charged) already well below room temperature. Acceptor states close to the valence band trap electrons and leave holes in the valence band.

In thermal equilibrium the occupation probability F of states with energy E at temperature T follows from Fermi statistics

$$F(E) = 1/\left(1 + \exp\left(E - E_{\mathrm{f}}\right)/kT\right), \tag{5.1}$$

with k the Boltzmann constant. The overall charge neutrality determines the Fermi level E_{f}.

Electrons bound in one of the localized donor states may be emitted into the conduction band by thermal excitation with a probability ε_n, thereby ionizing donors. Ionized donors may also capture electrons out of the conduction band. This process is described by a capture cross section σ_n. In thermal equilibrium these two processes have to balance each other. That condition allows to derive a relation between emission probability ε_n and capture cross section σ_n:

$$\varepsilon_\mathrm{n} = \sigma_\mathrm{n}\, \nu_{\mathrm{th\,n}}\, n_\mathrm{i}\, \exp\left(\left(E_\mathrm{d} - E_\mathrm{i}\right)/kT\right), \tag{5.2}$$

with $\nu_{\mathrm{th\,n}}$ thermal velocity of electrons in the conduction band, n_i intrinsic carrier concentration, E_d donor energy level, E_i intrinsic energy (Fermi level for an intrinsic semiconductor). This relation is valid more generally and can be applied to non-equilibrium conditions.

Electrons in the conduction band and holes in the valence band can move freely within the crystal lattice, their movement being only retarded by scattering on imperfections of the lattice. These imperfections may be due to lattice defects, doping atoms replacing regular atoms of the crystal (substitutional dopands) and distortions of the lattice due to thermal vibrations. The simplified way of describing these effects uses the assumption that charge carriers are accelerated by the electric field and lose all previous history at each scattering, starting with random thermal velocity again.

The movement due to the electric field is described by the drift velocity that for low fields can be assumed to be proportional to the electric field:

$$\nu_\mathrm{n} = \left(-q\tau_\mathrm{c}/m_\mathrm{n}\right) E = -\mu_\mathrm{n}\, E, \qquad \nu_\mathrm{p} = \left(q\tau_\mathrm{c}/m_\mathrm{p}\right) E = \mu_\mathrm{p}\, E, \tag{5.3}$$

with ν_n , ν_p , μ_n , μ_p being the drift velocities and low-field mobilities of electrons and holes, respectively, q elementary charge, τ_c average time between collisions, m_n , m_p effective masses of electrons and holes, and E electric field. For a high electric fields τ_c decreases and the drift velocity saturates.

At very high electric field electrons and holes may acquire sufficient energy in between collisions to generate additional electron hole pairs. This avalanche process can be the cause for an electrical breakdown of devices. It may also be used as an intrinsic amplification process in order to get sufficiently high signals from very small ionization.

For inhomogeneous carrier distributions charge carriers will preferably diffuse from high concentrations to regions of lower concentrations. This diffusion mechanism is described by

$$F_\mathrm{n} = -D_\mathrm{n}\, \nabla n, \qquad F_\mathrm{p} = -D_\mathrm{p}\, \nabla p. \tag{5.4}$$

With F_n , F_p flux of electrons and holes, D_n , D_p diffusion constant. Electron and hole current densities due to drift and diffusions are given by

$$J_n = -q\mu_n\, nE + q\, D_n\, \nabla n, \qquad J_p = q\mu_p\, pE - q\, D_p\, \nabla p. \tag{5.5}$$

Diffusion constant and mobility are related by Einstein's relation $D = (kT/q)\ \mu$. It can be derived from the requirement of zero current in thermal equilibrium of a device with non-uniform doping that has to have a constant Fermi level.

In the absence of magnetic fields charge carriers will move approximately parallel (holes) or antiparallel (electrons) to the electric field. The magnetic field adds a force perpendicular to the direction of motion and to the magnetic field direction so that the charge carriers move at an angle $\theta_p = \mu^{H}{}_p\, B$, $\theta_n = \mu^{H}{}_n\, B$ with respect to the drift direction. The Hall mobilities $\mu^{H}{}_p$ and $\mu^{H}{}_n$ differ from the drift mobilities μ_p and μ_n. B is the magnetic field component perpendicular to the electric field and the particle velocity.

5.4 Radiation Damage

Damage by ionizing and non-ionizing radiation, represents a major limitation for the use of silicon detectors in the harsh radiation environment of high-luminosity colliders, like the CERN LHC, where after its upgrade, fluences exceeding 10^{16} cm^{-2} and dose values up to 5 MGy will be reached. At high-brilliance X-ray sources, like the European X-ray Free-Electron Laser at Hamburg, dose values up to 1 GGy are expected. Radiation damage is classified in surface and bulk damage.

Surface damage is caused by ionization by charged particles and X-ray photons in the insulating layers, e.g. the SiO_2, required to fabricate silicon sensors. Like in the silicon bulk, ionizing radiation produces electron-hole pairs in the SiO_2. Whereas the mobility of electrons is sufficiently high so that they can move to a nearby electrode, holes are trapped, which results in a positive charge layer and interface traps at the Si-SiO_2 interface [4]. Positive surface charges can result in an electron accumulation layer in the Si at the interface, which can cause shorts between electrodes or break down. Interface traps, if exposed to an electric field, produce surface-generation currents. As the exact conditions at the Si-SiO_2 interface also depend on the potential on the outer SiO_2 surface, which in particular in dry conditions has a very high surface resistance (sheet resistance $> 10^{18}\ \Omega_{\square}$), it can take days until equilibrium is reached [5]. The result can be a breakdown after several days of operation or a humidity-dependent breakdown voltage. Surface radiation damage also depends on the dose-rate, which together with long time constants has to be taken into account, when studying surface damage or when testing silicon detectors. Surface damage is also technology dependent. In addition, already at room temperature significant annealing takes place. All these effects make a systematic study of surface-radiation damage difficult and time consuming. However, also thanks to the methods developed for radiation-hard electronics,

surface-radiation effects are sufficiently well understood and can be avoided by a proper design [6]. Nevertheless, there are many examples of improper designs and several unpleasant surprises due to surface damage.

Non-ionizing interactions, which knock out silicon atoms from their lattice points, are the main cause of bulk damage. A minimum energy transfer to the silicon atom of about 25 eV is required to produce such a primary defect. For energy transfers above 1 keV the silicon atom itself can knock out further silicon atoms, resulting in defect clusters, and for energies above 12 keV multiple clusters can be produced. These threshold numbers are the result of model calculations and only limited experimental information is available. The primary defects are mobile at room temperature. Some of them anneal, others diffuse to the silicon surface or interact with crystal defects and impurities and form stable defects. Using different spectroscopic methods a large number of defects could be identified and their properties, like donor- or acceptor-type, position in the band gap, cross-sections for electrons and holes and introduction rates determined [7]. The electrically active defects have three main consequences for detectors: (1) Increase of dark current, (2) trapping of signal charges thus reducing the charge collection, and (3) change of the electric field in the space charge region from which the signal charge is collected.

Typical introduction rates of stable defects are of order 1 cm^{-1}, i.e. a fluence if 1 particle per cm^2, produces 1 stable defect per cm^3. For fluences above about 10^{14} cm^{-2}, the density of defects exceeds by far the doping density, and the silicon properties change significantly: In non-depleted silicon the high generation-recombination rate results in an approximately equal density of holes and electrons, and the resistivity increases from the value determined by the dopant density to the value of intrinsic silicon, which is about 350 kΩ·cm at room temperature. The high dark current for a reverse biased diode, which is dominated by holes at the cathode and by electrons at the anode, results in a position-dependent filling of the defects and a completely different electric field distribution than in the detector before irradiation. High field regions appear at anodes and cathodes, a phenomenon called "double junction", and lower field regions in-between [8, 9]. Thus the concept of uniform doping breaks down and most of the methods used to characterize silicon before irradiation are no more applicable.

Based on a detailed and systematic study of silicon pad diodes with different doping and impurities irradiated by different particles and fluences, the phenomenological *Hamburg model* has been developed [10]. It parametrises the change of parameters like dark current and effective doping, used to characterise non-irradiated sensors, as a function of irradiation fluence and temperature history. Up to fluences of approximately 10^{14} cm^{-2}, which are presently (mid 2018) reached at the LHC, the model is remarkable successful in describing the observed effects of radiation damage. An extension of such a model to higher fluences is badly needed for monitoring the radiation fields at the LHC and for the planning of the experiments at the High-Luminosity LHC.

5.5 Semiconductor Detector Principles

The very basic and most common detector type, the reverse-biased diode, has already been sketched in Sect. 5.2. Here we will give some more information on this device and also present some more sophisticated principles, the semiconductor drift chamber and the DEPFET detector-amplification structure, while detectors based on the avalanche mechanism will be discussed in a later chapter.

5.5.1 Reverse Biased Diode (as Used in Strip and 3-D Detectors)

The principle of a reverse biased diode has already been sketched in Sect. 5.2. Here a more detailed discussion is given. Even without applying a bias the *p*-*n* junction develops a space charge region due to the diffusion of electrons and holes across the junction leading to a surplus of negative charge on the *p*-side and of positive charge on the *n*-side of the junction. This creates an electric field, a drift current and a space charge region on both sides of the junction. At any point of the device drift and diffusion currents cancel each other in equilibrium without external bias. Such a device can already be used as radiation detector since electron–hole pairs created in the space charge region will be separated by the electric field thus create a current across the junction.

Reverse biasing will increase the space charge region and therefore the electric field. For a strongly asymmetric, but in each region uniformly doped p^+n junction (as shown in Fig. 5.1) the depth of the space charge and therefore the sensitive region increases with the square root of the applied voltage.

Reverse biased diodes have been used as energy sensitive radiation detectors in Nuclear Physics for quite some time. The real breakthrough came with strip detectors in Particle Physics used for particle tracking with micro-meter accuracy. Many small strip-like diodes were integrated on the same wafer and each one connected to its own readout channel (Fig. 5.3). The particle position was given by the channel giving the signal. More sophisticated strip detectors will be described in Sect. 5.6.

Planar pixel detectors are obtained by shortening the individual strips so that they do not reach anymore the detector edge and form a two-dimensional pattern. Detectors with pixel sizes down to $15 \times 15\ \mu m^2$ have been built. The main difficulty of such detectors is their readout. Different realisations will be discussed later.

A different concept of diode detectors, the so called 3-D detectors [11], is shown in Fig. 5.4: Holes with diameters of a few micro-meters are etched into the crystal orthogonal to its surface, and alternate holes are n^+- and p^+-doped. A voltage difference between the n^+- and p^+-doped columns generates an electric field parallel to the crystal surface. The number of electron-hole pairs produced by a charged particle traversing the detector at large angles to the surface is given by

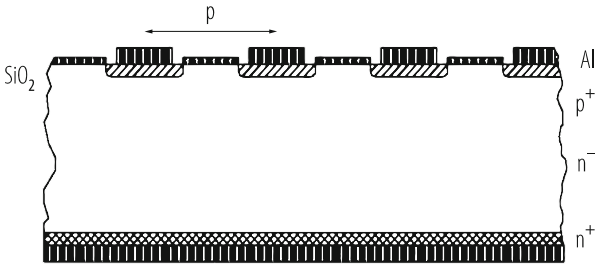


Fig. 5.3 Cross section of a silicon strip detector built on lightly phosphor doped (n^-) silicon bulk material. Strips are highly boron doped (p^+) and the backside highly phosphor doped (n^+)

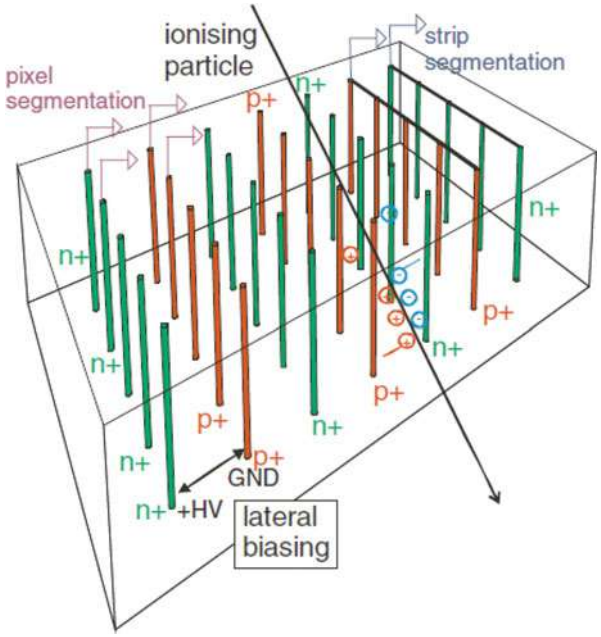


Fig. 5.4 Principle of the 3-D detector: Holes are etched into the silicon crystal orthogonal to the detector surface. Alternate holes are n^+- and p^+-doped. A positive voltage on the n^+-contacts, with the p^+-contacts grounded, generates an electric field parallel to the detector surface. In a 3-D detector the charge generated, given by the crystal thickness, and the charge collection distance, given by the distance between the holes, can be separately chosen (Book F. Hartmann Fig. 1.69)

the crystal thickness, whereas the charge collection distance is given by the column distance. In this way signal and charge collection distances can be chosen separately and the detector can be optimised for radiation tolerance. In addition, the operating voltage for 3D-detectors and thus the power heating the detector are significantly reduced compared to planar detectors. By connecting the p^+- and n^+-columns with

different metal patterns, strip- and pixel-sensors and other readout geometries can be realized.

5.5.2 Semiconductor Drift Chamber

The semiconductor drift chamber has been invented by Emilio Gatti and Pavel Rehak [1]. This device (Fig. 5.5) makes use of the sideward depletion principle, having diode junctions on both surfaces and a bulk contact on the fringe. Fully depleting the device by applying a reverse bias voltage between p- and n-contacts creates a potential valley for electrons in the middle plane. Electrons created by ionizing radiation will assemble in this valley and subsequently diffuse until they eventually reach the n-doped anode. Faster and controlled collection is achieved by adding a horizontal drift field. This is obtained by dividing the diodes into strips and applying from strip to strip increasing voltages.

This device is able to measure position (by means of the time difference between particle interaction and arrival of the signal at the anode) as well as the energy from the amount of signal charge. In many applications the latter aspect is the important one. Here one profits from the small electric capacitance of the anode compared to the planar diode shown in Fig. 5.1, which acts as capacitive load to the readout amplifier. Large area detectors can therefore be operated with excellent energy resolution at high rates.

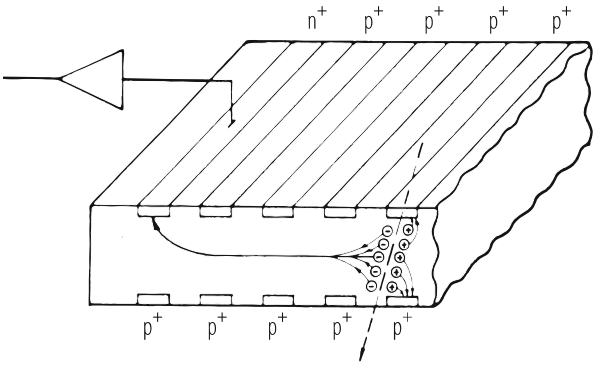


Fig. 5.5 Semiconductor drift chamber using the sideward depletion method. Dividing the p^+ doped diodes into strips and applying a potential which increases from strip to strip superimposes a horizontal field in the potential valley that drives the electrons towards the n^+ anode which is connected to the readout electronics. Upon arrival of the signal charge at the n^+ anode the amount of charge and the arrival time can be measured

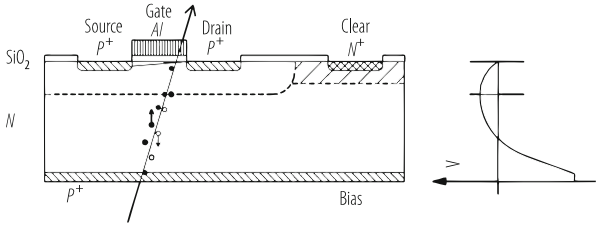


Fig. 5.6 The concept of a DEPFET: Simplified device structure (*left*) and potential distribution along a cut across the wafer in the gate region of the transistor (*right*)

5.5.3 DEPFET Detector-Amplification Structure

The *DEPFET* structure which simultaneously possesses detector and amplification properties has been proposed by J. Kemmer and G. Lutz in 1987 [3] and has subsequently been confirmed experimentally [12]. It is based on the combination of the sideward depletion method—as used in a semiconductor drift chamber shown in Fig. 5.5—and the field effect transistor principle.

In Fig. 5.6 a *p*-channel transistor is located on a fully depleted *n*-type bulk. Compared to Fig. 5.5 the potential valley has been moved close to the top side. Signal electrons generated in the fully depleted bulk assemble in a potential minimum for electrons ("internal gate") and increase the transistor channel conductivity in a similar way as by changing the (external) gate voltage. The device can be reset by applying a large positive voltage on the clear electrode.

The DEPFET has several interesting properties:

- Combined function of sensor and amplifier;
- Full sensitivity over the complete wafer;
- Low capacitance and low noise;
- Non-destructive repeated readout;
- Complete clearing of the signal charge: No reset noise.

These properties make it an ideal building block for an X-ray pixel detector, or for a pixel detector for the precision tracking of charged particles.

5.6 Silicon Strip Detectors (Used in Tracking)

Silicon strip and pixel detectors are the most common semiconductor detectors in Particle Physics, mostly used for particle tracking. There one profits from the precise position measurement (few μm) at very data rates (up to tenths of MHz per detection element). In its simplest form they are narrow strip diodes put next to each other on the same semiconductor substrate, each strip having its own readout channel. Typical charge collection times are about 10 ns. Due to diffusion, track

inclination and the Lorenz force in a magnetic field, the charge of one track may be distributed over two or more strips. This can be exploited to improve the accuracy of the position measurement well below the value given by the strip pitch. It is then limited by fluctuations of the ionization process and in particular the generation of delta electrons and the electronics noise. A measurement precision down to about 1 μm has been achieved.

5.6.1 Strip Detector Readout

In the conceptually simplest version each strip is connected to its own electronic readout channel and the position is determined by the number of the strip providing a signal.

Binary (yes/no) *readout* may be used if no energy information is required and if the position accuracy given by the strip pitch is sufficient. One also does not lose position resolution compared with analogue readout if the strip pitch is large with respect to the width of the diffusion cloud.

Analogue (signal amplitude) *readout* of every channel may lead to a substantial improvement of the position measurement precision if the strip spacing matches the charge spread due to diffusion during collection. (Charge spread can also be due to track inclination or the Lorenz angle in a magnetic field.) In addition, the simultaneous measurement of energy loss becomes possible.

Charge division readout reduces the number of readout channels as only a fraction of the strips is connected to a readout amplifier (Fig. 5.7). Charge collected at the other (interpolation) strips is divided between the two neighbouring readout channels according to the relative position. Charge division is due to the capacitors between neighbouring strips. For charge division to work, it is necessary to hold the intermediate strips at the same potential as the readout strips. This can be accomplished by adding high ohmic resistors or with other methods. If the intermediate strips were left floating, they would adjust themselves to a potential

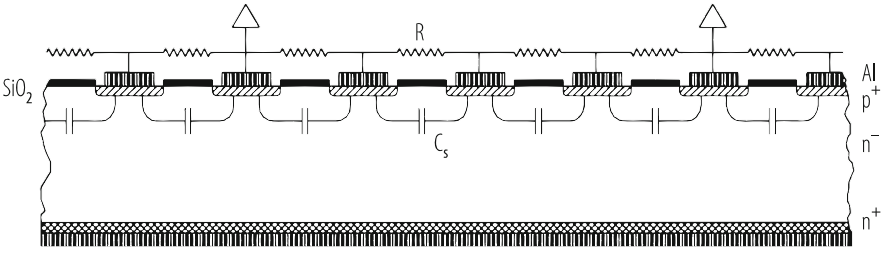


Fig. 5.7 Charge division readout. The interstrip capacitors between the readout strips act as capacitive charge divider. The high-ohmic resistors are required to keep all strips at the same potential

such that they would collect no signal charge, and thus charge division would cease to function.

5.6.2 *Strip Detectors with Double-Sided Readout*

As shown in Fig. 5.8 it is possible to segment the electrodes on both sides of the wafer. This double-sided readout has the obvious advantage of providing twice the information for the same amount of scattering material. With crossed strips on the two detector faces, a projective two-dimensional measurement is obtained from a single detector.

For a traversing particle, a spatial point can be reconstructed as both projections are obtained from the same initial charge cloud. With analogue readout it is furthermore possible (to some degree) to correlate signals from the two sides, making use of Landau fluctuations and the equality of the charge induced on both sides for each ionizing particle. This can be of interest for resolving ambiguities when several particles traverse simultaneously the detector.

A problem in producing double-sided detectors is the insulation of neighbouring strips on both detector sides. The naive solution of only providing highly doped *n*- and *p*-doped strips on the two sides of the detector (Fig. 5.8) fails because of the build-up of an electron-accumulation layer (an inversion layer on *p*-type silicon) between the *n*-strips below the insulating SiO_2 (Fig. 5.9a). This electron layer results in an electrical shortening of neighbouring strips. It is caused by the positive charges that are always present at the silicon-oxide interface. As discussed in Sect. 5.4 ionizing radiation results in a further increase of positive charges.

There are three possibilities for curing the problem:

1. Large-area *p*-type surface doping. In this case the oxide charges are compensated by the negative acceptor ions and the build-up of the electron layer is prevented (Fig. 5.9b). This method requires a delicate choice of *p*-type doping concentration and profile. A too large doping results in high electric fields and in a possible

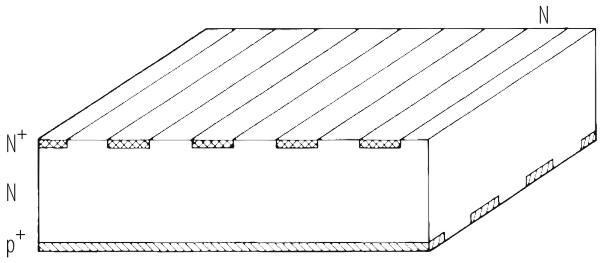


Fig. 5.8 Double sided strip detector (naive solution)

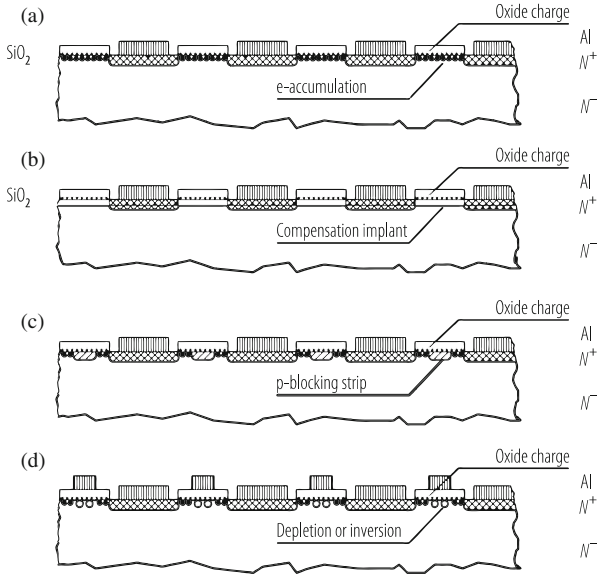


Fig. 5.9 Insulation problem for *n*-strips in silicon, due to electrical shortening by an electron accumulation layer (**a**), and three possible solutions: Large area *p*-implantation (**b**); interleaved *p*-strips (**c**) and negatively biased MOS structures (**d**)

electrical breakdown at the strip edges. This problem is alleviated by the other two solutions presented below.

2. Disruption of the electron layer by implantation of *p*-strips between the *n*-doped charge-collection strips (Fig. 5.9c); and
3. Disruption of the electron layer by a suitably biased (negatively with respect to the *n*-strips) MOS structure (Fig. 5.9d). For moderate biasing neither electrons nor holes will accumulate underneath the MOS structure, while for a high negative bias a hole layer (inversion on *n*-type, accumulation on *p*-type silicon) will form.

5.6.3 Strip Detectors with Integrated Capacitive Readout Coupling and Strip Biasing

Capacitive-coupled (AC) readout (Fig. 5.10, *right*) has the obvious advantage of shielding the electronics from dark current, whereas direct coupling (DC, Fig. 5.10, *left*) can lead to pedestal shifts, a reduction of the dynamic range, drive the electronics into saturation or requires a dark-current compensation.

As it is difficult to fabricate high-ohmic resistors, and almost impossible to produce sufficiently large capacitors in LSI electronics, it seemed natural to integrate

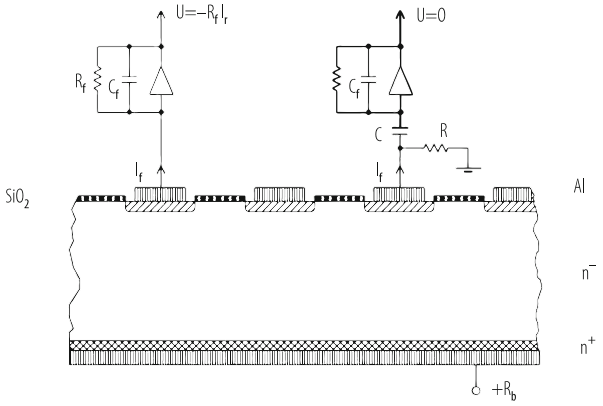


Fig. 5.10 Direct and capacitive coupling of electronics to the detector. With direct coupling (*left*) the detector reverse bias current I_f has to be absorbed by the electronics. With capacitive coupling (*right*), only the AC part of the detector current reaches the electronics, while the DC part flows through the resistor R

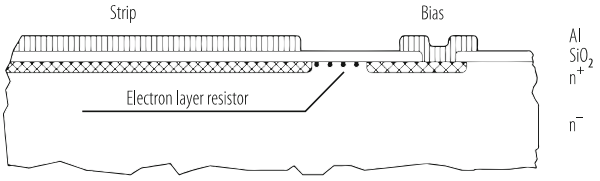


Fig. 5.11 *n*-strip biasing by an electron-accumulation-layer resistor. The diagram shows a cut along the strip direction. The electron layer is induced by the always present positive oxide charges that attract electrons towards the Si-SiO_2 interface. It is sidewise enclosed by p-implants so as to prevent electrical shortening between neighbouring strips. Bias and strip implants are at nearly the same potential

these elements into the detector. This has been done in a collaborative effort by a CERN group with the Center of Industrial Research in Oslo [13], where the detectors were produced. Capacitances have been built by separating implantation and metallization of the strips by a thin SiO_2 layer. Biasing resistors were made of lightly doped polysilicon, a technology that is used in microelectronics. The detectors gave very satisfactory results. The strip detectors of several particle physics experiments use this design.

A different method of supplying the bias voltage to the detector has been developed and used for double-sided readout by a Munich group [3, 12]. It leads to a considerable simplification of the technology as it does not require resistors but only uses technological steps that are already required for DC coupled detectors. The polysilicon technology can be avoided altogether; instead, the voltage is supplied through the silicon bulk. Two methods can be applied either using the resistance of an electron accumulation layer (Fig. 5.11) that is induced by the positive oxide charge or a punch through mechanism that occurs between two closely spaced *p*-

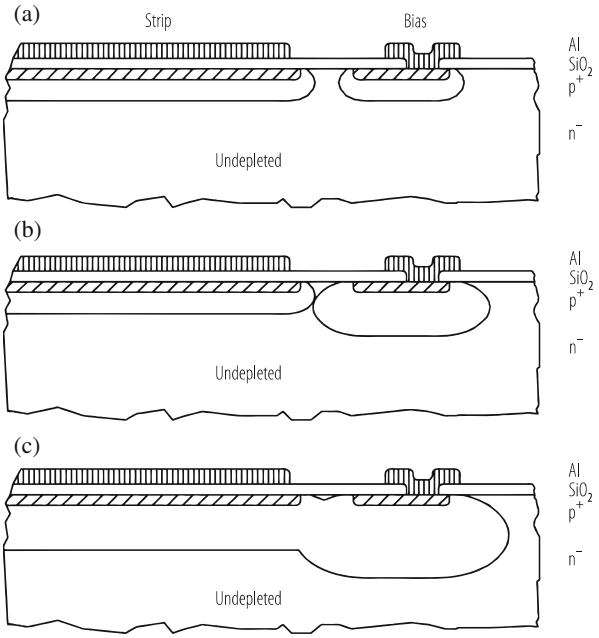


Fig. 5.12 *p*-strip punch-through biasing. The diagrams show cuts along the strip direction: (**a**) Before applying a bias voltage, where the space-charge regions around the strip and the bias implant are isolated; (**b**) at onset of punch-through, where the space-charge region around the bias implant has grown and just touches the space-charge region of the strip. The potential barrier between strip and bias implants has diminished, but is just large enough to prevent the thermal emission of holes towards the bias strip; (**c**) at larger bias voltage, where the space-charge region has grown deeper into the bulk. Holes generated in the space-charge region and collected at the strip implant are thermally emitted towards the bias strip. The voltage difference between strip implant and bias depends on geometry, doping and bias voltage. A weak dependence on oxide charge is also present

electrodes (Fig. 5.12). These biasing methods can be used for single sided and also for double sided readout where *p*- and *n*-strips are located at opposite surfaces of the wafer as was the case in the ALEPH experiment. In all cases the capacitors are built by interleaving a thin oxide layer between implantation and metal strips.

A word of caution on the operation of capacitive-coupled detectors and in particular of double sided detectors will be given at this point since it has been overlooked in a couple of experiments causing detector breakdown. At first glance it seems that one can choose the voltages on implant and metal strips independently. However this can result in shortening of neighbouring strips or electrical breakdown due to the build-up of accumulation layers at the Si-SiO_2 interface. Although the SiO_2 is not covered with an ohmic layer its surface will slowly charge up to a potential close to the neighbouring metal electrodes, because of a high but finite surface resistivity, as discussed in Sect. 5.4.

5.7 Detector Front-End Electronics

Before discussing more sophisticated detectors we now turn to readout electronics, a subject relevant to all detectors. As there is a close interplay between a detector and its electronics, both components have to be considered together when designing a detector for a specific application. In most cases a signal charge produced by photons or ionizing radiation has to be measured as precisely as possible in a predefined time interval and with tolerable power consumption. Readout uses in most cases large scale integrated (LSI) electronics adapted to the needs of the special application.

5.7.1 Operating Principles of Transistors

Transistors are commonly classified into unipolar and bipolar, depending on whether only one or both types of charge carriers participate in the current flow. As a consequence of the difference in operating principles, their properties—and therefore their suitability for specific applications—differ greatly. Bipolar transistors are well suited for high-speed applications and for driving large currents. Unipolar transistors are common in moderate-speed low-noise applications (JFETs) and are most prominent in digital circuitry (MOSFETs).

We use as an example the *n*-channel MOSFET (Metal-Oxide-Semiconductor Field Effect Transistor). Figure 5.13 shows a cross section along the channel. Two n^+p diodes are connected by a MOS structure. Applying a high enough positive potential on the gate an inversion (electron) layer will connect source and drain and for non-zero drain-source voltage an electron current will flow from source to drain. The strength of this current can be controlled by the gate potential and also

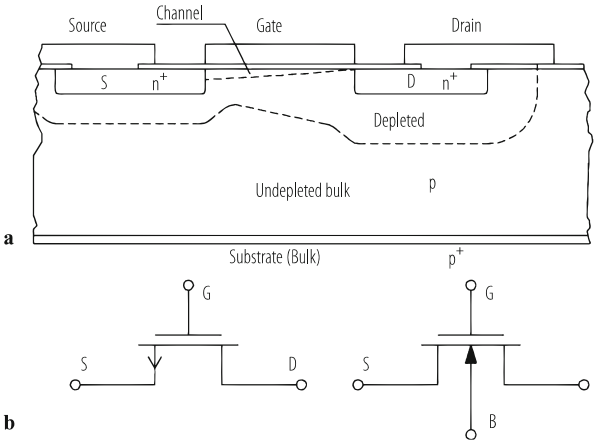


Fig. 5.13 *n*-channel MOSFET: Cross-section (**a**) and device symbol (**b**). The separation of the space-charge region from the channel below the gate and from the undepleted bulk is indicated by the dashed lines

by the drain voltage. A resistive voltage drop along the channel is responsible for the current saturation that occurs once this voltage drop equals the effective gate voltage (voltage above the threshold necessary to create inversion).

Important parameters of the transistor to be used in noise considerations are the transistor (output) conductance $g = \mathrm{d}I_\mathrm{d}/\mathrm{d}V_\mathrm{d}$ and transconductance $g_\mathrm{m} = \mathrm{d}I_\mathrm{d}/\mathrm{d}V_\mathrm{g}$. These and other parameters can be modelled using the graded channel approximation which relies on the assumption that changes along the channel are much smaller than those occurring in the transverse direction. It allows deriving scaling laws for changes in geometry. However, for microelectronics with minimal feature size these are of limited validity. Instead, two- and three-dimensional numerical device simulations are needed.

Measurement precision is limited by noise. There are several noise mechanisms present. Considering a resistor with resistance R for example, the thermal motion of electrons will result in a statistical fluctuation of the charge distribution in the conductor, leading to a noise voltage density of $\mathrm{d}\langle v_\mathrm{n}^2\rangle/\mathrm{d}f = 4\ kT{\cdot}R$ between the terminals of the resistor. The resistance of the MOSFET channel is a source of white noise too. It is customary to represent this noise by a voltage at the gate $\mathrm{d}\langle v_\mathrm{n}^2\rangle/\mathrm{d}f = 4\ kT(2/3)(1/g_\mathrm{m})$ for the operation of the transistor in the saturation region.

A further mechanism of noise is the capture and delayed release of single charge carriers in the channel. While being captured the drain current decreases, returning to the initial value when released. For a single trapping centre with characteristic average capture and release times a Lorentzian noise spectrum as function of frequency results. Having many different trapping centres, as is the case for traps at the $Si\text{-}SiO_2$ interface where trapping and detrapping occurs by means of tunnelling, the result of the superposition of Lorentzian noise spectra is a $1/f$ spectrum $\mathrm{d}v_\mathrm{n}^2/\mathrm{d}f = A_f\ /f$ with A_f a constant, which depends on the technology and the geometric parameters of the transistor. A_f is usually obtained from measurements and parameterized as $A_f = K_\mathrm{F}\ /(WLC_\mathrm{ox}{}^2)$. K_F characterizes the technology, W and L are channel width and length, and C_ox the oxide capacitance per unit area. Note that the *1/f* noise is independent of the transistor current.

5.7.2 The Measurement of Charge

The standard problem in the readout of a semiconductor detector is the low-noise measurement of the signal charge, usually under severe constraints such as high-speed operation, low power consumption, restricted space and frequently high radiation levels. In this section the general problems of charge measurement will be addressed, while specific solutions for the electronics will be considered later.

The *charge-sensitive amplifier* (CSA), invented by Emilio Gatti [14] and represented in Fig. 5.14, consists of an inverting amplifying circuit which—in the ideal case—delivers an output voltage proportional to the input ($U_\mathrm{out} = -A\ U_\mathrm{in}$) and a feedback capacitor C_f. In addition, a high-resistance feedback or a switch is needed in the feedback loop, in order to bring the circuit into its operating condition. C_D

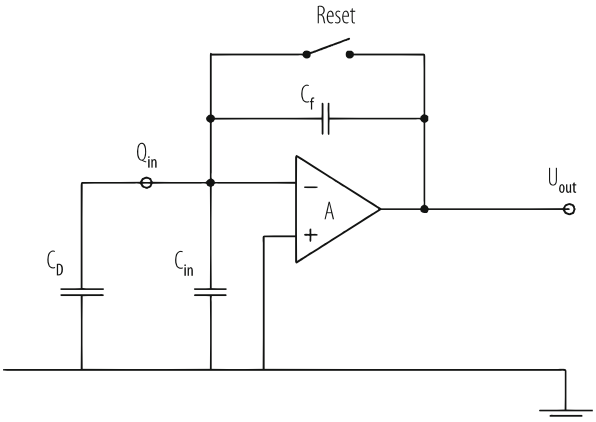


Fig. 5.14 Principle of a Charge Sensitive Amplifier (CSA). The inverting amplifier has gain A and a capacitive feedback. The reset switch is only used for bringing the system into its operating condition, and is often replaced by a high-ohmic resistor

represents the capacitive load of the detector at the input, C_{in} the capacitive load to ground in the amplifier, which is usually dominated by the gate capacitance of the input transistor.

Putting a charge Q_{in} at the input will result in an output voltage change of $U_{\mathrm{out}} = -Q_{\mathrm{in}}/(C_{\mathrm{f}} + (C_{\mathrm{D}} + C_{\mathrm{in}} + C_{\mathrm{f}})/A)$ which for large amplification is given by the ratio of signal charge over feed-back capacitance, indicating that the charge has been transferred completely from the detector to the feedback capacitor. For low frequencies the input impedance of the CSA will be represented by a capacitance of the value $C_{\mathrm{eff}} = (A+1)\ C_{\mathrm{f}} + C_{\mathrm{in}}$. A high value of $C_{\mathrm{eff}} > C_{\mathrm{D}}$, i.e. a low input impedance, is important because when C_{eff} is only of the same order of magnitude as the detector capacitance C_{D} the charge is incompletely transferred to the electronics. This results in a loss of sensitivity and possibly crosstalk within the detector to neighbouring channels.

Turning now to the question of measurement precision, respectively noise in the detector-amplifier system, we remark that it is customary to represent the effect of all amplifier noise sources by a single noise voltage U_{n} placed at the input (Fig. 5.15). As this noise voltage generator is in series with detector and amplifier it is called serial noise. The presence of the serial noise voltage U_{n} will result in an output voltage even if there is no signal charge present. For an evaluation of the serial noise charge, it is easiest to consider the charge necessary to compensate for the effect of the noise voltage, such that the output voltage remains at zero. The value can be immediately read from Fig. 5.15: $Q_{\mathrm{n}} = U_{\mathrm{n}}\ (C_{\mathrm{D}} + C_{\mathrm{in}} + C_{\mathrm{f}}) = C_{\mathrm{T}} U_{\mathrm{n}}$ with C_{T} the total "cold" input capacitance.

Notice that the serial noise is generated in the amplifier, the influence of the detector is due to the capacitive load at the amplifier input only. The detector itself produces noise due to statistical fluctuations of its leakage current I. This parallel

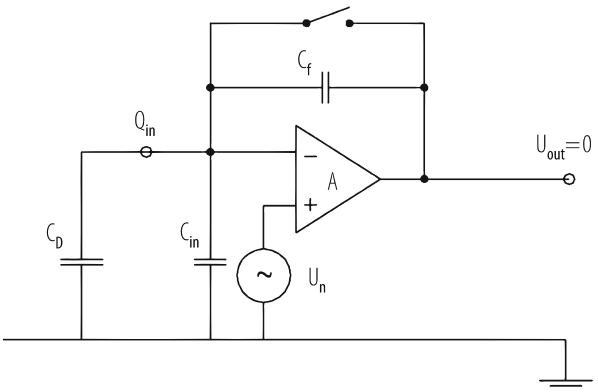


Fig. 5.15 The effect of amplifier serial noise in a detector-amplifier system

noise is represented by a noise current source of density $d<i_n^2>/df = 2I \cdot q$ in parallel to the detector capacitance C_D. To estimate the charge measuring precision one has to follow separately signal and noise through the complete readout chain and compare their respective output signals.

The signal produced by the amplifier will usually not be used directly; it will be further amplified and shaped, in order to optimize the ratio of signal to noise and to reduce the interference between subsequent signals. We will only consider a few very simple cases, the simplest being an idealized charge-sensitive amplifier followed by an RCCR filter. For a more elaborate treatment, the reader is referred to the literature (e.g. [15]).

The arrangement of a CSA followed by an RCCR filter is shown in Fig. 5.16. The output of the CSA is a voltage step given for very high amplification as Q/C_f. The shaper does an RC integration followed by a CR differentiation. This procedure results in a signal peak, which for the same integration and differentiation time constant $\tau = R_1C_1 = R_2C_2$ has the shape $U_{out}(t) = (Q/C_f) \cdot (t/\tau) \cdot \exp(-t/\tau)$ with a peak value $U_{peak} = (Q/C_f) \cdot \exp(-1)$. The height of this peak is a measure of the signal charge. Superimposed on the signal is the noise voltage, and we are interested in the signal-to-noise ratio, which is defined as the ratio of the height of the peak value to the root-mean-square value of the noise voltage measured at the same point in the circuit.

In order to find the noise voltage at the output, each noise source in the circuit has to be traced to the output and the resulting voltages added in quadrature. Doing so, one finds the important result that, for white (thermal) serial noise, the ratio of noise to signal (N/S) decreases with the square root of the shaping time constant τ, while for *1/f* noise this ratio remains constant. Parallel noise, given as a time integral over current fluctuations, increases with the square root of the shaping time.

More sophisticated continuous time filtering methods use (for example) Gaussian shape filtering, which can be approximated by several sequential RC integration and differentiation steps. Especially important in integrated electronics are the

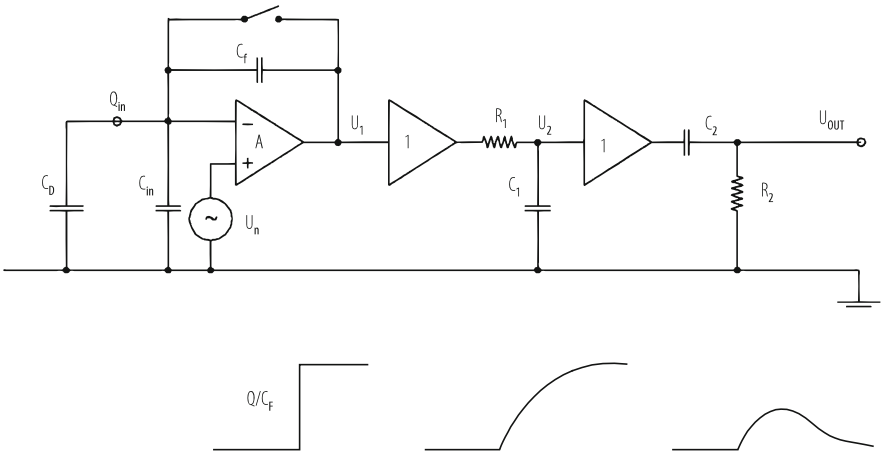


Fig. 5.16 Noise filtering and signal shaping in an RCCR filter following a charge-sensitive amplifier (top). The two unity gain amplifiers have been introduced in order to completely decouple the functions of the CSA, the integration (RC) and the differentiation (CR) stages. The signal form is indicated for each stage (bottom)

techniques in which the output signal is sampled several times and mathematical manipulations of the samples are performed. This can be done either after the measurement by numerical processing or directly by the local readout electronics. In the latter case, it is usually achieved by using switched capacitor techniques for analogue algebraic manipulations. Common to both methods, however, is the need to sample the signal at fixed (or, at least, known) times with respect to its generation. Alternatively with frequent enough sampling, the arrival time of the signal can be extracted from the data and filtering can be done afterwards by selecting the relevant samples before and after arrival of the signal. In all cases, however, the fact that the three noise components (white serial, 1/*f* and white parallel noise) scale with the available readout time in the described way remains valid.

As a further example we discuss double correlating sampling realized in switched capacitor technology which is most naturally realizable in integrated circuit technology. It is applicable if the signal arrival time is known in advance, as is the case for example in collider physics experiments.

The circuit (Fig. 5.17) consists of two sequential charge-sensitive amplifiers connected by a coupling capacitor C_s and switch S_c. Initially all switches are closed. Thus both CSAs have reset their input and output voltages to proper working conditions and a possible offset voltage between CSA1 and CSA2 is stored on capacitor C_s. The following operations are performed in sequence: (1) opening switch S_1 at time t_1, resulting in an unwanted charge injection into the input of CSA1 and therefore an output voltage change that will be stored on capacitance C_s and thus made invisible to the input of CSA2; (2) opening of reset switch S_2. Any voltage change on the output of CSA1 (e.g. signal or noise) is also seen in the output of CSA2, amplified by the ratio C_s/C_{f2}; (3) signal charge Q_s generation at time t_3

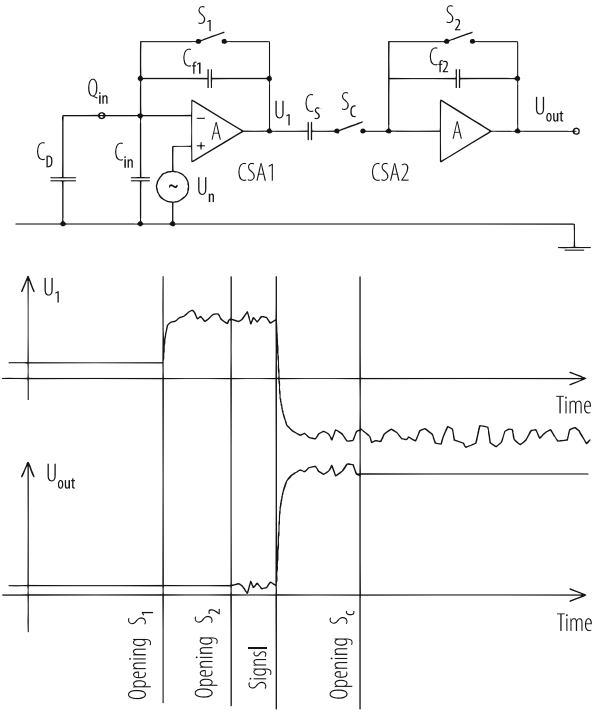


Fig. 5.17 Double-correlated sampling of the output of a charge-sensitive amplifier (CSA1) with the help of a coupling capacitor C_s and a second CSA2

changes the output of CSA1 by $\Delta U_1 = Q_s/C_{f1}$ and the output voltage by $\Delta U_{out} = Q_s\ (C_s\ /(C_{f1}C_{f2}))$; and (4) opening of switch S_c at time t_4 inhibits further change of the output voltage. The difference of the output voltage of CSA1 (amplified by C_{f2}/C_s) between times t_2 and t_4 remains present at the output of the circuit.

Double correlated sampling suppresses the reset noise due to operating switch S_1 and also suppresses low frequency noise but enhances the noise at higher frequencies. As a result white noise is not suppressed. This has to be done by limiting the frequency range of the amplifier. More sophisticated schemes of switched capacitor filtering, taking several samples (sometimes with different weights), have also been implemented.

5.7.3 Integrated Circuits for Strip Detectors

The development of integrated detector readout electronics was initiated by the simultaneous requirements of high density, low power and low noise for use with silicon strip detectors in the tight space environment of elementary particle physics

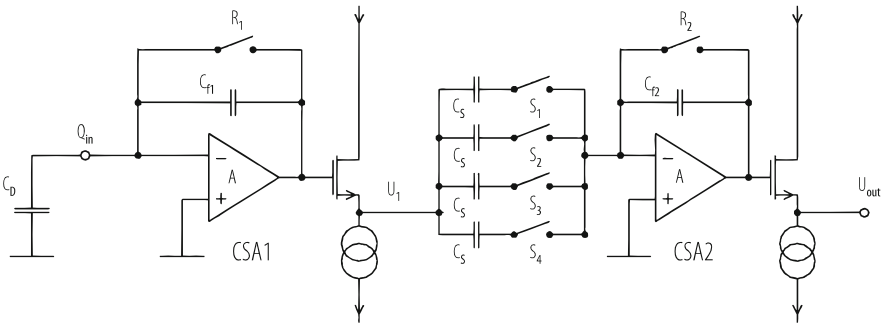


Fig. 5.18 Single channel readout schematics of the CAMEX64 strip-detector readout circuit

collider experiments. A variety of circuits has been developed for this purpose, the basic principle of essentially all of them being: (1) parallel amplification using a charge-sensitive amplifier at each input; (2) parallel signal filtering combined with second-stage amplification and parallel storage within capacitive hold circuits; and (3) serial readout through one single output channel.

We will present only one of the developments [16]. This was not only one of the first to be started but is still in use and has been further developed for many important applications.

The basic functional principle of a single channel is shown in Fig. 5.18. It consists of two charge-sensitive amplifiers, each followed by a source follower, and four sets of capacitors and switches that connect the output of the first amplifier with the input of the second amplifier. The circuit is rather similar to the one shown in Fig. 5.17, but the essential difference is the fourfold multiplication of the capacitive coupling between the amplifiers. In this way it is possible to perform fourfold double-correlated sampling at times that are shifted relative to each other. This procedure provides a good approximation to trapezoidal shaping, which means averaging the output over time intervals before and after signal arrival and taking the difference between the averaged samples.

The switching sequence that performs this function is the following: (1) Close R_1 and R_2. The charges on the feedback capacitances C_{f1} and C_{f2} are cleared. (2) Open R_1: some (unwanted) charge will be injected into the input by the switching procedure, producing an offset in U_1. (3) Close and open in sequence S_1 to S_4. The U_1 offset values at the four times t_1 to t_4 will be stored on the four capacitors C_s. (4) Open switch R_2. A small offset voltage appears at the output. (5) Deposit signal charge Q_{sig} at input. U_1 changes by an amount of $\Delta U_1 = Q_{sig}/C_{f1}$. (6) Close and open S_1 to S_4 in sequence at times t_1 to t_4. A charge $C_s \Delta U_1$ is inserted into the second amplifier at each sample. The total output voltage is $4C_s \Delta U_1/C_{f2} = Q_{sig} 4C_s/(C_{f1} C_{f2})$.

The complete chip, containing 64 channels, also comprises additional electronics, as shown in Fig. 5.19. Three test inputs allow injection of a defined charge through test capacitors. Digital steering signals are regenerated by comparators. The

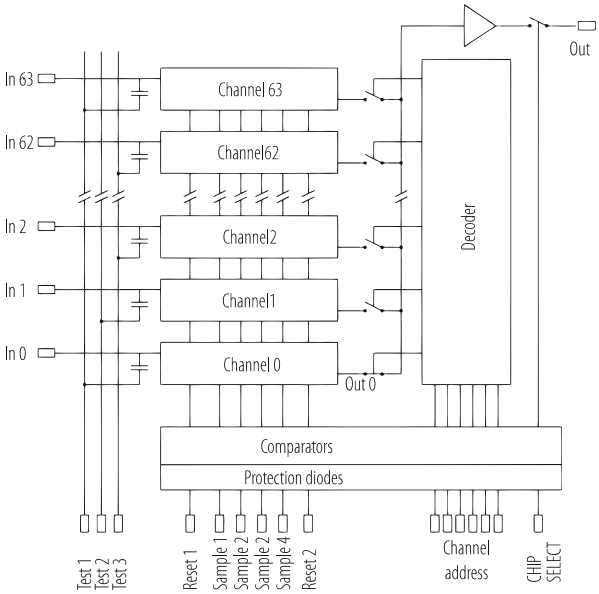


Fig. 5.19 Block diagram of the CAMEX64 strip-detector readout chip

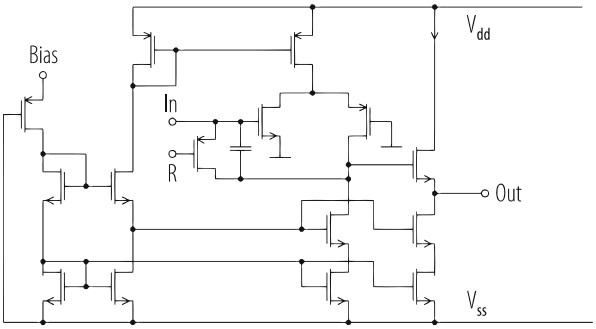


Fig. 5.20 Circuit diagram of the amplifier, including source follower and biasing circuit, of the CAMEX64 strip-detector readout chip

decoder switches one signal at a time on the single output line where a driving circuit for the external load is attached.

A circuit diagram valid for all charge sensitive amplifiers used is shown in Fig. 5.20. The current in all transistors can be scaled by a reference bias current (Bias). The input (In) can be shorted to the output with the reset switch (R) that lies in parallel to the feedback capacitor. The CSA output is connected to a source follower driving the output node (Out).

The circuits mentioned so far have been designed for moderate speed of applications in low-radiation environments. For the CERN Large Hadron Collider

(LHC), where the time difference between consecutive crossings of particle bunches (25 ns) is much shorter than the time it takes to decide whether or not the data of a particular event needs to be kept (approximately 2 μs), chips with high-speed operation and radiation hardness have been developed successfully. In addition to fast low-noise amplifiers and radiation hardness, it is required to store the information for approximately hundred bunch crossings.

The task of designing radiation hard electronics has been considerably eased by the industrial development of submicron integrated circuit technology which, due to the use of ultra-thin oxide, to a large extend has eliminated the problem of radiation induced threshold shifts in MOS transistors [17]. Taking some precautions in the design these technologies can be considered "intrinsically radiation hard".

5.8 Silicon Drift Detectors

The semiconductor drift detector was invented by E. Gatti and P. Rehak [1]. First satisfactorily working devices in silicon were realised in a collaborative effort by J. Kemmer at the Technical University Munich, the Max Planck Institute for Physics in Munich and the inventors [18].

The working principle may be explained by starting from the diode (Figs. 5.1 and 5.21a) if one realizes that the ohmic n^+ contact does not have to extend over the full area of one wafer side but can instead be placed anywhere on the undepleted conducting bulk (Fig. 5.21b). Then there is space to put diodes on both sides of the wafer (Fig. 5.21c). At small voltages applied to the n^+ electrode, there are two space-charge regions separated by the conducting undepleted bulk region (hatched in Fig. 5.21). At sufficiently high voltages (Fig. 5.21d) the two space-charge regions will touch each other and the conductive bulk region will retract towards the vicinity of the n^+ electrode. Thus it is possible to obtain a potential valley for electrons in which thermally or otherwise generated electrons assemble and move by diffusion only, until they eventually reach the n^+ electrode (anode), while holes are drifting rapidly in the electric field towards the p^+ electrodes.

Based on this double-diode structure the concept of the drift detector is realised by adding an additional electric field component parallel to the surface of the wafer in order to provide for a drift of electrons in the valley towards the anode. This can be accomplished by dividing the diodes into strips and applying a graded potential to these strips on both sides of the wafer (Fig. 5.5).

Other drift field configurations (e.g. radial drift) can be obtained by suitable shapes of the electrodes. Drift chambers may be used for position and/or energy measurement of ionizing radiation. In the first case the position is determined from the drift time. Furthermore, segmenting the n^+-strip anode in Fig. 5.5 into pads, a two-dimensional position measurement is achieved.

Due to the small capacitive load of the readout electrode to the readout amplifier, drift detectors are well suited for high precision energy measurement.

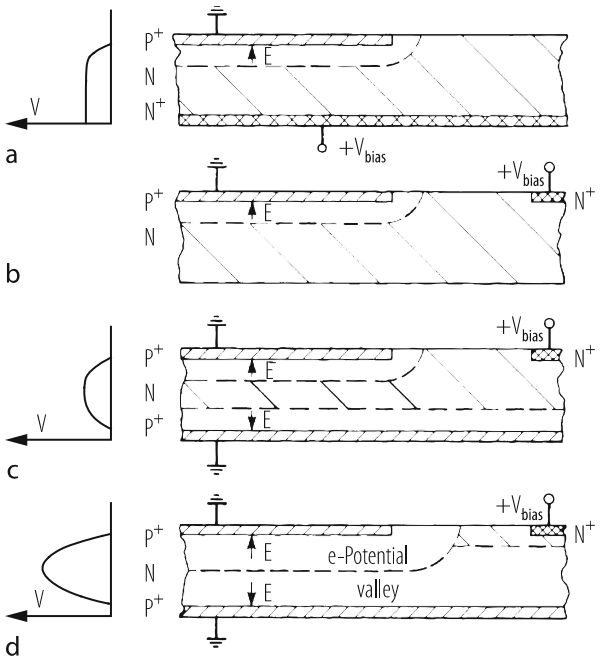


Fig. 5.21 Basic structures leading towards the drift detector: diode partially depleted (**a**); diode with depletion from the side (**b**); double diode partially depleted (**c**); double diode completely depleted (**d**)

5.8.1 Linear Drift Devices

Although linear devices seem to be the most straightforward realisation of the drift detector principle, one encounters some nontrivial problems. They are due to the finite length of the biasing strips and the increasing potential to be applied to these strips, which leads to a very large voltage of several hundred or (for very large drift length) a few thousand volts. Therefore guard structures have to be implemented which provide a controlled transition from the high voltage to the non-depleted region at the edges of the device.

A schematic drawing of the first operational silicon drift detector [18] is shown in Fig. 5.22. Anodes placed on the left and right side of the drift region collect the signal electrons generated by the ionizing radiation. The most negative potential is applied to the field-shaping electrode in the centre. Electrons created to the left (right) of this electrode will drift to the left (right) anode. The p^+-doped field electrodes do not simply end on the side, but some of them are connected to the symmetrical strip on the other half of the detector. In this way one insures that the high negative potential of the field strips drops in a controlled manner towards the potential of the undepleted bulk on the rim of the detector.

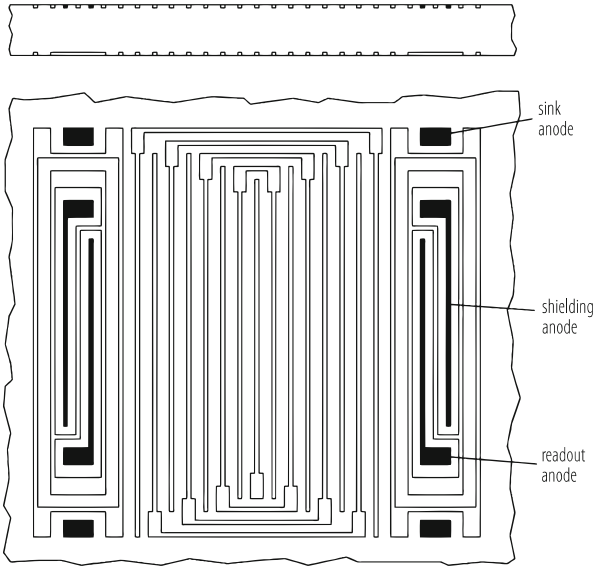


Fig. 5.22 Schematic cross-section and top view of a linear drift detector with *p*-doped field-shaping electrodes (light) and two *n*-doped (double) anodes (dark)

Looking closely at the anodes (Fig. 5.22), it can be seen that there are pairs of *n*-doped strips. Each pair is surrounded by a *p*-doped ring, which also functions as the field-shaping electrode closest to the anode. The two *n*-doped strips are separated by a *p*-doped strip that also connects to the ring surrounding the anode. Surrounding the *n*-strips completely by *p*-doped regions ensures that the adjacent *n*-doped anodes are electrically disconnected to each other and to the other regions of the detector (such as the non-depleted bulk). The outer *n*-strips are used to drain away electrons from the high voltage protection region, while the inner strips measure the signals created in the active detector region.

The opposite side of the silicon wafer is for the large part identically structured. Differences are only in the anode region, where the *n*-implantation is replaced by *p*-doped strips. In the main part of the detector, the strips on opposite sides of the wafers are kept at the same potential, thus assuring a symmetrical parabolic potential distribution across the wafer (Fig. 5.23a). Near the anode an increasing potential difference between the two wafer surfaces moves the potential valley for electrons to the front side until it ends at the anode (Fig. 5.23b).

The linear drift detectors described so far allow one dimensional position measurement only. Dividing the anode of a linear drift detector into pads (Fig. 5.24) leads to a two-dimensional position measurement. One coordinate is obtained from the drift time, the other from the pads on which the signals appear. The second coordinate may be further improved by interpolation using the signal in neighbouring pads. The signal will be distributed over more than one pad if the

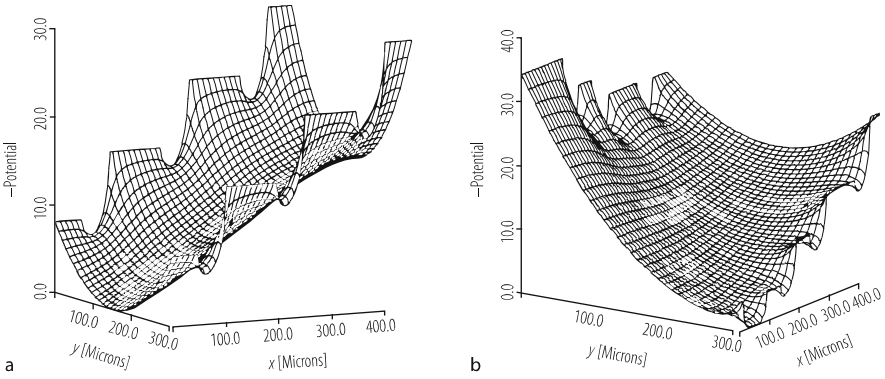


Fig. 5.23 Electron-potential distribution in the linear region (**a**) and close to the anode region where the potential valley is directed towards the surface (**b**)

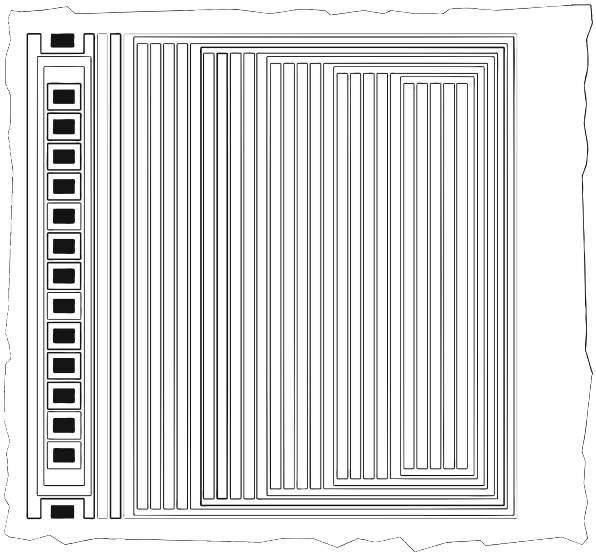


Fig. 5.24 Two-dimensional drift detector with the anode strip divided into pads. The dark pad anodes are embedded in a *p*-doped grid that provides insulation between neighbouring pads

diffusion during the drift time leads to a charge cloud at the anode that is comparable to the spacing of the pads.

For very long drift distances and/or low drift fields, the signal charge will be spread over more than two readout pads. This is an undesirable feature when measuring closely spaced signals. Lateral diffusion can be suppressed by creating deep strip-like p-implanted regions parallel to the nominal drift direction [19]. In this way deviations from the nominal drift direction due to non-uniform doping of the silicon are also avoided.

5.8.2 *Radial and Single Side Structured Drift Devices*

Radial drift devices are in some sense simpler to design than linear devices because the problem of proper termination of the field-shaping strips does not occur. Radial devices are especially interesting for energy measurement. A small point-like anode with extremely small capacitance may be placed into the centre of the device. The small capacitance results in low electronic noise and as a consequence very good energy resolution.

In one special case radial drift to the outside has been realized with a circular anode divided into pads, thus arriving at two-dimensional position measurement in cylindrical coordinates. An interesting feature of such an arrangement is the high position accuracy at small radius in the azimuthal direction. The position in this second coordinate is obtained from the charge distribution measured in the anode pads by projecting it back in the radial direction. A large-area device of this type [20], with a hole in the centre for the passage of the particle beam, has been produced for the CERES particle physics experiment at CERN. The device also uses a method to drain the current generated at the oxide-silicon interface between the field-shaping rings to an *n*-doped drain contact, separated from the signal-collecting anode [21]. In this manner the anode leakage current is reduced and the measurement precision increased.

The S*ilicon Drift Diode* (SDD) [3] combines radial drift with a homogeneous unstructured backside radiation entrance window (Fig. 5.25). Its principal field of application is in (X-ray) spectroscopy where excellent energy resolution is required. A further significant improvement was obtained by integrating a readout transistor into the device (Fig. 5.26). In contrast to the original drift chamber with the electron potential valley located parallel to the wafer surfaces now only one structured surface provides the drift field in the valley which now is at an angle with respect to the wafer surface.

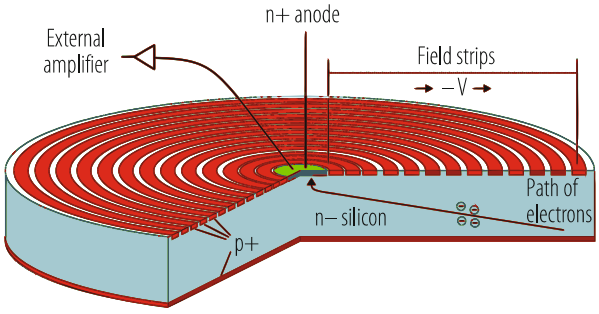


Fig. 5.25 Cylindrical silicon drift detector. The entire silicon wafer is sensitive to radiation. Electrons are guided by an electric field to the small collecting anode in the centre

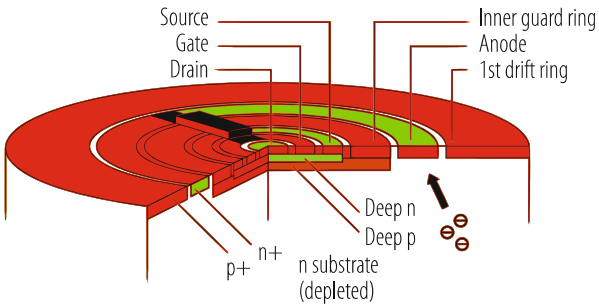


Fig. 5.26 On-chip single sided junction FET coupled to the readout node of a cylindrical silicon drift detector

Having only one surface structured allows using the unstructured surface of the fully depleted device as radiation entrance window. Not having to take other functions into considerations, this radiation entrance window can be made very thin and uniform [22]. The circular geometry with a very small charge-collecting anode in its centre reduces the capacitive load to the amplifier and therefore the noise.

Having the first transistor integrated into the device [23], the capacitance of the detector-amplifier system is minimized by eliminating bond wires between detector and amplifier. In this way stray capacitances between the readout node and ground are avoided, which makes the system faster and less noisy. Further advantages are evident as electrical pickup is significantly reduced and microphony i.e. noise introduced by mechanical vibrations, is excluded. In order to work on the lowly doped and fully depleted substrate, a non-standard "Single Sided Junction Field Effect Transistor" (SSJFET) has been developed [24].

Drift detectors with an integrated transistor are commercially available. They can also be obtained as modules assembled with a Peltier cooler in a gas-tight housing with a thin radiation entrance window (Fig. 5.27). To demonstrate the excellent spectroscopic performance achieved with such devices a spectrum obtained with an ^{55}Fe source and the quantum efficiency are presented in Fig. 5.28 for a cylindrical SDD with a sensitive area of 5 mm^2. The detector temperature, important for the leakage current, was set to −20 °C and the signal shaping time to 1 μs. The $Mn_{K\alpha}$ line at 5.9 keV and the $Mn_{K\beta}$ line at 6.5 keV are clearly separated and their widths are only slightly above the intrinsic Fano limit given by the pair generation process in silicon.

Cylindrical silicon drift diodes with integrated SSJFETs have been manufactured with sensitive areas in the range from 5 mm^2 to 1 cm^2.

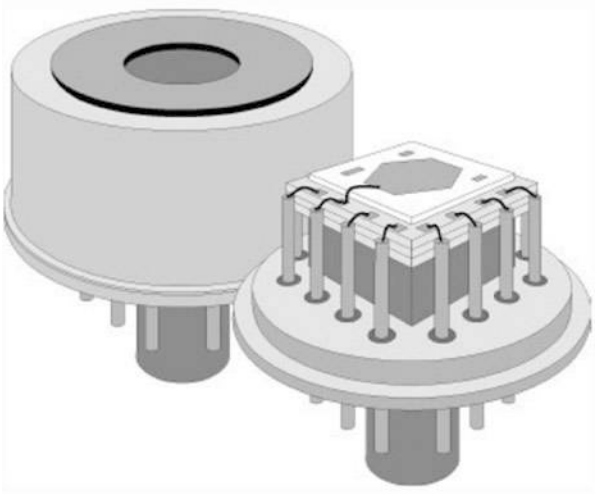


Fig. 5.27 Perspective view of a module consisting of a single-sided structured cylindrical drift detector with integrated SSJFET transistor, cooled by a Peltier element

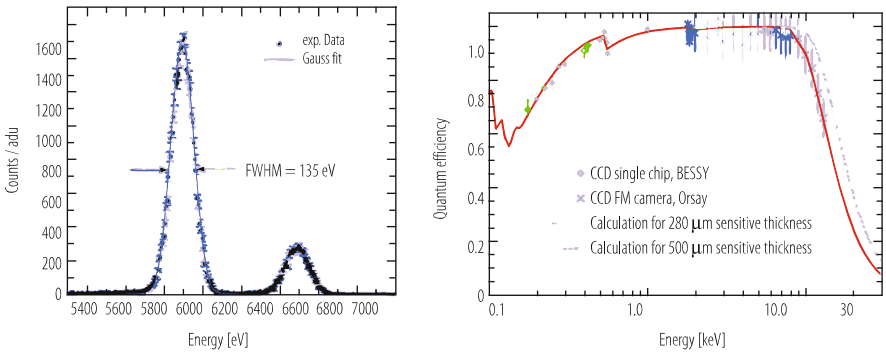


Fig. 5.28 $Mn_{K\alpha}$ – $Mn_{K\beta}$ spectrum (*left*) and quantum efficiency as function of X-ray energy (*right*) of a 5 mm^2 drift diode. The device was operated at −20 °C with a shaping time of 1 μs

5.9 Charge Coupled Devices

Charge coupled devices (CCDs) have for a long time been used as optical sensors, most noticeably as imaging devices in video cameras. Some years ago they also found their application as particle detectors in Particle Physics [25], where specially selected optical CCDs were used. Meanwhile detector systems have been constructed for measuring tracks in electron-positron collisions [26].

p-n CCDs for the special purpose of particle and X-ray detection have been developed [2]. They are based on the principle of side-wards depletion of a double-

diode structure, which is also used in the semiconductor drift chamber. Their first use was in two space-based X-ray telescopes: XMM [27] and ABRIXAS [28].

CCDs are non-equilibrium detectors. Signal charge is stored in potential pockets within a space-charge region, the content of which is then transferred to a collecting readout electrode. In order to retain the thermal non-equilibrium condition, thermally generated charge that also assembles in the potential pockets has to be removed from time to time. Usually this is done during the readout cycle of the device.

While in conventional MOS CCDs minority carriers (electrons in a *p*-type bulk) are collected, the *p*-*n* CCDs are majority carrier (electrons in an *n*-type bulk) devices. The conventional MOS CCDs to be described in the following for didactic purposes store and transfer the charge directly at the semiconductor-insulator interface. These devices are in practice not used anymore and have been replaced by buried-channel CCDs, in which the store-and-transfer region is moved a small distance away from the surface. As a result they are less sensitive to surface radiation damage. In *p*-*n* CCDs, this region is moved a considerable distance into the bulk.

5.9.1 MOS CCDs

The CCD transfer mechanism is explained in Fig. 5.29 that shows a cut along the transfer channel. The top part of the *p*-type bulk is depleted of charge carriers and the potential along the Si-SiO_2 interface is modulated in a periodic fashion with the help of the metal electrodes on top of the SiO_2. Electrons created in the sensitive bulk region assemble in the potential maxima (minima for electrons) at the Si-SiO_2 interface.

The charge can now be moved towards the readout electrode by a periodic change of the voltages ϕ_1, ϕ_2, and ϕ_3, as shown in the figure. First ϕ_2 is increased to the same level as ϕ_1 and the signal charge will spread between ϕ_1 and ϕ_2. If now ϕ_1 is lowered, the signal charge will transfer below the electrodes ϕ_2. If this procedure is followed for ϕ_2 and ϕ_3 and then again for ϕ_3 and ϕ_1, the signal charge is transferred by a complete cell. After several cycles the charge will finally arrive at the anode, where it can be measured.

Placing many of these channels next to each other and separating them by so called channel stops one arrives at a matrix CCD. Channel stops prevent the spreading of signal charge to neighbour channels. They can be realized by doping variations as for example an increased *p*-doping between channels. Usually charge is transferred into one additional charge transfer channel oriented perpendicular to the matrix channel (Fig. 5.30) so that the pixel charge can be shifted towards a single output node.

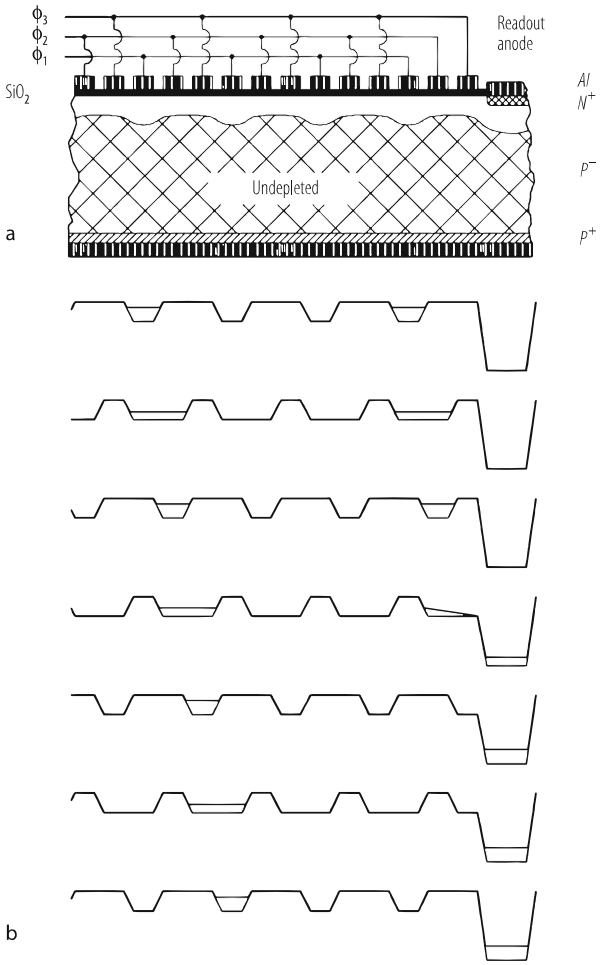


Fig. 5.29 Working principle of a three-phase MOS CCD: layout (**a**); charge-transfer (**b**): Every third gate electrode is connected to the same potential (ϕ_1, ϕ_2, ϕ_3) so that a periodic potential appears below the gates at the Si-SiO_2 interface. Electrons are collected in the maxima of the potential distribution. They can be shifted towards the readout anode by changing the potentials, as shown in (**b**)

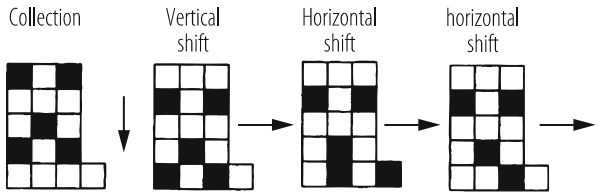


Fig. 5.30 Matrix CCD and the principle of the charge-transfer sequence. Charge is shifted in the vertical direction with all pixels of the matrix in parallel, the lowest row being transferred into a horizontal linear CCD. This horizontal CCD is then read out through a single output node

5.9.2 *Fully Depleted pn-CCDs*

pn-CCDs were originally developed for X-ray imaging in space. A 6×6 cm^2 size device is used as focal imager in one of the three X-ray mirror telescopes at the European XMM/Newton X-ray observatory [29]. From 2000 until the end of the mission in 2018 it has produced high quality X-ray images of the sky [30].

The *pn*-CCD principle, derived from the silicon drift chamber, has already been shown in Fig. 5.5. The layout of the XMM focal plane detector is shown in Fig. 5.31. Twelve 1×3 cm^2 CCDs with 150×150 μm^2 pixel size are monolithically integrated into a single device placed on a 4 inch silicon wafer of 300 μm thickness. Each column of pixels has its own readout channel allowing for fast parallel readout.

Figure 5.32 shows a cross section of a *pn*-CCD along the transfer channel. Here one sees in greater detail the functioning of the device. Contrary to standard MOS-CCDs the registers are formed as *pn*-junctions and the radiation sensitive oxide plays only a minor role. The device is fully depleted with a higher *n*-type doping concentration in the epitaxial layer below the top surface. This leads to a potential distribution shown in the right part of the figure and prevents holes from the p^+-doped registers to be emitted across the wafer towards the backside *p*-doped entrance window. Charge storage and transfer occurs in a depth of approximately 10 μm in contrast to MOS CCDs where this happens at the Si-SiO_2 interface. Fast and efficient charge transfer by drift is therefore possible even for large pixel sizes.

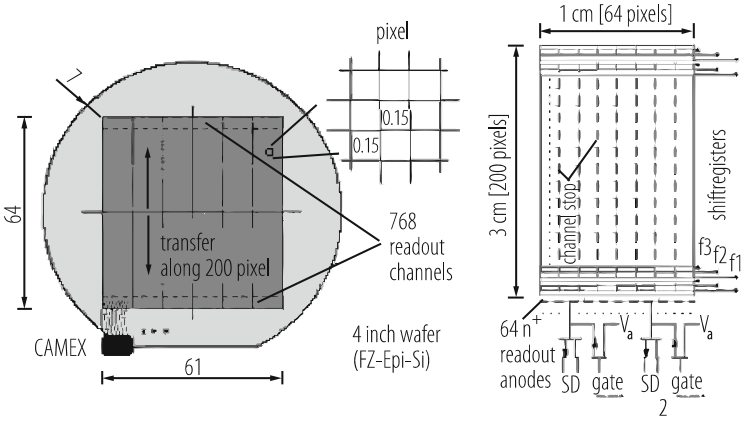


Fig. 5.31 Layout of the XMM *pn*-CCD. 12 logically separate *pn*-CCDs of 1×3 cm^2 area are monolithically fabricated on a 4 in. wafer to a 6×6 cm^2 device with a common backside entrance window. The pixel size is 150×150 μm^2

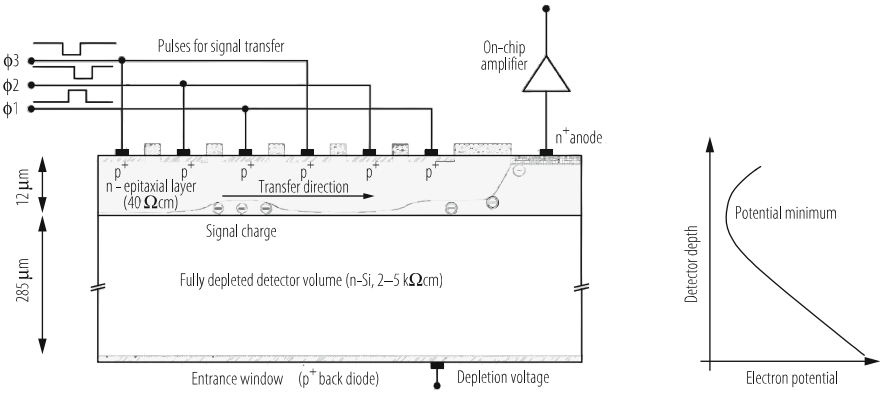


Fig. 5.32 Cross section through the CCD along the transfer channel

5.9.3 CCD Applications

MOS CCDs have a long history in optical imaging. They have been used in camcorders but also in optical astronomy. In particle physics they were first used by the ACCMOR collaboration in the NA11 experiment at CERN where they were successfully employed for heavy flavour decay detection and measurement. They then found their way to collider physics at SLAC and also to X-ray astronomy, where thinning for backside illumination was necessary to achieve sensitivity for low energy X-rays.

Thinning reduces the sensitive volume and therefore the sensitivity at higher X-ray energies. This disadvantage is avoided with *pn*-CCDs that have a typical thickness of 500 µm and, in addition are built with a ultra-thin entrance window so that high quantum efficiency at both low (100 eV) and high (20 keV) X-ray energies is reached. Good radiation tolerance for X-rays is due to two reasons, the absence of sensitive MOS registers and the absorption of X-rays within the bulk before they reach the sensitive charge transfer region (self-shielding). At XMM/Newton *pn*-CCDs have been operating in space for 18 years without noticeable performance degradation.

Compared to MOS CCDs the readout speed is significantly increased due to the larger pixel size, the higher charge transfer speed and parallel column readout. Very large pixel sizes cannot be realized in MOS CCDs that transfer charges very close to the Si-SiO_2 interface.

Use in a further X-ray mission is in preparation: eROSITA (extended ROentgen Survey with an Imaging Telescope Array). Here the CCD is split into an image collecting area and a frame store area. After collection, the complete image is transferred very fast into the frame store area from where it is read with moderate speed row by row while at the same time the next image is collected. The typical image frame readout takes 1 ms, while for MOS CCDs it is in the range of 1 s.

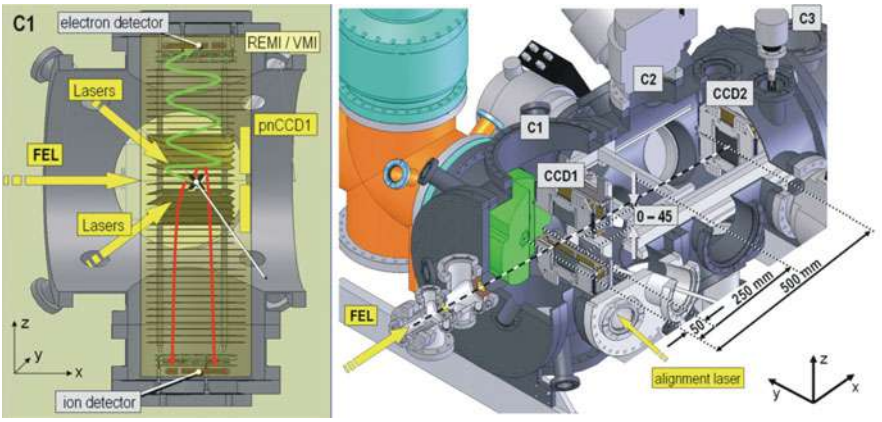


Fig. 5.33 Schematic section through the CAMP detector. The reaction electron and ion detectors with the first CCD sensor plane are depicted on the left hand side. The *pn*-CCD detectors shown in perspective view on the right can detect all photons emerging from the target. In addition, the design allows feeding in other lasers for alignment or pump-probe purposes, as well as for mounting other high-resolution, small-solid-angle electron TOF or crystal spectrometers. The pnCCD1 can be moved in all three directions with a maximum distance of 25 cm along the beam trajectory

Although *pn*-CCDs have been developed for X-ray astronomy they are also visible-light detectors. One application is in adaptive optics that corrects in real time mirror geometries of optical telescopes in order to compensate for atmospheric turbulences at frequencies of approximately 1 kHz.

pn-CCDs are also used in experiments at accelerator-based light sources in particular at X-ray Free Electron Lasers (e.g. FLASH and the European XFEL at Hamburg and LCLS at SLAC). The Center of Free Electron Science (CFEL) in Hamburg has designed the CFEL-ASG Multi Purpose (CAMP) chamber (Fig. 5.33) [31], which combines electron and ion momentum imaging spectrometers with large area, broadband (50 eV to 25 keV), high dynamic range, single photon counting and imaging X-ray detectors based on *pn*-CCDs. The excellent low energy response of *pn*-CCDs has been demonstrated by measuring the response to 90 eV photons at FLASH (Fig. 5.34).

5.10 Active Pixel Detectors

The CCDs discussed in the previous chapter collect charges in pixels during their charge collection period and transport them during the transfer period pixel by pixel to a readout node. Charges produced during the transfer cycle will also be read but the assigned position will be wrong. In active pixel detectors each pixel has its own readout channel and the charge will be assigned to the pixel where it was generated. There are four types of active pixel detectors:

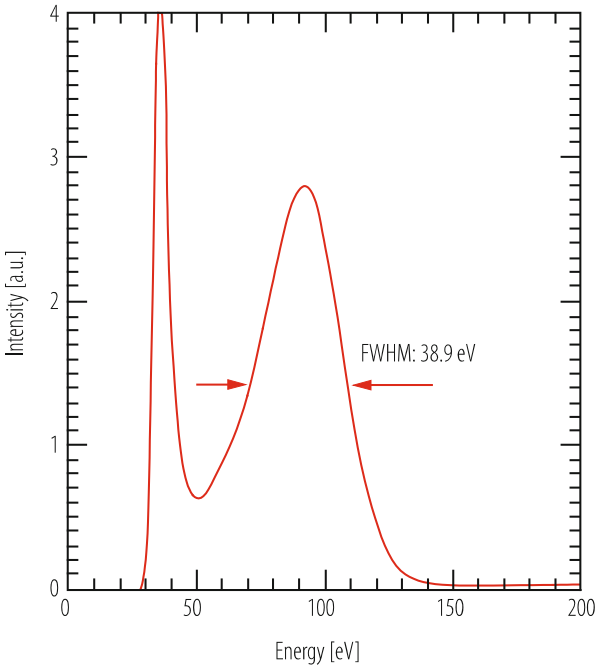


Fig. 5.34 Energy resolution measured at FLASH with 90 eV photons. Every photon generates approximately 25 electron-hole pairs, which are detected with a read-out noise of 2.5 electrons (rms). The measured FWHM energy resolution is only 38.9 eV

(a) Hybrid pixel detectors are diode arrays bonded to an electronics chip produced on a separate wafer so that each pixel has its own readout channel.
(b) MAPS (Monolithic Active Pixel Sensors) are pixel arrays with readout for every pixel directly integrated on the same chip.
(c) DEPFET pixel detectors are two dimensional arrays of DEPFETs with parallel charge collection in the DEPFETs and serial delayed readout of the charges stored in the internal gates.
(d) DEPFET Macro Pixel detectors, pixel detectors with large cell size combine DEPFETs with drift detectors.

All these detector types exist in many variations. Hybrid pixel detectors and MAPS allow parallel data processing and can perform complex tasks thanks to the miniaturized VLSI electronics. This however has a price in power consumption. DEPFET pixel detectors so far are built in a technology of moderately large feature size. Thus complex data processing is not foreseen. Its advantages are sensitivity over the whole bulk, high energy resolution and very low power consumption.

5.10.1 Hybrid Pixel Detectors

Hybrid pixel detectors are used at the Large Hadron Collider (LHC) as the tracking detectors closest to the beam, where the track densities is highest and the radiation exposure most severe. They also became a standard detector for X-ray imaging, in particular at accelerator driven X-ray sources. In their simplest form they consist of a detector wafer with a two dimensional diode array and separate electronics wafers as shown in Fig. 5.35. Every diode is individually connected by bump bonding to its own readout channel. Other connection techniques, including capacitive coupling, have been demonstrated. As readout and sensor are separate, the sensor material can be freely chosen, e.g. a high-Z sensor for the detection of high-energy X-rays.

The main challenge in such a device lies in the electronics that has to provide several functions as for example low noise charge readout and high dynamic range, and—depending on the application—data storage, zero suppression and transmission to the external electronics in analogue or digital form. These functions have to be implemented on an area of the pixel size. Frequently very high speed operation at low power is required as is the case for example in the LHC at CERN. Reaching these goals has been possible by profiting from the dramatic industrial progress in submicron electronics and adapting it to the specific needs. The use of submicron electronics that uses very thin gate oxides has also alleviated the problems with respect to radiation damage.

The typical pixel dimension for the hybrid pixel sensors presently operating at the CERN LHC are of order $100 \times 100\ \mu m^2$. The modules of the ATLAS vertex

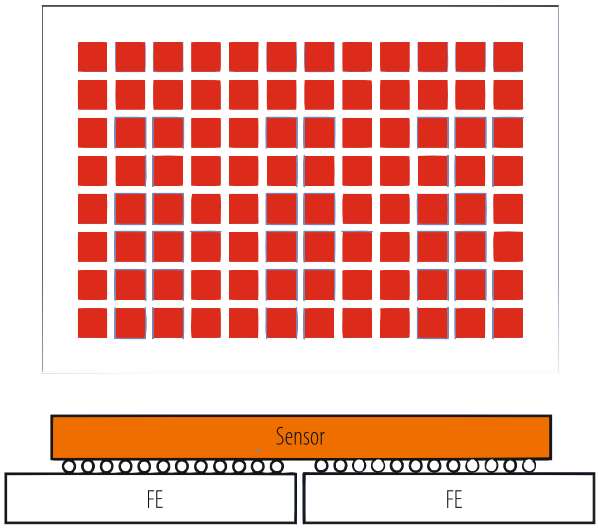


Fig. 5.35 Concept of a Hybrid Pixel Detector consisting of a diode array "flip chip" bonded to several readout chips

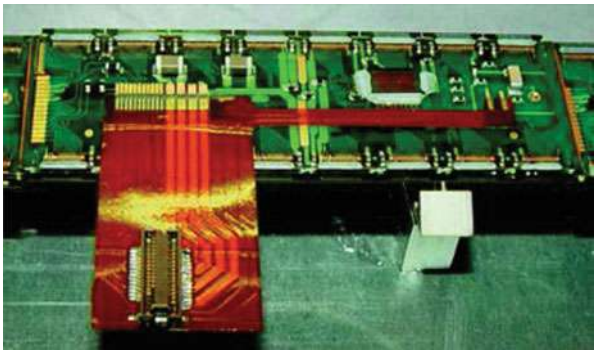


Fig. 5.36 Photo of an ATLAS pixel detector module

detector, shown in Fig. 5.36, have a pixel size of 50 μm × 250 (400) μm, the ones of CMS 100 μm × 150 μm For the High-Luminosity LHC hybrid pixel detectors with pixel sizes of 50 μm × 50 μm and 25 μm × 100 μm are under development.

The hybrid pixel detectors used for X-ray science face somewhat different challenges and follow different concepts. AGIPD (Adaptive Gain Integrating Pixel Detector) [32], which operates at the European XFEL at Hamburg, where X-rays are delivered in pulse-trains with 220 ns distance between pulses, is designed to detect single and up to 10^4 photons with energies in the range 5–15 keV per pulse in pixels of 200 μm × 200 μm, and store 350 frames to be read out in between the pulse trains. This is achieved by signal-driven switching into four gain ranges. In addition, the 500 μm thick pixel sensor is designed for a breakdown voltage above 900 V for ionizing doses up to 1 GGy. There are many applications in X-ray science, where the recording of individual frames is not required, but the number of hits above a given threshold or in a given energy interval are counted for every pixel or the integrated charge for a given time interval recorded. As the electronics takes significantly less space than required for recording and storing individual frames, pixel sizes as small as 55 μm × 55 μm have been achieved. Outstanding examples for such detectors are PILATUS [33] developed at PSI, and the MEDIPIX series [34], developed by a collaboration centred at CERN.

5.10.2 Monolithic Active Pixel Sensors (MAPS)

This name is used for pixel sensors produced with integrated circuit technology on a single wafer using part of the substrate as detector material. One advantage of MAPS is the significantly easier fabrication of detector modules resulting in a significant cost reduction; another is that MAPS can be produced in CMOS Fabs, which includes a fast turn-around time for the development. However, MAPS are very complex devices and achieving all the requirements of the experiments at high-luminosity, including their radiation performance remains a challenge.

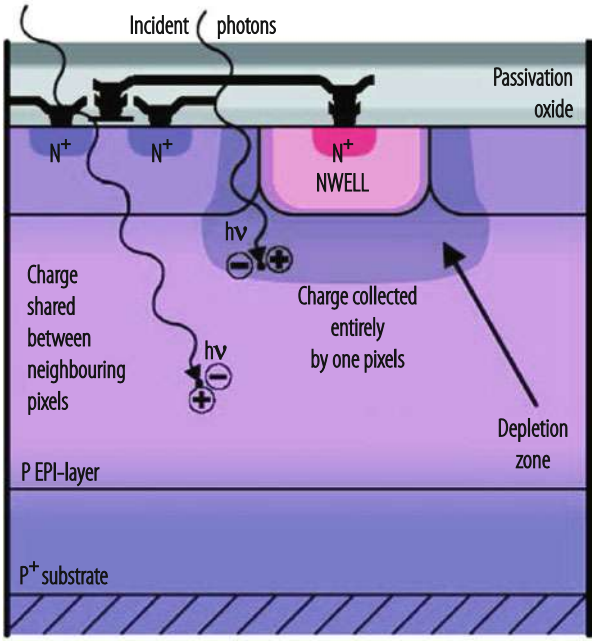


Fig. 5.37 Cross section through a pixel of a MAPS fabricated on CMOS technology but using only NMOS transistors

A first successful demonstration of MAPS operating in an experiment is the EUDET beam telescope [35], with MAPS using only *n*-channel transistors out of an original CMOS technology. Figure 5.37 shows the cross section through a MAPS pixel cell. The *n*-well is used as collecting electrode and all transistors are placed within the *p*-wells. A small volume next to the *n*-well is depleted of charge carriers. In this region signal electrons are collected by drift, but, the major part of the sensitive volume—the *p*-epitaxial layer—is field-free. Thus most of the charge is collected by diffusion, which is intrinsically slow and leads to a large spread of charge into neighbouring cells. There are good reasons why *p*-type transistors are avoided. They would have to be placed into an *n*-well. If this well were separated from the charge collecting electrode it—depending on the *n*-well potentials—would collect signal electrons in competition to the signal electrode or might even inject electrons into the bulk. If it were put into the same well as the collecting electrode it would induce charge directly into the input of the pixel.

For photon detection—as shown in the figure—in addition the material on the top as for example the conducting leads as well as the thick insensitive well zones will absorb part of the incident radiation.

The pixel circuitry (Fig. 5.38) is rather simple. It consists of an NMOS input transistor, a reset transistor and an output select switch. Signal charge is stored at the

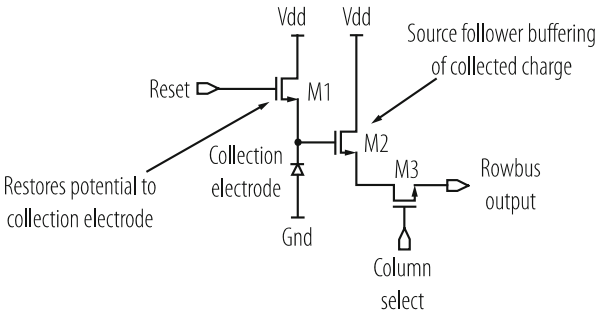


Fig. 5.38 Pixel circuitry of MAPS based on CMOS technology but using only three NMOS transistors. The collecting electrode is directly connected to the gate of a source follower (M2) whose load is common to all pixels of a column and activated by the column select switch. The input node is reset with the reset transistor M1

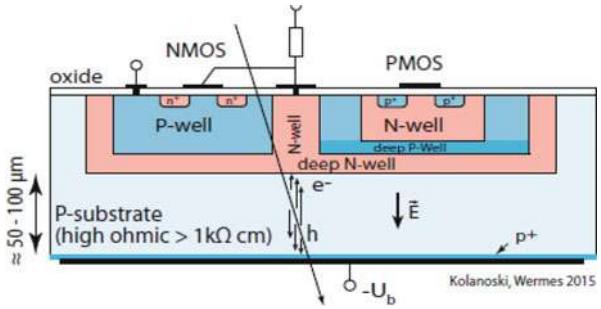


Fig. 5.39 DMAPS with large collection electrodes (figure from Wermes-Kolanoski)

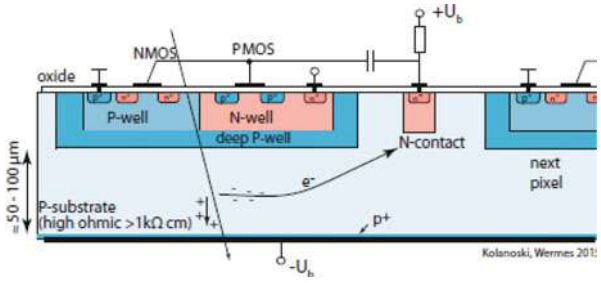


Fig. 5.40 DMAPS with small collection electrodes

input node, read out sequentially and cleared afterwards. MAPS using both CMOS types have also been developed [36].

To overcome the problem of slow charge collection by diffusion, which also makes the sensor sensitive to bulk radiation damage, DMAPS (Depleted CMOS Active Pixel Sensors), are being developed [37]. They are fabricated on substrates with resistivity between 100 Ω·cm and a few kΩ·cm and operated with depletion depths of typically 50–200 µm. As shown in Figs. 5.39 and 5.40, two approaches

are followed: Large Collection Electrode (a) and Small Collection Electrode (b). Design (a) has the advantage of a more uniform electric field resulting in shorter drift distances, and thus a good radiation tolerance is expected. Its disadvantage is the large capacitance of about 100 fF per pixel and an additional well-to-well capacitance of similar value, which results in increased noise, reduced speed, higher power consumption and possibly cross-talk between sensor and digital electronics. Design (b) has a small electrode adjacent to the well in which the electronics is embedded. This has the advantage of a small capacitance of about a few fF and thus improved noise and speed at low power. However, the electric field in the sensor is not uniform with low field regions. This makes them more sensitive to radiation damage. DMAPS of both types have been fabricated by different foundries in 150 nm, 180 nm and 350 nm technologies. They show impressive results even after irradiation with hadrons to fluences exceeding a few 10^{15} cm^{-2}.

5.10.3 DEPFET Active Pixel Sensors

The Depleted Field Effect Transistor structure shown in Fig. 5.6 is a natural building element for a pixel detector. It acts simultaneously as detector and as amplifier. A variety of DEPFET designs can be constructed. Figure 5.41 shows two examples, one with cylindrical, the other with linear geometry.

Arranging many of these devices in a matrix and connecting them in such a way that selected DEPFETs can be turned on, one arrives at a pixel detector with charge

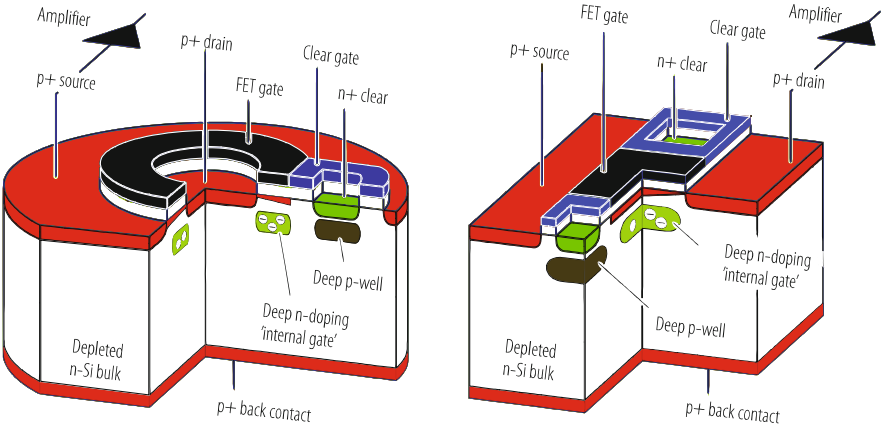


Fig. 5.41 Schematic drawings of MOS-type DEPFETs with circular (*left*) and linear (*right*) geometry. The signal charge is collected in a potential well ("internal gate") below the FET gate, thereby increasing the conductivity of and thus the current in the transistor channel. The collected charges can be drained towards the clear contact by applying voltage pulses to the clear contact and/or the clear gate

storing capability. Before turning to the matrix arrangement the main properties of the DEPFETs are summarised:

- Combined function of sensor and amplifier;
- Full sensitivity the over complete wafer, low capacitance and low noise, non-destructive repeated readout, complete clearing of the signal charge and thus no reset noise.
- Continuous (real time) and integrating (charge storage) operating modes can be chosen.

The signal can be read out either at the source as indicated in the left figure or at the drain as shown in the linear example. With source readout one compensates the increase of channel conduction due to the charge in the internal gate by a reduction of the external gate-source voltage, seen as voltage change of the source. In the drain readout the source potential is kept constant and the drain-current change can be directly observed. An important property in pixel detector applications is the fact that the signal charge collection occurs not only for current carrying DEPFETs but also for those which have been turned off with the help of the external FET gate.

DEPFET pixel sensors have been developed at the MPI Semiconductor Laboratory in Munich for several purposes, as focal sensors of the proposed European X-ray observatory XEUS [38] and as vertex detector for the BELLE-II experiment at KEK in Japan and the proposed International Linear Collider ILC. In XEUS the combined functions of imaging and spectroscopy are of importance, for the vertex detectors the measurement of position of charged tracks is of prime interest. This however has to be done with very high precision (few μm) and at high readout speed. The position measurement requirement in XEUS is not as stringent; it is matched to the expected quality of X-ray imaging. However, highest emphasis is given to spectroscopic quality and quantum efficiency and data readout speed is still large.

As a consequence of these and further requirements circular geometries have been chosen for XEUS and linear ones for the vertex detectors (see Fig. 5.41). The excellent spectroscopic capabilities of DEPFETs can be appreciated from the ^{55}Fe source spectrum taken with a single circular pixel cell (Fig. 5.42).

The DEPFET with its capability of creating, storing and amplifying signal charge is an ideal building block for a pixel detector. A large number of DEPFETs can be arranged in a matrix in such a way as to power selected DEPFETs for reading and clearing the collected signal charge. Figure 5.43 shows a rectangular arrangement of DEPFETs. Their drains are connected column wise while gates and clear electrodes are connected row wise. Each row has its individual readout channel. A row at a time is turned on with the help of the gate voltage while all other DEPFETs have zero current. Charge collection does not require a current within the DEPFET.

Readout can be performed in double correlated mode: Turning on the current with a negative voltage on the gate is followed by a first reading of the current, a clearing of the signal charge in the internal gate with a positive pulse at the clear contact and a second current reading before the current is turned down again and reading is switched to the next row. The difference of first and second current

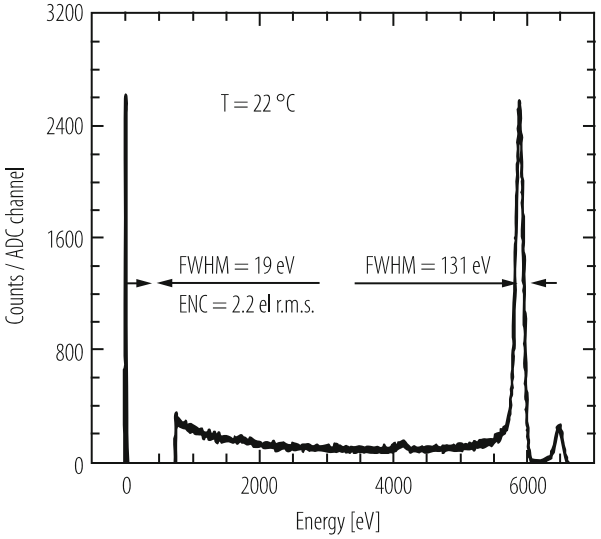


Fig. 5.42 ^{55}Fe spectrum measured with a single circular (XEUS-type) DEPFET. A spectral resolution of 131 eV has been obtained with room temperature operation and 6 μs Gaussian shaping. The separately measured noise peak has a FWHM of 19 eV corresponding to an electronic noise of 2.2 electrons r.m.s

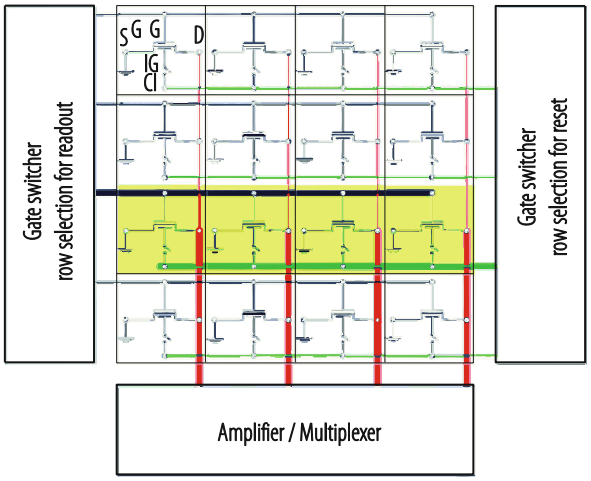


Fig. 5.43 Circuit diagram of a DEPFET pixel detector with parallel row-wise readout of the drain current

reading is a measure for the signal charge in the pixel cell. Alternatively to the procedure described above, sources may be connected column wise and source voltages measured instead of drain currents. Figure 5.44 shows the spectroscopic quality reached with a 64 × 64 DEPFET matrix of 50 × 50 μm^2 pixels.

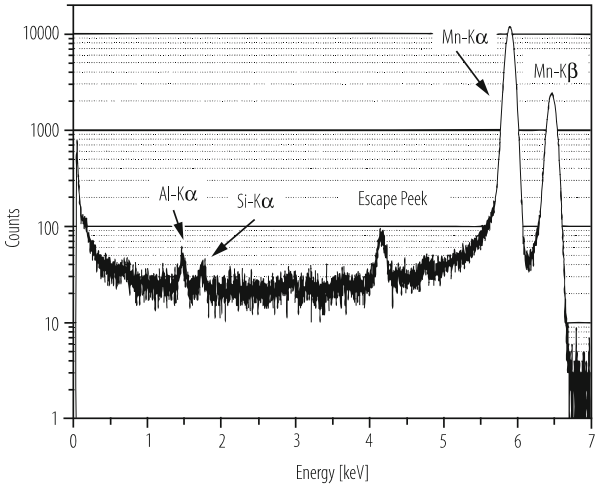


Fig. 5.44 ^{55}Fe spectrum measured at −28 °C with a 64 × 64 cell DEPFET pixel matrix with 50 μm pixel size

Pixel sensors with large pixels can be constructed by combining DEPFET structure and drift chamber principle. Large pixel may be preferred in order to increase the readout speed and reduce the number of readout channels and power consumption. It is advisable to match the pixel size to the properties of the rest of the system. Over-sampling may increase the electronic noise lead to a worse performance.

Macro Pixel DEPFET Sensors

Figure 5.45 shows the principle with a cut and a top view of a cell. The circular DEPFET structure is located in the centre of a cylindrical drift detector. Electrons created anywhere in the fully depleted bulk are driven by the suitably shaped drift field towards the internal gate below the transistor channel. For this device a new type of DEPFET has been invented that allows clearing of the signal charge with substantially lower voltage by putting the clear electrode inside the drain region located in the centre. The drain region does not consist of a highly doped *p* region but is formed by an inversion layer that is controlled by a gate voltage and automatically connected to the small drain contact. Putting a sufficiently high positive voltage on this gate, the drain assumes the role of the clear electrode, which is automatically connected to the *n*-doped clear contact.

Single pixel cells and a 4 × 4 1 mm^2 pixel matrix (Fig. 5.46) have been tested successfully. Figure 5.47 shows an ^{55}Fe spectrum taken at room temperature. Here one notices a somewhat worse spectroscopic resolution than with the small-pixel devices. This is due to the leakage current which now is collected from a volume which is larger by a factor 400. The leakage current can be suppressed by lowering the operating temperature.

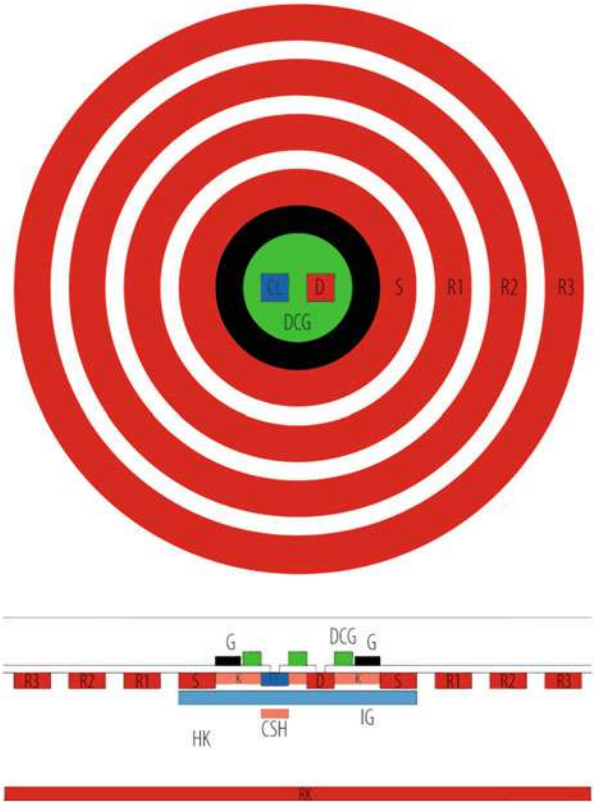


Fig. 5.45 Principle of a macro-pixel cell: A DEPFET located at the centre of a drift detector serves as storage and readout device

New DEPFET Developments

The DEPFET concept allows a variety of further functionalities that have partially been proven experimentally but not yet implemented into a large area pixel detector:

(a) As signal charge is not destroyed by the readout process this charge can be read repeatedly and the measurement precision improves with the square root of the number of measurements. This has been verified with a pair of neighbouring DEPFET transistors arranged in such a way as to allow the transfer of signal charge from one internal gate to the other and in reverse direction. A measurement precision of 0.25 electrons has been achieved independently of the amount of signal charge [39].

(b) Gatable DEPFETs [40] are developed for applications in High Time Resolution Astronomy (HTRA) and Adaptive Optics. They collect signals in preselected time intervals only, whereas the charge generated outside of these gate periods are drained towards a clear electrode.

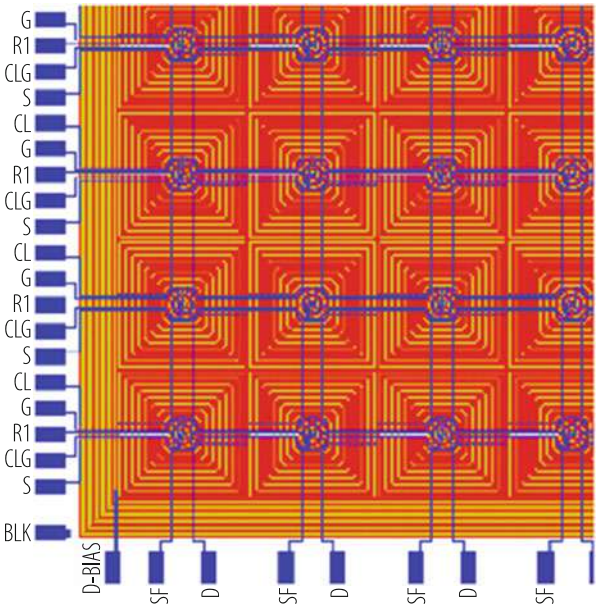


Fig. 5.46 Layout of a macro pixel matrix

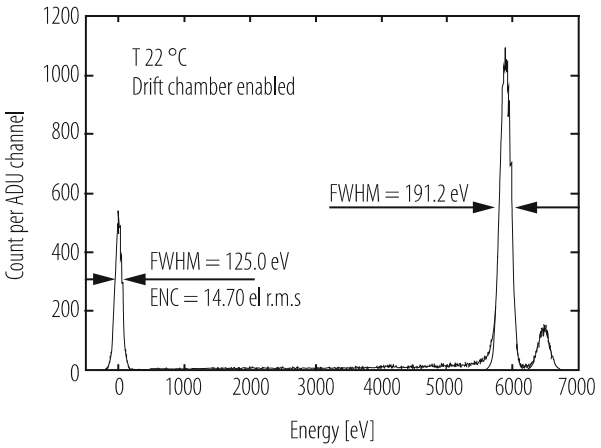


Fig. 5.47 ^{55}Fe spectrum measured at room temperature in a 1×1 mm^2 pixel of an 8×8 macro pixel matrix with 6 μs shaping. The increase of the noise compared to single DEPFET cells is due to the leakage current in the large sensitive volume of $1 \times 1 \times 0.45$ mm^3, which can be reduced by cooling

(c) Nonlinear DEPFETs [41] developed for applications at the European X-ray Free Electron Laser (EuXFEL) at Hamburg. Their non-linear characteristics and high-speed capability combines simultaneously single X-ray-photon sensitivity and very high dynamic range at the 5 MHz EuXFEL repetition rate.

DEPFET Pixel Detector Applications

In the last years DEPFET pixel detectors have been developed at the MPI Semiconductor Laboratory for the following projects:

Bepi Colombo, a mission for observing mercury [42], XEUS/IXO a space based X-ray observatory that will succeed the XMM/Newton and vertex detectors for the International Linear Collider (ILC) and the BELLE-II experiment at the KEK e^+e^- collider.

As an example for the application in X-ray detection Fig. 5.48 shows spectra at high readout rates taken with a Bepi Colombo prototype macro pixel detector. In the final detectors (Fig. 5.49) the pixel size is reduced to $300 \times 300\ \mu m^2$. An X-ray image obtained by illumination through a mask (Fig. 5.50) demonstrates functioning of the full detector.

The ILC and BELLE vertex detectors [44, 45] require fast readout (10 μs frame time), excellent spatial resolution (5 μm) and minimal material thickness to

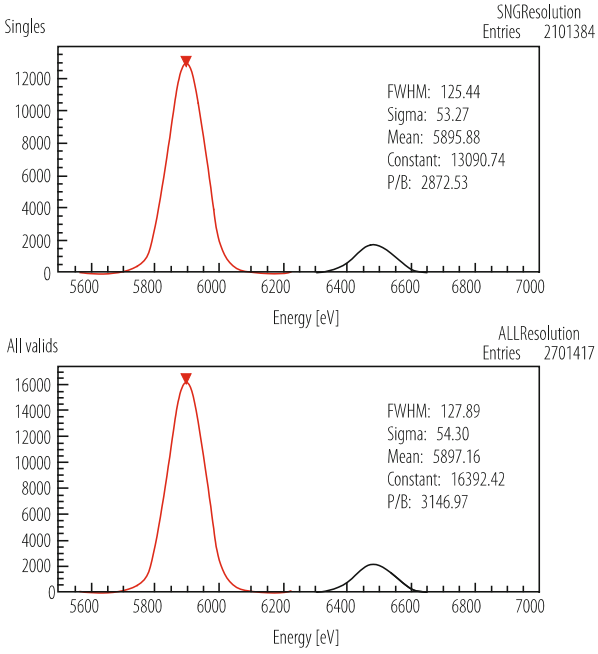


Fig. 5.48 Spectroscopic resolution of Bepi Colombo macro-pixel detectors with 64 × 64 pixels of 500 × 500 μm^2 size on a 500 μm fully depleted substrate with ultra-thin backside radiation entrance window. The top figure is restricted to photons contained in single pixels. while in the lower part signals split between neighbour pixels are included. Readout was with the ASTEROID pixel chip [43] that averages the DEPFET signals over an "integration time" once before and once after clearing and takes their difference as a measure for the deposited charge. The measured width of 125 eV FWHM with 0.9 μs integration time corresponds to an electronic noise of 4 electrons r.m.s. Reducing the integration time from 0.9 to 0.25 μs increases the width to 163 eV FWHM corresponding to 13 electron charges r.m.s

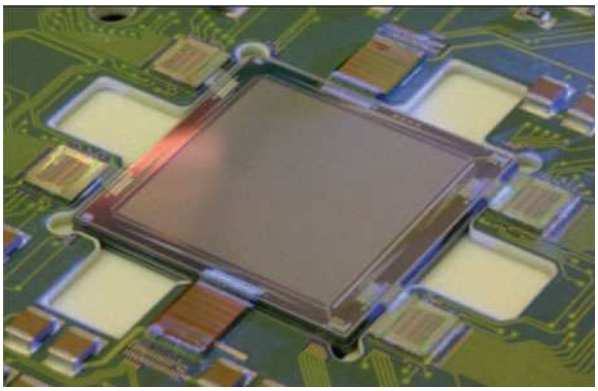


Fig. 5.49 Photo of an assembled macro-pixel detector with two 64 channel ASTEROID readout chips on top and bottom and four steering chips

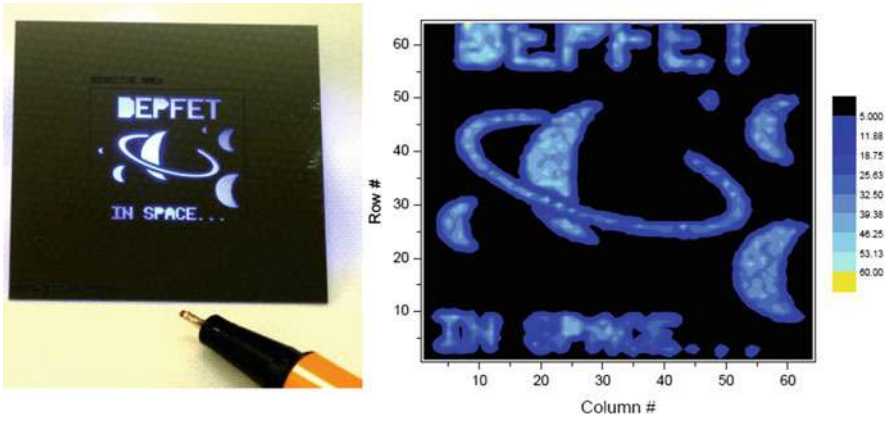


Fig. 5.50 X-ray image (right) obtained with the mask shown on the left

minimize the scattering of charged particles. Consequently the pixel size has been chosen as $25 \times 25\ \mu m^2$ for ILC and $50 \times 75\ \mu m^2$ for BELLE-II. A new method for wafer thinning based on wafer bonding technique has been developed in order to produce thin (50 μm) self-supporting all silicon modules [46].

5.11 Detectors with Intrinsic Amplification

Contrary to gas detectors, semiconductor detectors usually provide only the primary ionization as signal charge. This mode of operation is possible because of the low energy needed for producing an electron-hole pair (3.6 eV in silicon, whereas the ionization energy for gases is about 30 eV) and the availability of low noise

electronics. The measurement of the primary ionization without gain avoids any effect of gain variation or amplification noise, and thus leads to stable operation in spectroscopic measurements. However, high speed and very low noise requirements, detection of single photons, compensation for charge losses due to radiation damage or timing accuracies of the order of tens of picoseconds, make an intrinsic amplification of the detectors desirable.

A rather old and well known device is the avalanche diode, with several different operating modes. In the last two decades arrays of avalanche diodes operated in the Geiger mode (SiPMs—Silicon Photo Multipliers) have become photo-detectors of choice for many applications, and more recently tracking detectors with gain (LGAD—Low Gain Avalanche Detectors) are developed with the aim to combine precision position with precision timing in the harsh radiation environment of the high-luminosity LHC at CERN.

5.11.1 Avalanche Diode

An avalanche diode has a region with a field of sufficient strength to cause charge multiplication. An example of such a device is shown in Fig. 5.51. The base material is low doped p-type silicon. The junction, consisting of a thin highly doped n-type layer on top of a moderately doped p-layer, may also be used as entrance window for radiation, especially when the bulk material is only partially depleted.

An enlarged view of the central top region of Fig. 5.51, in which multiplication takes place, is shown in Fig. 5.52. Also shown are charge density, electric field and potential for the idealized assumption of uniform doping in the n^+-, p- and p^--regions ignoring diffusion. The middle p region is fully depleted and the space-charge region extends into the thin n^+ top region and the low doped p^--bulk. The maximum of the electric field is at the n^+p junction.

Electrons produced below the n^+p junction (and holes produced above the junction) will pass the high field region of the junction when drifting in the electric

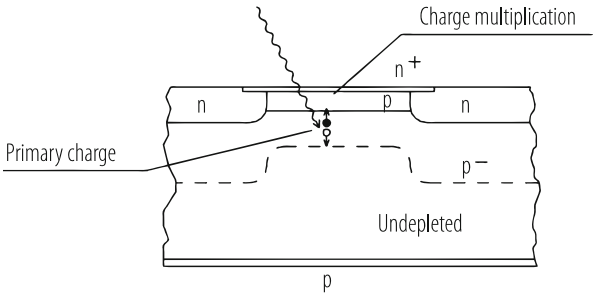


Fig. 5.51 Avalanche diode built on p-type silicon with a high-field region right below the top surface

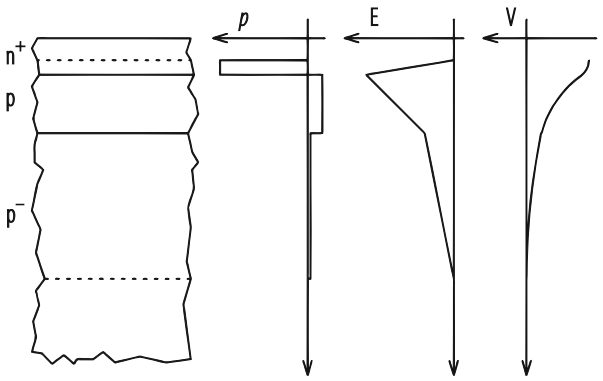


Fig. 5.52 Amplification region of the avalanche diode shown in Fig. 5.51. Also shown are charge density ρ, electric field E, and potential V

field towards the collecting electrode on top (on bottom). If the electric field is strong enough to accelerate electrons (or holes) between collisions with the lattice imperfections so that the kinetic energy is sufficient to create another electron-hole pair, the charge produced by the primary ionization is amplified.

One important aspect to be considered in designing or operating avalanche diodes is the different behaviour of electrons and holes with respect to charge multiplication. In silicon, the onset of charge amplification for holes occurs at higher electric fields than for electrons. The situation is opposite in germanium, while in GaAs the difference between electrons and holes is comparatively small.

Therefore several working regimes exist that vary depending on the strength and extension of the high electric field region. In the case of silicon one finds: (a) At low electric field, no secondary electron-hole pairs are generated. The device has the characteristics of a simple diode. (b) At higher electric field only electrons generate secondary electron-hole pairs. The amplified signal will be proportional to the primary ionization signal, with some statistical fluctuation from the multiplication process added to the fluctuation in the primary ionization process. (c) At even higher field, holes will also start to generate secondary electron-hole pairs. Secondary electrons generated by holes will again pass through (part of) the amplification region, thereby possibly generating other (tertiary) electron-hole pairs. This avalanche process will continue until it is either stopped by a statistical fluctuation in the multiplication process or by a sufficiently large drop of the externally supplied voltage. This drop may be due to the increased current passing through a bias resistor or an external enforcement by, for example, a feedback circuit. The generation of a large number of free charge carriers in the multiplication region also reduces the electric field strength and therefore decreases charge multiplication in later stages of the avalanche generation. In this operation mode the output signal is no more proportional to the primary charge; however, single photon detection becomes possible.

5.11.2 Low Intensity Light Detection

An optical photon in its primary interaction will create a single electron-hole pair, a charge too small to be detected by standard electronics. However, intrinsic amplification in an avalanche process makes single photon detection possible. The avalanche diode of Fig. 5.51 is such a device. Operation in proportional mode will result in an output signal proportional to the number of (optical) photons, with some statistical fluctuations of the avalanche process added and additional contributions from the non-uniformity of the electric field in the avalanche region. Operation in limited Geiger mode will result in a signal independent of the number of incident photons. The charge signal will be approximately given by the product of the diode capacitance times the difference of the applied voltage and the voltage at which the avalanche process stops.

As the charge multiplication probability is a strong function of the electric field strength, high uniformity over the active area is required and high field regions at the edge of the device have to be avoided by proper design. Edge breakdown is avoided in Fig. 5.51 by the less strongly doped n region at the rim. This leads to a space-charge region extending deeper into the bulk and to a reduction of the maximum field.

If the structure of Fig. 5.51 is to be operated in proportional mode (with only electrons multiplying), primary charge produced by radiation entering from the top has to be generated below the high field multiplication region in order to be properly amplified. Therefore for blue light, with its submicron penetration depth, the efficiency is low for this design.

In choosing the width of the depleted region, one has to consider several partially conflicting requirements. Based on noise considerations, this region should be large in order to reduce the capacitive load to the amplifier. The same is required for the detection of deeply penetrating radiation such as X-rays or energetic charged particles. One may even extend the depleted region all the way to the bottom surface. Then the backside p-doped surface can also be used as a radiation entrance window. This can be an advantage for low penetrating radiation such as optical photons, since such an entrance window can be made thin. The disadvantage of a large depleted region is the large volume for thermal generation of electron-hole pairs, the electrons being capable of initializing the avalanche process and, depending on the application, a not wanted sensitivity to deeply penetrating radiation.

The electric field configuration in the avalanche region is shown in an idealized way in Fig. 5.52, assuming abrupt doping changes. Such a distribution is not only unrealistic but also far from optimal for proportional operation: Breakdown should be avoided as much as possible which can be achieved by an extended amplification region and lower hole-to-electron multiplication ratios, as is the case for lower fields. Such a design can be realised by suitably doping the avalanche region.

5.11.3 Solid-State Photo Multipliers: SiPMs

In the last decade a new type of avalanche photon detector has reached maturity and is now commercially available, the Solid State Photo Multiplier, also referred to as SiPM (Silicon Photo Multiplier), G-APD (Geiger Mode Avalanche Photo Diode) or MPPC (Multi Pixel Photon Counter) [47]. It consists of two dimensional arrays of 100–10,000 single photon avalanche diodes (SPADs), called pixels, with typical dimension between $(10\ \mu m)^2$ and $(100\ \mu m)^2$. The pixels are operated in limiting Geiger mode and every pixel gives approximately the same signal, independent of the number of photons which have produced simultaneously electron-hole pairs in the amplification region of the pixel. The sum of the pixel signals is equal to the number of pixels with Geiger discharges, from which the number of incident photons can be determined. As the output charge for a single Geiger discharge is typically larger than 10^5 elementary charges, 0, 1, 2, and more Geiger discharges can be easily distinguished, enabling the detection of single optical photons with high efficiency and sub-nanosecond timing. The quenching of the Geiger discharge is either achieved by a resistor in series with each pixel or an active feedback.

Two types of SiPMs have been developed: Analogue and Digital. In Analogue SiPMs [47] the individual pixels are connected to a common readout and the SiPM delivers the summed analogue signal. In Digital SiPMs [48] each pixel has its own digital switch to a multi-channel readout system and the output is the digitized pulse height and precise time information for the pixels with Geiger discharges. Digital SiPMs also allow disabling pixels with high dark-count rates.

The pulse shape and the gain of SiPMs are explained with the help of Figs. 5.53 and 5.54: A schematic cross section of a single pixel is shown in Fig. 5.53, and an electrical model of a pixel with resistor quenching, in Fig. 5.54. The bias voltage is denoted V_{bias}, the single pixel capacitance C_{pix}, and the quenching resistance R_q. Frequently, in particular for SiPMs with larger pixel sizes, a capacitance C_q parallel to R_q is implemented. In the quiescent state the voltage over C_{pix} is V_{bias}. When an electron-hole pair in the amplification region starts a Geiger discharge, in the model the switch is closed and C_{pix} is discharged through the current source until the turn-off voltage V_{off} is reached, at which the Geiger discharge stops and the switch opens. The assumption of a constant current source is certainly oversimplified. However the sub-nanosecond discharge time is so short, that details of the time dependence of the discharge current hardly affect the results of the simulation. If a finite capacitance C_q is present, a fast pulse with charge $C_q \cdot (V_{bias} - V_{off})$ appears. After the switch opens, C_{pix} is charged up to V_{bias} with the time constant $\tau \approx R_q \cdot C_{pix}$ and the total signal charge is approximately $(C_{pix} + C_q) \cdot (V_{bias} - V_{off})$. Figures 5.55 and 5.56 show two examples of pulse shapes: (a) For a KETEK SiPM with $(15\ \mu m)^2$ pixels and negligible C_q, and (b) for a KETEK SiPM with similar doping profiles however with $(50\ \mu m)^2$ pixels and a finite C_q. The value of R_q has to be sufficiently high to quench the Geiger discharge. As C_{pix} increases with increasing pixel area, $\tau = R_q \cdot C_{pix}$ also increases, and a finite C_q has to be introduced to achieve a good timing performance and an increased pulse height if fast pulse shaping is used.

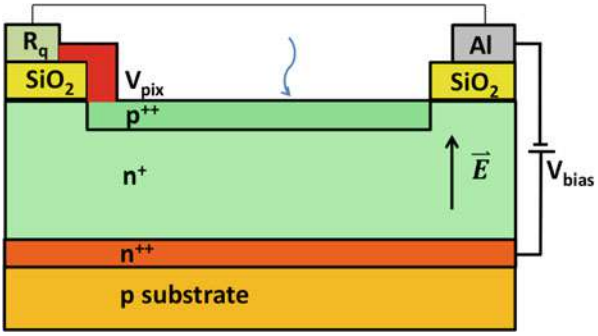


Fig. 5.53 Example of the schematic layout of a SiPM pixel. The Geiger breakdown occurs in the high-field n^+ region, which has a depth of order 1–2 μm. The p^{++}-electrode of every pixel is connected through the quenching resistance (R_q) to the biasing lines (Al) to which the biasing voltage V_{bias} is applied The photons enter through the transparent p^{++}-electrode

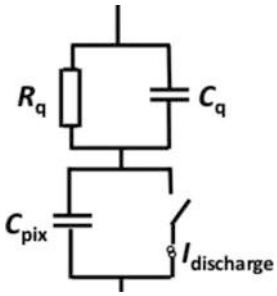


Fig. 5.54 Electrical model of a single SiPM pixel

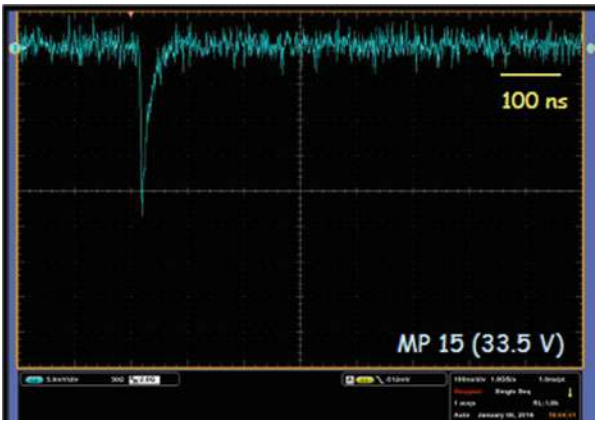


Fig. 5.55 Pulse shape from a single photon for a KETEK SiPM with 4384 pixels of $(15\ \mu m)^2$: A single exponential with $\tau = R_q \cdot C_{pix} \approx 20$ ns

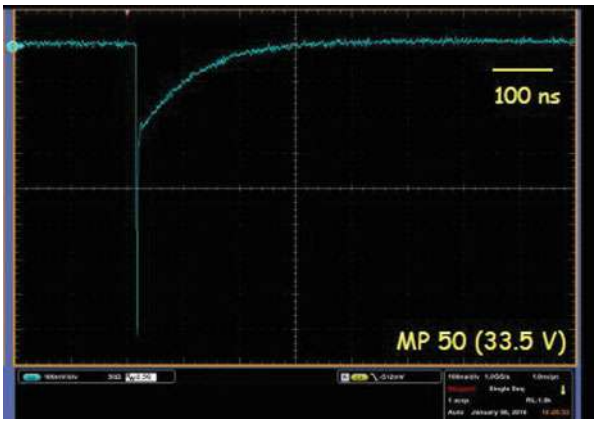


Fig. 5.56 Single photon pulse shape for a KETEK SiPM with 400 pixels of $(50\ \mu m)^2$: A prompt signal due to the finite value of C_q and a slow component with the time constant $\tau = R_q \cdot C_{pix} \approx 110$ ns is observed

In our discussion we distinguish between the breakdown voltage V_{bd}, the threshold voltage for a Geiger discharge, and the turn-off voltage V_{off}, the voltage at which the Geiger discharge stops. Differences $V_{bd} - V_{off}$ of up to about 1 V have been observed [49]. They should be taken into account when characterising or modelling SiPMs. We note that V_{bd} can be obtained from I–V measurements, as the voltage at which the current rises quickly due to the onset of Geiger discharges or the voltage at which the photon detection efficiency starts to differ from zero, and V_{off} can be determined from the dependence of SiPM Gain on V_{bias} by extrapolating the linear Gain(V_{bias}) dependence to Gain = 1.

One outstanding feature of SiPMs is the single-photon resolution, as demonstrated in the charge spectrum shown in Figs. 5.57 and 5.58 [50]. 0, 1, ... up to >30 simultaneous Geiger discharges can be distinguished allowing for straight-forward

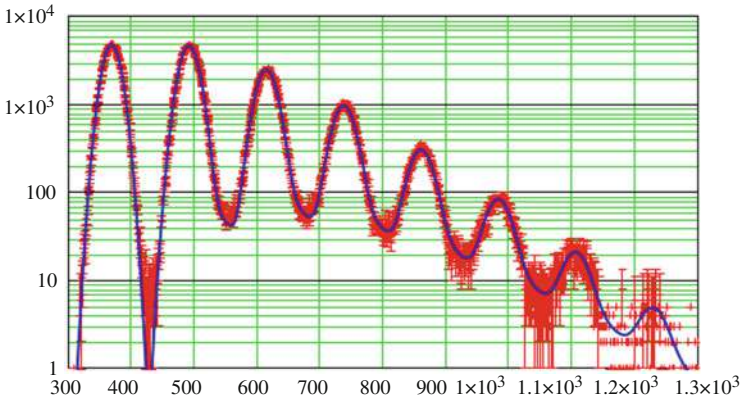


Fig. 5.57 Pulse height spectrum for a pulsed picosecond-laser measured with a KETEK SiPM with 4384 pixels of $(15\ \mu m)^2$. The solid curve is a model fit to the data. The average number of photons producing an initial Geiger discharge is 1.15

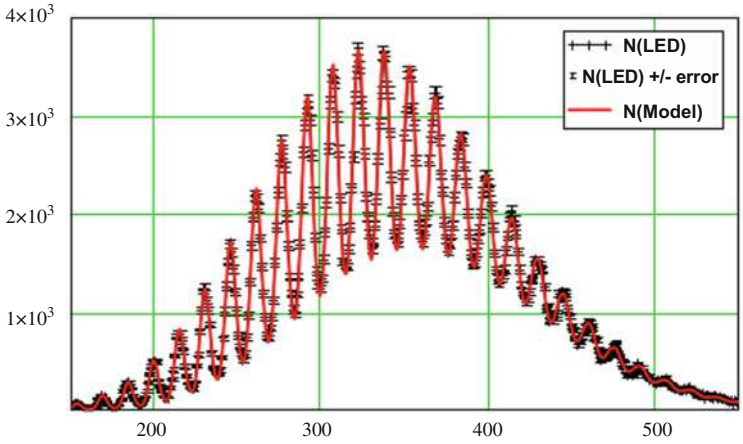


Fig. 5.58 Same as Fig. 5.57, however with an average number of photons producing an initial Geiger discharge of 18.6

calibration methods. The high photon-detection efficiency, where after careful optimisation values in excess of 60 % for wavelengths between 250 and 600 nm have been reached, the high gain of typically 10^6, and the intrinsic timing resolution of a few picoseconds, are other attractive performance parameters. In addition, SiPMs are not affected by magnetic fields, operate in a wide temperature range, are very robust, and work at moderate bias voltages ($\approx$ 25–75 V). Also, thanks to the microelectronics technology, SiPMs have highly reproducible performance parameters and are relatively inexpensive.

Limitations of SiPMs are their size, which is typically below 1 cm^2, and their limited dynamic range, essentially determined by the number of pixels. In addition, the measurement of the number of photons is affected by two sources of excess noise, which worsen the resolution beyond Poisson statistics: After-pulsing and Cross-talk. After-pulses are the result of charge carriers which are produced in the Geiger discharge and trapped in defect states. Depending on the energy in the silicon band gap and the properties of the defect states, they are released with different de-trapping time constants and cause additional signal fluctuations, which depend on the integration time of the readout electronics. In Figs. 5.57 and 5.58, which show pulse-height spectra recorded with a 100 ns gate at room temperature, after-pulses can be seen as entries in-between the peaks. Cross-talk is produced by the photons from the accelerated charges in the Geiger discharge, which generate electron-hole pairs in adjacent SiPM pixels. The photon path can be inside of the silicon but also via reflection in the protective layer of the SiPM or a light guide. This light path is so short that this cross-talk can be considered as prompt. Implementing trenches filled with absorbing material in-between the pixels reduces the prompt cross-talk significantly. The photons from the Geiger discharge can also generate electron-hole pairs in the non-depleted region of the SiPM, which can diffuse into

the amplification region and cause delayed cross-talk. The result of prompt cross-talk is that the number of entries in the peaks does not follow a Poisson distribution, even if the number of photons causing initial Geiger discharges does. As shown in [51] the result of cross-talk is that the number of entries in the peaks follows a Generalised-Poisson instead of a Poisson distribution. We note that the solid curve shown in Figs. 5.57 and 5.58 is the result of a model fit which includes both after-pulsing and prompt cross-talk simulated by a Generalised Poisson distribution. The model provides a fair description of the measurements and gives a precise determination of the SiPM parameters [50]. As both, after-pulses and cross-talk are related to the number of charge carriers in the Geiger discharge and thus to the Gain, the corresponding probabilities are expected to be approximately proportional to $V_{bias} - V_{off}$, which is also observed. Typical values at $V_{bias} - V_{off} = 5$ V for after-pulsing as well as prompt cross-talk are 5 % resulting in an excess noise factor, the ratio of the square of the relative resolution to the Poisson expectation, ENF $= [(\sigma_{meas}/mean_{meas})/(\sigma_{Poisson}/mean_{Poisson})]^2$ of ≈ 1.08. As the photon detection efficiency increases with voltage and finally saturates, whereas Gain and ENF continue to increase, there is a voltage at which the photon number measurement is optimal.

Dark counts are another limitation of SiPMs. Typical dark count rates (DCR) for SiPMs before irradiation are between 10 and 100 kHz/mm^2 at room temperature. Cooling reduces the DCR by about a factor 2 for an 8 °C reduction in temperature. Ionizing radiation, which mainly causes damage to the SiO_2, hardly affects the DCR. However non-ionizing radiation, like neutrons or high energy (> 5 MeV) particles, significantly affect the performance. At sufficiently high fluences (Φ) the DCR is so high that most pixels are in a state of Geiger discharge, the photon-detection efficiency decreases and finally the SiPM stops working as a photo-detector. Whereas V_{bd} and the electrical SiPM parameters hardly change up to $\Phi = 5 \times 10^{13}$ cm^{-2}, DCR increases by many orders of magnitude: For a KETEK SiPM with 15 μm pitch at –30 °C and $(V_{bias} - V_{off}) = 5$ V, DCR increases from ≈ 10 kHz/mm^2 before irradiation to ≈ 200 GHz/mm^2 after irradiation by reactor neutrons to $\Phi = 5 \times 10^{13}$ cm^{-2} [52, 53]. It is found that the increase in DCR is approximately proportional to Φ. It is also observed that after irradiation the increase of DCR with excess voltage is significantly steeper and the decrease with temperature slower after than before irradiation. As a result of the increased DCR, the signal baseline shows large fluctuations and single photon detection becomes impossible. Finally the occupancy of the pixels by dark counts is so high that the probability of a photon hitting a pixel which is already busy increases and the photon detection efficiency degrades. For the KETEK SiPM with 15 μm pitch at –30 °C the photon detection efficiency due to dark counts is reduced by a factor 2 for $\Phi = 5 \times 10^{13}$ cm^{-2} at $(V_{bias} - V_{off}) \approx 2.5$ V, and essentially zero for $\Phi = 5 \times 10^{14}$ cm^{-2} [53]. At these high fluences the dark currents exceed several tens of mA and thermal run-away has to be avoided.

After irradiation a significant reduction of DCR by annealing occurs. Annealing is a strong function of temperature: The typical reduction of DCR is a factor 2–3 after several days at room temperature, and a factor 10–50 at 175 °C. A systematic

study of different annealing scenarios, which allows to optimise the temperature cycling for operating SiPMs in high radiation fields, as available for silicon tracking detectors without gain [7, 10], is so far not available. In [54] it is demonstrated that SiPMs produced by Hamamatsu and SENSL, after irradiation to a fluence of 10^{14} cm^{-2} and annealed at 175 °C can achieve single photon detection at 77 K with a DCR below 1 kHz/cm^2.

The values of V_{bd} and V_{off} have a temperature dependence of order 20 mV/°C, which results in a temperature-dependent gain. However this is not a real problem and several feedback systems for gain stabilisations have been designed and are used.

Due to the vast application potential, which spans from research, over industrial applications to medicine, several firms develop and manufacture SiPMs. In close collaboration with research institutions, in particular working in particle physics, a rapid development and major improvements of SiPMs are presently under way.

5.11.4 Ultrafast Tracking Detectors: LGADs

At the HL-LHC (High-Luminosity Large Hadron Collider at CERN planned to start operation in 2026) in the large collider experiments ATLAS and CMS there will be on average ≈ 200 interactions with vertices distributed over ≈ 10 cm along the beam direction for every bunch crossing. For the complete kinematic reconstruction of the most interesting interactions in a bunch crossing, the information of the individual detector components has to be assigned to the correct interaction vertices. To illustrate the problem, Fig. 5.59 shows the reconstructed tracks extrapolated to the interaction region for a single bunch crossing with 50 interactions recorded in 2012. For a few vertices the interaction times, which are spread over ≈ ±200 ps, as obtained from a simulation, are given. For an efficient assignment of tracks to vertices, tracking detectors with high efficiency, 5 μm position resolution, 20 ps

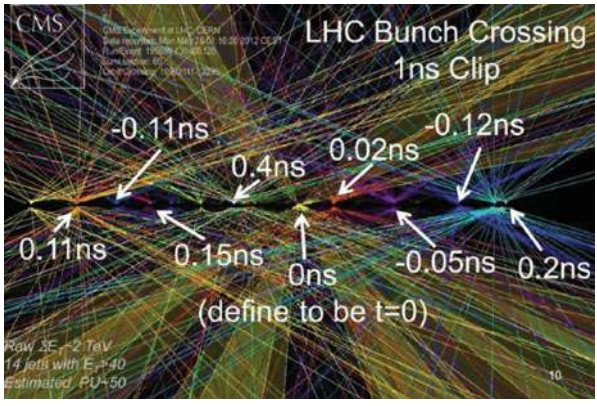


Fig. 5.59 Interaction times of a number of proton-proton vertices in a single bunch crossing with 50 interactions [55]. The data have been recorded by the CMS experiment in 2012. At the HL-LHC, the average number of interactions per bunch crossing is expected to be about 200

timing accuracy and 25 ns pulse shaping, are required. From simulations [55] it is concluded that pixel sensors with 50 μm active thickness and a doping profile similar to the one shown for APDs in Fig. 5.52 and operated at a gain of ≈ 20 can reach the required performance. These detectors are called Low-Gain-Avalanche Detectors, LGADs.

Different to optical photons which generate single electron-hole pairs, minimum-ionizing produce about 75 charge pairs per micro-meter and a high gain is not required. In addition to increasing fluctuations, high gain causes also practical difficulties and increases the shot noise from the dark current. Thin sensors have the additional advantage of smaller dark currents and a pulse rise time which increases with decreasing sensor thickness.

The effects which influence the timing accuracy can be grouped in five categories: (1) Position-dependent fluctuations of the charge carriers produced by the charged particle to be measured, (2) excess noise of the amplification mechanism, (3) position dependent drift field and coupling of the of the drifting charges to the readout electrodes, (4) electronics noise, and (5) digitisation error of the time-to-digital convertor.

A major issue for LGADs is the control of the gain after irradiation. The change of the effective doping by dopant removal and defect states, and the decrease of the mobilities and amplification coefficients of electrons and holes due to radiation damage appear to present major problems. These are addressed in an extensive R&D program which started in 2012 and has already given first encouraging results.

5.12 Summary and Outlook

Different concepts of solid silicon sensors and the electronics required for their readout have been described in this contribution. Although a detailed theoretical understanding of silicon devices had already been achieved in the 1960s, silicon detectors remained a niche application, used mainly in Nuclear Physics. This changed around 1980, when Josef Kemmer adapted the planar technology of micro-electronics to sensor fabrication and the ACCMOR Collaboration demonstrated the reliable long-term operation and excellent physics performance of silicon strip detectors. Based on these results, many groups started to develop and use silicon detectors, and today there is hardly a particle physics experiment, which does not rely heavily on them. The areas covered by silicon detectors in the particle physics experiments increased from tens of cm^2 to hundreds of m^2. Large areas of silicon detectors are even used on satellites for space experiments. In parallel to silicon detectors, the development of low-noise ASICs and connection technology started. They are required for reading out the more and more complex silicon sensors. In addition, a number of industrial producers, in closed collaboration with academia, developed and fabricated silicon sensors. Today silicon radiation detectors are a quite big market. Initially developed for Particle Physics, the use of silicon detectors spread into many different fields of science, medicine and industrial applications.

Since 1980 several new detector concepts were proposed, realised and used for a variety of measurement tasks. Outstanding examples are drift detectors, fully depleted CCDs, DEPFETs, MAPSs 3-D sensors, APDs and SiPMs. The different devices have their advantages and shortcomings, but offer high-performance solutions for most measurement tasks. In recent years radiation damage for the use of silicon sensors at high flux or high luminosity colliders has become more and more of a concern. Whereas radiation damage by X-rays can be controlled by a proper sensor design, the question up to which fluence of high-energy radiation silicon detectors can be used is a field of intense research. Unfortunately other sensor materials, like crystalline diamond or GaAs seem not to be a solution. Defect engineering, by doping crystals with different impurities has resulted in some improvements. However, a breakthrough for high fluences could not be demonstrated. Therefore the only approach appears to optimise the sensor layout for radiation tolerance. The recipe followed are high fields and low charge collection distances. How far intrinsic amplification can help remains an open question. For the design optimisation, complex TCAD (Technology Computer-Aided Design) simulations are performed. In spite of some first successes, a major progress is still required. As far as the electronics, which is exposed to the same fluences, is concerned, the sub-micron technology with nano-meter dielectric layers resulted in a big step in radiation tolerance.

For the future there is the strong hope that detectors can be fabricated which achieve the challenging performance parameters in the high radiation fields of the HL-LHC and future high-luminosity colliders. The field of solid state detectors will also profit very much from the ongoing industrial R&D efforts, in particular of 3-D integration technology and nano-electronics. Last but not least I very much hope that, like in the past, radically new ideas will come up and expand further the applications of solid state detectors.

References

1. E. Gatti, P. Rehak: *Semiconductor Drift Chamber - An Application of a Novel Charge Transport Scheme*, Nucl. Instrum. Meth. 225 (1984) 608-614; E. Gatti et al.: *Silicon Drift Chambers - First results and optimum processing of signals*, Nucl. Instrum. Meth. 226 (1984) 129-141; E. Gatti et al.: *Semiconductor Drift Chambers*, IEEE Trans. Nucl. Sci. 32 (1985) 1204-1208.
2. L. Strüder et al.: *The MPI/AIT X-ray imager (MAXI) - high speed pn-CCDs for X-ray detection*, Nucl. Instrum. Meth. A 288 (1990) 227-235; L. Strüder et al.: *First results with the pn-CCD detector system for the XMM satellite mission*, Nucl. Instrum. Meth. A 326 (1993) 129-135; L. Strüder et al.: *A 36 cm^2 large monolythic pn-charge coupled device X-ray detector for the European XMM satellite mission*, Rev. Sci. Instrum. 68 (1997) 4271-4274.
3. J. Kemmer, G. Lutz: *New semiconductor detector concepts*, Nucl. Instrum. Meth. A 253 (1987) 356-377.
4. J. Zhang et al.: *Study of radiation damage induced by 12 keV X-rays in MOS structures built on high-resistivity n-type silicon,* Journal of Synchrotron Radiation 19 (2012) 340-376.
5. T. Poehlsen, et al.: *Charge losses in segmented silicon sensors at the Si-SiO2 interface*, Nucl. Instrum. Meth. A 700 (2013) 22-39.

6. J. Schwandt et al.: *Design and First Tests of a Radiation-Hard Pixel Sensor for the European X-Ray Free-Electron Laser,* IEEE TRANSACTIONS ON NUCLEAR SCIENCE, VOL. 61, NO. 4, AUGUST 2014 1894-1901.
7. M. Moll: *Displacement Damage in Silicon Detectors for High Energy Physics*, Manuscript accepted for Publication in IEEE Transactions on Nuclear Science, DOI:https://doi.org/10.1109/TNS.2018. 2819506.
8. V. Eremin, E. Verbitskaya, Z. Li: *The origin of double peak electric field distribution in heavily irradiated silicon detectors,* Nucl. Instrum. Meth. A 476 (2002) 556-564.
9. R. Klanner et al.: *Determination of the electric field in highly-irradiated silicon sensors using edge-TCT measurements,* Nucl. Instrum. Meth. A 951 (2020) 162987.
10. R. Wunstorf et al.: *Results on Radiation Hardness of Silicon Detectors up to Neutron Fluences of* $10^{15} n/cm^{-2}$, Nucl. Instrum. Meth. A 315 (1992) 149-155.
11. S.I Parker, C.J. Kenney and J. Segal: *3D - A proposed new architecture for solid state radiation detectors*, Nucl. Instrum. Meth. A 395 (1997) 328.
12. J. Kemmer et al.: *Experimental confirmation of a new semiconductor detector principle*, Nucl. Instrum. Meth. A 288 (1990) 92-98.
13. M. Caccia et al.: *A Si Strip Detector with Integrated Coupling Capacitors*, Nucl. Instrum. Meth. A 260 (1987) 124-131.
14. C. Cottini, E. Gatti, G. Gianelli, G. Rozzi: *Minimum noise preamplifiers for fast ionization chamber*, Nuovo Cimento (1956) 473-483.
15. E. Gatti, P.F. Manfredi: *Processing the signals from solid state detectors in elementary particle physics*, Rivista di Nuovo Cimento 9, Ser. 3 (1986) 1-145.
16. W. Buttler et al.: *Low-noise, low power monolithic multiplexing readout electronics for silicon strip detectors*, Nucl. Instrum. Meth. A 273 (1988) 778-783.
17. P. Jarron, et al.: *Deep submicron CMOS technologies for the LHC experiments*, Nucl. Phys. B Proc. Suppl. 78, no. 1-3 (1999) 625-634.
18. P. Rehak et al.: *Semiconductor drift chambers for position and energy measurements*, Nucl. Instrum. Meth A 235 (1985) 223-234.
19. A. Castoldi et al.: *A new drift detector with reduced lateral diffusion*, Nucl. Instrum. Meth A 377 (1996) 375-380.
20. W. Chen et al.: *Large area cylindrical silicon drift detector*, IEEE Trans. Nucl. Sci. 39 (1992) 619-628.
21. P. Rehak et al.: *Spiral silicon drift detectors*, IEEE Trans. Nucl. Sci. 36 (1989) 203-209.
22. R. Hartmann, et al.: *Ultrathin entrance windows for silicon drift detectors*, Nucl. Instrum. Meth. A 387 (1997) 250-254.
23. R. Hartmann et al.: *Design and test at room temperature of the first silicon drift detector with on-chip electronics*, IEDM Technical Digest (1994) 535-539.
24. V. Radeka et al.: *Implanted silicon JFET on Completely Depleted High Resistivity Devices*, IEEE El. Dev. Lett. 10, nb. 2 (1989) 91-95; E. Pinotti et al.: *The pn-CCD On-Chip Electronics*, Nucl. Instrum. Meth. A 326 (1993) 85-92.
25. R. Bailey et al.: *First Measurements of Efficiency and Precision of CCD Detectors for High Energy Physics*, Nucl. Instrum. Meth. 213 (1983) 201-215; C.J.S. Damerell et al.: *CCDs for Vertex Detection in High Energy Physics*, Nucl. Instrum. Meth. A 253 (1987) 478-481; C.J.S. Damerell et al.: *A CCD based vertex detector for SLD*, Nucl. Instrum. Meth. A 288 (1990) 236-239.
26. K. Abe et al.: *Design and performance of the SLD vertex detector: a 307 Mpixel tracking system*, Nucl. Instrum. Meth. A 400 (1997) 287-343.
27. D.H. Lumb et al.: *X-ray Multi-Mirror Mission - an overview*, SPIE 2808 (1997) 326-337.
28. G. Richter et al.: *ABRIXAS, A Broadband Imaging X-ray All-sky Survey*, (L. Bassani, G. di Cocco, eds.): Imaging in High Energy Astronomy, Experim. Astron. (1996) 159.
29. N. Meidinger et al.: *The PN-CCD detector for XMM and ABRIXAS*, SPIE 3765 (1999) 192-203.
30. L. Strüder et al.: *pnCCDs on XMM-Newton – 42 months in orbit*, Nucl. Instrum. Meth. A 512 (2003) 386-400.

31. L. Strüder et al.: *Large format, high-speed, X-ray pnCCDs combined with electron and ion imaging spectrometers in a multipurpose chamber for experiments at 4th generation light sources*, Nucl. Instrum. Meth A614 (2010) 483-496.
32. B. Henrich et al.: *The adaptive gain integrating pixel detector AGIPD: A detector for the European XFEL*, Nucl. Instrum. Meth. A 633 (2011) S11-S14; A. *Allahgholi et al.*: *The adaptive gain integrating pixel detector*, JINST 11 (2016) C02066.
33. B. Henrich et al.: *PILATUS: A single photon counting pixel detector for X-ray applications*, Nucl. Instrum. Meth. A 607 (2009) 247-249.
34. R. Ballabriga, M. Campbell, X. Llopart, *ASIC Developments for radiation imaging applications: The medipix and timepix family*, Nucl. Instrum. Meth. A 878 (2018) 10-23.
35. W. Dulinski et al.: *Beam telescope for medium energy particles based on thin, submicron precision MAPS*, in: Nuclear Science Symposium Conference Record, 2007, NSS '07, IEEE, Vol. 2, 995-1002.
36. L. Ratti et al., *CMOS MAPS with fully integrated, hybrid-pixel-like analog front-end electronic*, eConf C0604032 (2006) S.0008.
37. I. Peric, *A novel monolithic pixelated particle detector implemented in high-voltage CMOS technology*, Nucl. Instrum. Meth. A 582 (2007) 876-885; I Peric et al.: *High-voltage pixel detectors in commercial CMOS technologies for ATLAS, CLIC and Mu3e experiments*, Nucl. Instrum. Meth. A 731 (2013) 131-136.
38. XEUS Astrophysics working group: *X-ray Evolving - Universe Spectroscopy - The XEUS scientific case*, ESA SP-1238 (1999), 30 pages.
39. S. Wölfel et al.: *Sub electron noise measurements on repetitive non-destructive readout devices*, Nucl. Instrum. Meth. A 566 (2006) 536-539.
40. G. Lutz, R.H. Richter, L. Strüder: *Halbleiterstruktur, insbesondere in einem Halbleiterdetektor, und zugehöriges Betriebsverfahren*, EU Patent 1 873 834; G. Lutz, et al.: *DEPFET detector-amplifier structure: Properties, achievements and new developments, concepts and applications*, in: Nuclear Science Symposium Conference Record, 2007, NSS '07, IEEE, Vol. 2, 988-994.
41. G. Lutz, L. Strüder: *DEPFET Transistor mit großem Dynamikbereich und Halbleiterdetektor*, DE Patent 10 2007 048 890; G. Lutz et al.: *DEPFET Sensor with intrinsic signal compression developed for use at the XFEL free electron laser radiation source*, Nucl. Instrum. Meth. A 624 (2010) 528-532.
42. J. Treis et al,: *DEPFET based instrumentation for the MIXS focal plane on BepiColombo*, in: Instrumentation and Methods for Astrobiology and Planetary Missions XII (R.B. Hoover, G.V. Levin, A. Yu Rozanov, K. Retherford, eds.), Proc. SPIE 7441 (2009) 774116.
43. M. Porro et al.: *Performance of ASTEROID: A 64 channel ASIC for source follower readout of DEPFET matrices for X-ray astronomy*, IEEE Nuclear Science Symposium Conference Record 2008, pp. 1830-1835.
44. R.H. Richter et al.: *Design and technology of DEPFET pixel sensors for linear collider applications*, Nucl. Instrum. Meth. A511 (2003) 250-256.
45. L. Andricek et al.: *The MOS-type DEPFET pixel sensor for the ILC environment*, Nucl. Instrum. Meth. A 565 (2006) 165-171.
46. L. Andricek, G. Lutz, M. Reiche, R.H. Richter: *Processing of ultra-thin silicon sensors for future e^+e^- linear collider experiments*, IEEE Trans. Nucl. Sci. 51 (2004) 1117-1120.
47. D. Renker: *Geiger-Mode Avalanche Photodiodes, history properties and problems*, Nucl. Instrum. Meth. A 567 (2006) 48-56; D. Renker and E. Lorenz: *Advances in solid state photon detectors*, JINST 4 (2009) P04004.

48. T. Frach et al.: *The Digital Silicon Photomultiplier – Principle of Operation and Intrinsic Detector Performance*, in: Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC), N34-4 (2012) 1959-1965; C. Degenhardt et al.: *The digital Silicon Photomultiplier – A novel sensor for the detection of scintillation light*, in Proceeding to IEEE NSS-MIC conference, Orlando U.S.A. October 25–31 2009, Proc. IEEE 2009 (2009) 2383-2386; S. Mandai, E. Charbon: *Multi-channel digital SiPMs: concept, analysis and implementation*, in: Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC), N34-4 (2012) 1840-1844.
49. V. Chmill et al.: *Study of the breakdown voltage of SiPMs*, Nucl. Instr. Meth. A 845 (2017) 56-59.
50. V. Chmill, et al.: *On the characterisation of SiPMs from pulse-height spectra*, Nucl. Instr. Meth. A 854 (2017) 70-81.
51. S. Vinogradov: *Analytical models of probability distribution and excess noise factor of solid state photomultiplier signals with crosstalk*, Nucl. Instrum. Meth. A 695 (2012) 247-251.
52. Yu. Musienko et al.: *Radiation damage studies of silicon photomultipliers for the CMS HCAL phase I upgrade,* Nucl. Instr. Meth. A787 (2015) 319–322; Yu. Musienko et al.: *Effects of very high radiation on SiPMs*, Nucl. Instr. Meth. A824 (2016) 111-114.
53. M. Centis Vignali et al.: *Neutron irradiation effect on SiPMs up to* $\Phi = 5\text{x}10^{14}$ cm^{-2}, Nucl. Instr. Meth. A912 (2018) 137.
54. M. Calvi et al.: *Single photon detection with SiPMs irradiated up to* 10^{14} cm^{-2} *1-MeV-equivalent neutron fluence*, arXiv:1805.07154 (2018).
55. N. Cartiglia et al.: *Tracking in 4 dimensions*, Nucl. Instr. Meth. A845 (2017) 47-51.

Chapter 6
Calorimetry

Particle Detectors and Detector Systems

C. W. Fabjan and D. Fournier

6.1 Introduction, Definitions

In particle physics, calorimetry refers to the absorption of a particle and the transformation of its energy into a measurable signal related to the energy of the particle. In contrast to tracking a calorimetric measurement implies that the particle is completely absorbed and is thus no longer available for subsequent measurements.

If the energy of the initial particle is much above the threshold of inelastic reactions between this particle and the detector medium, the energy loss process leads to a cascade of lower energy particles, in number commensurate with the incident energy. The charged particles in the shower ultimately lose their energy through the elementary processes mainly by ionization and atomic level excitation. The neutral components in the cascade (γ, n,..) contribute through processes described later in this section.

The sum of the elementary losses builds up the calorimetric signal, which can be of ionization or of scintillation nature (or Cherenkov) or sometimes involve several types of response.

While the definition of calorimetry applies to both the low energy case (no showering) and the high energy case (showering), this section deals mostly with the showering case. Examples of calorimetry without showering are discussed in Sect. 6.2.3.

C. W. Fabjan
Austrian Academy of Sciences and University of Technology, Vienna, Austria
e-mail: Chris.Fabjan@cern.ch

D. Fournier (✉)
IJCLab, Université Paris-Saclay, CNRS/IN2P3, Orsay, France
e-mail: daniel.fournier@cern.ch

C. W. Fabjan, H. Schopper (eds.), *Particle Physics Reference Library*,
https://doi.org/10.1007/978-3-030-35318-6_6

Only electromagnetic and strong interactions contribute to calorimetric signals, the weak (and gravitational) interaction being much too small to contribute. Particles with only weak (or gravitational interaction) will escape direct calorimetric detection. An exception are the neutrino detectors discussed in Sect. 6.4: statistically, when a very large number of neutrinos cross a detector, a tiny fraction of them will interact (weakly) with matter and will lead to particle production which can be measured by different methods, including calorimetry.

The measurement of the energy of a particle is the primary goal of calorimetry. In addition, several other important quantities can be extracted, such as impact position and timing, particle direction and identification. These issues are considered in Sects. 6.4–6.6, before addressing specific examples in Sect. 6.7.

In Sect. 6.2 the fundamentals of calorimetry are presented, followed by a discussion of signal formation obtained from the energy deposition (Sect. 6.3).

In recent years, calorimetry in the ATLAS and CMS detectors at the LHC played an essential role in the discovery of the Higgs boson, announced in July 2012.

6.2 Calorimetry: Fundamental Phenomena

Given the large differences between electromagnetic interactions and strong interactions, the following subsections start with electrons and photons, which have only electromagnetic interactions (see however the end of this section), before addressing the case of particles with strong interactions, also called hadrons. The case of muons is considered in a separate subsection.

6.2.1 Interactions of Electrons and Photons with Matter

Several elementary interaction processes of the electrons with the medium contribute to the energy loss $-\mathrm{d}E$ of an electron of energy E after a path dx in a medium: Møller scattering, ionization and scattering off the nuclei of the medium: bremsstrahlung (Fig. 6.1). Electron-electron scattering is considered as ionization (Møller) if the energy lost is smaller (larger) than $m_e c^2/2$. It is customary to include

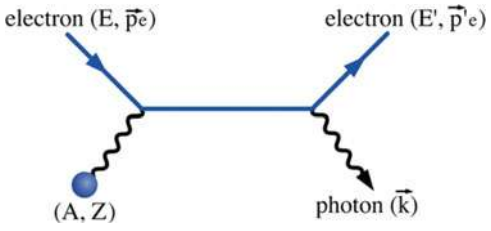


Fig. 6.1 Photon radiation from electron interaction with a nucleus (A, Z)

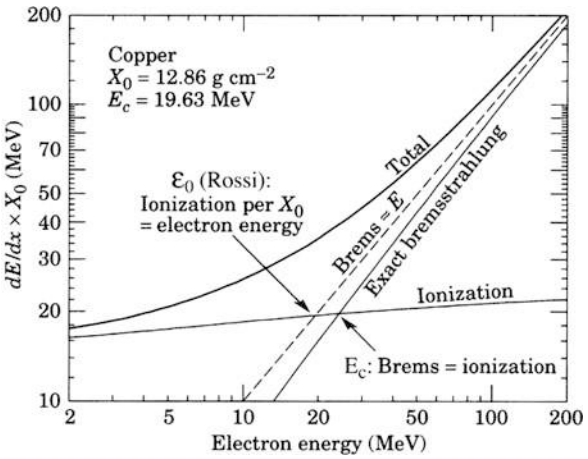


Fig. 6.2 Average energy loss of electrons in copper by ionization and bremsstrahlung. Two definitions of the critical energy (E and ε (Rossi)) are shown by arrows

in energy loss by ionization atomic excitations, some of which lead to light emission (scintillation). For positrons the Møller scattering is replaced by Bhabha scattering.

The calculated average energy loss is shown in Fig. 6.2 for copper and the average fractional energy loss ($-1/E\ \mathrm{d}E/\mathrm{d}x$) is plotted in Fig. 6.3 for lead [1].

Figure 6.2 illustrates that the average energy lost by electrons (and positrons see Fig. 6.3) by ionization is almost independent of their incident energy (above ~1 MeV), with however a small logarithmic increase. For electrons [1, 2].

$$\frac{-\mathrm{dE}}{\mathrm{dx}} = \mathrm{k}\frac{\mathrm{Z}}{\mathrm{A}}\frac{1}{\beta^2}\left\{\ln\frac{\gamma \mathrm{m_e c^2}\beta\sqrt{\gamma-1}}{\mathrm{I}\sqrt{2}} + \frac{1}{2}\left(1-\beta^2\right) - \frac{2\gamma-1}{2\gamma^2}\ln 2 + \frac{1}{16}\left(\frac{\gamma-1}{\gamma}\right)^2\right\}\left(\mathrm{MeV}/\left(\mathrm{g/cm^2}\right)\right) \tag{6.1}$$

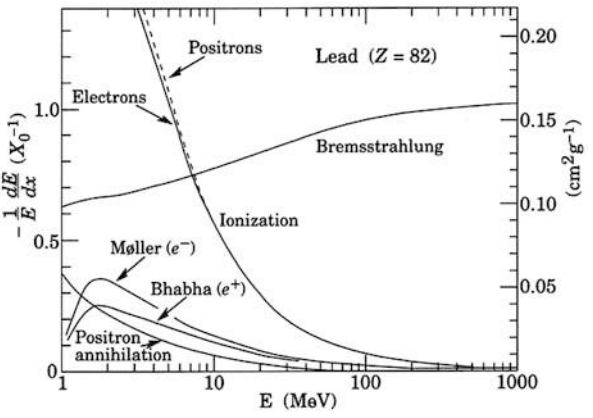


Fig. 6.3 Relative energy loss of electrons and positron in lead with the contributions of ionization, bremsstrahlung, Møller (e^-) and Bhabha (e^+) scattering and positron annihilation

and for positrons

$$-\frac{\mathrm{dE}}{\mathrm{dx}} = \mathrm{k}\frac{\mathrm{Z}}{\mathrm{A}}\frac{1}{\beta^2}\left[\ln\frac{\gamma \mathrm{m_e c^2}\beta\sqrt{\gamma-1}}{\mathrm{I}\sqrt{2}} - \frac{\beta^2}{24}\left(23+\frac{14}{\gamma+1}+\frac{10}{(\gamma+1)^2}+\frac{4}{(\gamma+1)^3}\right)\right]\left(\mathrm{MeV}/\left(\mathrm{g/cm^2}\right)\right) \tag{6.2}$$

In these formula, A (Z) are the number of nucleons (protons) in the nuclei of the medium, I is the mean excitation energy of the medium—often approximated by 16 $Z^{\mathbf{0.9}}$ eV—the constant $k = 4\pi N_\mathrm{A} {r_\mathrm{e}}^2 m_\mathrm{e} c^2 = 0.3071$ MeV/(g/cm^2), N_A the Avogadro number and $\mathrm{r_e} = \frac{1}{4\pi\varepsilon_0}\cdot\frac{\mathrm{e}^2}{\mathrm{m_e c^2}} = 2.818\ 10^{-15}$ m the classical radius of the electron.

For positrons the annihilation with an electron of the medium has to be considered. The cross section of this process ($\sigma_{\mathrm{an}} = Z\pi\,{\mathrm{r}_\mathrm{e}}^2/\gamma$ for $\gamma >> 1$) decreases rapidly with increasing energy of the positron. At very low energy, the annihilation rate is:

$$R = NZ\ \pi\ {\mathrm{r_e}}^2\ c\ \left[\mathrm{s}^{-1}\right], \tag{6.3}$$

with $N = \rho\ N_\mathrm{A}/A$, the number of atoms per unit volume.

This rate corresponds to a lifetime in lead of about 10^{-10}s [3]. Positron annihilation plays a key role in some technical applications (Positron Emission Tomography, Chap. 7).

Figure 6.2 shows that the average energy loss by bremsstrahlung (photon emission in the electromagnetic field of a nucleus) increases almost linearly as a function of incident energy (meaning that the fractional energy loss is almost constant, as shown in Fig. 6.3).

This is described by introducing the *radiation length* X_0 defined by:

$$-\mathrm{d}E/E = \mathrm{d}x/X_0 \tag{6.4}$$

It follows from the definition that X_0 is the mean distance after which an electron has lost, by radiation, all but a fraction 1/e of its initial energy. X_0 also has a simple meaning in terms of photon conversion (see below).

While X_0 should show a small increase at low energy corresponding to a small drop in the fractional energy loss visible in Fig. 6.3, it soon reaches a high-energy limit which has been calculated by Bethe and Heitler [3, 4] and more recently by Tsai [5] and tabulated by Dahl [1] for different materials. In the seminal book by Rossi [6] the formula for X_0, based on the Bethe–Heitler formalism reads:

$$1/X_0 = 4\,\alpha\ (N_\mathrm{A}/A)\ \left\{Z\,(Z+1)\,{r_\mathrm{e}}^2 \ln\left(183 Z^{-1/3}\right)\right\}\ \left[\mathrm{cm^2\ g^{-1}}\right] \tag{6.5}$$

The Z^2 term reflects the fact that the bremsstrahlung results from a coupling of the initial electron to the electromagnetic field of the nucleus, somewhat screened by the electrons (*log* term), and augmented by a direct contribution from the electrons (Z^2 replaced by $Z\,(Z+1)$).

The radiation length of a compound, or mixture, can be calculated using:

$$1/X_0 = \Sigma\ w_\mathrm{j}/X_\mathrm{j} \tag{6.6}$$

where the w_j are the fractions by weight of the nuclear species j of the mixture or of the compound.

The spectrum of photons with energy k radiated by an electron of energy E traversing a thin slab of material (expressed as a function of $y = k/E$) has the characteristic "bremsstrahlung" spectrum:

$$\mathrm{d}\sigma/\mathrm{d}k = A/\left(X_0 N_\mathrm{A} k\right) \cdot \left(4/3 – 4/3\mathrm{y} + \mathrm{y}^2\right). \tag{6.7}$$

At very high energies a number of effects, considered at the end of this subsection, modify the spectrum.

Another important quantity, the *critical energy* can be introduced examining Fig. 6.2. The critical energy E_c for *electrons* (or *positrons*) in a given medium is defined as the energy at which energy loss by radiation in a thin slab equals the energy loss by ionization. A slightly different definition ε_0, introduced by Rossi, results from considering the relative energy loss as fully independent of energy (see Fig. 6.2). The critical energy ε_0 is well described in dense materials (see Fig. 6.4) by:

$$\varepsilon_0 = 610\ \mathrm{MeV}/\left(Z + 1.24\right). \tag{6.8}$$

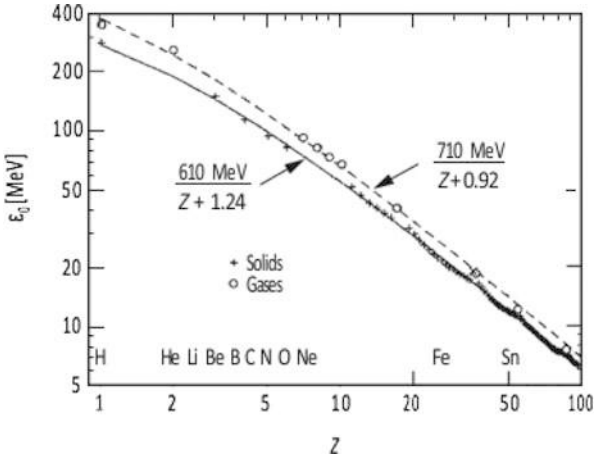


Fig. 6.4 Critical energy for the chemical elements, using Rossi's definition [6]. The fits shown are for solids and liquids (solid line) and gases (dashed line)

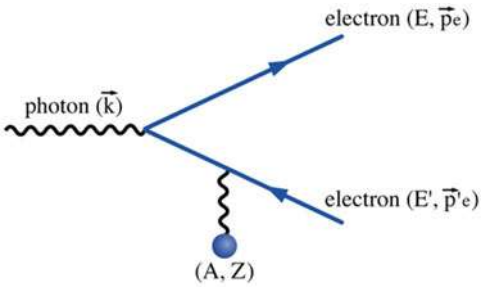


Fig. 6.5 Electron-positron pair creation in the field of a nucleus (A, Z)

As will be seen below, X_0 and E_c (or ε_0) are among the important parameters characterizing the formation of electromagnetic showers.

Several processes contribute to the interaction of photons with matter, the relative importance of which depends primarily on their energy.

Pair Production

This process is dominant as soon as photon energies are above a few times 2 $m_e c^2$. The graph responsible of the process (Fig. 6.5) shares the vertices of the bremsstrahlung graph.

The dominant part (Z^2) is due to the nucleus, while the electrons contribute proportionally to Z. The process of pair production has been studied in detail [7]. The pair production cross section can be written, in the complete screening limit at high energy as:

$$\mathrm{d}\sigma/\mathrm{d}x = A/(X_0 N_\mathrm{A}) \cdot (1\text{–}4/3x\ \ (1-x))\,, \tag{6.9}$$

where $x = E/k$ is the fraction of the photon energy k taken by the electron of the pair. Integrating the cross section over E gives the pair production cross section:

$$\sigma = 7/9\ A/(X_0 N_\mathrm{A})\,. \tag{6.10}$$

After 9/7 of an X_0, the probability that a high-energy photon survives without having materialized into an electron-positron pair is 1/e. In the pair production process the energy of the recoil nucleus is small, typically of the order of $m_e c^2$, implying that at high photon energy ($k >> m_e c^2$) the electron and the positron are both collinear with the incident photon. When the reaction takes place with an electron, the momentum transfer can be much higher leading to "triplets" with one positron and two electrons in the final state.

As for bremsstrahlung the cross section is affected at very high energy by processes considered later.

Compton Effect

The QED cross-section for the photon-electron scattering (Klein-Nishina [8]) can be written in the limit of $k >> m_e c^2$, using $x = k/m_e c^2$,

$$\sigma = \pi {r_e}^2\,(\log\ 2x + 1/2)/x\ \left[\mathrm{cm}^2\right]. \tag{6.11a}$$

The related probability for Compton scattering after the traversal of a material slab of thickness dt and mass per unit volume ρ is:

$$\phi = \sigma\rho \; N_{\mathrm{A}} \; Z/A \; \mathrm{d}t. \tag{6.11b}$$

For high Z (e.g. lead) the maximum of the Compton cross section and the pair production cross-section are of the same order of magnitude, while for lighter materials the maximum of the Compton cross section is higher. This is illustrated in Fig. 6.6 (from [1]) where carbon and lead are compared.

The differential Compton cross-section, with θ denoting the scattering angle between the initial and final photon, and η the angle between the vector perpen-

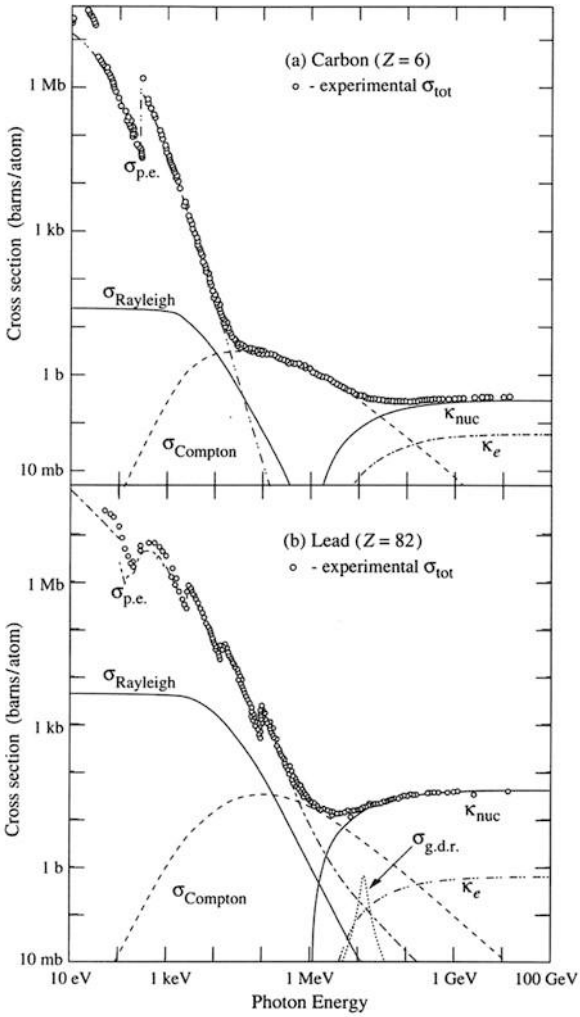


Fig. 6.6 Photon total cross section as a function of the photon energy in carbon and lead, with the contributions of different processes. $\sigma_{\mathrm{p.e.}}$ corresponds to the atomic photoelectric effect and κ_{nuc} (κ_{e}) corresponds to pair production in the nuclear (electron) field

dicular to the scattering plane and the polarization vector of the initial photon (in case it is linearly polarized) reads (ε being the ratio between the scattered and the incident photon energy $\varepsilon = 1/(1 + k/m_e c^2(1 - \cos\theta))$:

$$\mathrm{d}\sigma/\mathrm{d}\Omega = 0.5\, r_e^2 \left(\varepsilon + 1/\varepsilon - 2\sin^2\theta\cos^2\eta\right). \tag{6.12}$$

At low energy (k not larger than a few MeV), the η-dependence can be exploited for polarization measurements (Compton polarimetry). In the same energy range the probability of backward scattering is also sizeable.

Photoelectric Effect

For sufficiently low photon energies the atomic electrons can no longer be considered as free. The cross section for photon absorption, followed by electron emission (photoelectric effect) presents discontinuities whenever the photon energy crosses the electron binding energy of a deeper shell.

Explicit calculations [4] show that above the K-shell the cross section decreases like $E^{-3.5}$.

In the section devoted to shower formation, the relevance of the photoelectric effect will be considered. The coherent scattering (or Rayleigh scattering) is comparatively smaller than the photoelectric effect and its role negligible for shower formation.

High Energy Effects (LPM)

In the collinear approximation of bremsstrahlung, the longitudinal momentum difference $q_{||}$ between the initial electron (energy E) and the sum of the final electron and photon (energy k) is equal to

$$q_{||} = {m_e}^2 c^3 k/2E\,(E - k)\,. \tag{6.13}$$

This quantity can be extremely small, being for example 0.002 eV/c for a 25 GeV electron radiating a 10 MeV photon. Such a small longitudinal momentum transfer implies a large formation length, L_f ($L_f\ q_{||} \geq h/2\pi$), about 100 μm in the above example. Secondary interactions (like multiple scattering) taking place over this distance will perturb the final state and will in general diminish the bremsstrahlung cross section and the pair production cross section in case of photon interactions. Coherent interaction of the produced photons with the medium (dielectric effect) also affects, and reduces, the bremsstrahlung cross-section.

Such effects, already anticipated by Landau and Pomeranchuk [9] were considered in detail by several authors, and were measured by the experiment E146 at SLAC. A recent overview is given in [10]. The high k/E part of the bremsstrahlung spectrum is comparatively less affected (because of much larger $q_{||}$ values) while the low k/E part is significantly influenced for E above ~100 GeV, see Fig. 6.7. Only at much higher energies (>10 TeV) is the pair production cross-section affected.

In crystalline media the strong intercrystalline electrical fields may result in coherent suppression or enhancement of bremsstrahlung. Net effects depend on the

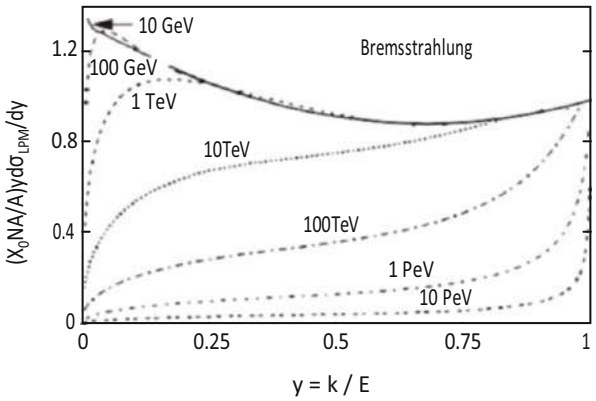


Fig. 6.7 Normalized Bremsstrahlung cross-section $k\,\mathrm{d}\sigma/\mathrm{d}k$ in lead as a function of the fraction of momentum taken by the radiated photon

propagation direction of the particle with respect to the principal axes of the crystal [11].

Hadronic Interactions of Photons

Photons with energies above a few GeV can behave similarly to Vector Mesons (ρ, ω and ϕ) with the same quantum numbers and in this way develop strong interactions with hadronic matter. They can be parameterized with the Vector Meson Dominance model. Using the Current-Field Identity [12], the amplitude for interactions of virtual photons γ^* of transverse momentum q is:

$$\begin{aligned}\mathcal{A}\left(\gamma^* \mathrm{A} \to \mathrm{B}\right) = \left(e/2\gamma_\rho\right)\ \mathrm{m}^2{}_\rho/\left(\mathrm{m}^2{}_\rho - q^2\right)\mathcal{A}\left(\rho \mathrm{A} \to \mathrm{B}\right) \\ +\text{equiv.terms for } \omega \text{ and } \phi \text{ mesons.}\end{aligned} \tag{6.14}$$

Various photo- and electro-production cross sections were calculated and confronted with experiment. As an example the ratio of hadron production to electron-positron pair production in the interaction of a 20 GeV photon is about 1/200 for hydrogen and 1/2500 for lead [13]. While this ratio is small, the effect on shower characteristics and on particle identification can in certain cases be significant [for example—see Ref. 14—when studying CP violating $\pi\pi$ final states in $\mathrm{K_L}$ decays, for which $\pi e \nu$ decays are a background source].

6.2.2 Electromagnetic Showers

When a high energy electron, positron or photon impinges on a thick absorber, it initiates an electromagnetic cascade as pair production, bremsstrahlung and Compton effects generate electrons/positrons and photons of lower energy. Electron/positron energies eventually fall below the critical energy, and then dissipate their energy

by ionization and excitation rather than by particle production. Photons propagate somewhat deeper into the material, being ultimately absorbed primarily via the photoelectric process.

Given the large number of particles (electrons, positrons, photons) present in a high energy electromagnetic cascade (more than one thousand for a 10 GeV electron or photon in lead), global variables have been sought to describe the average shower behaviour. Scale variables, such as X_0 as unit length, can be used to parameterize the radiation effects. However, since energy losses by $\mathrm{d}E/\mathrm{d}x$ and by radiation depend in a different way on material characteristics, one should not expect perfect 'scaling'.

Analytical Description

In an analytical description [6] a first simplification consists in 'factorizing' the longitudinal development and the lateral spread of showers, with the assumption that the lateral excursion of electrons and photons around the direction of the initial particle does not affect the longitudinal behaviour and in particular the 'total track length' (see below).

As for any statistical process the first goal is to obtain analytical expressions for average quantities. Particularly relevant (for a shower of initial energy E_0) are: $c(E_0,E,t)$ the average number of electrons plus positrons with energy between E and $E + \mathrm{d}E$ at depth t (expressed in radiation length), and the integral distribution $C(E_0,E,t) = \int_0^E c(E_0,E',t)\mathrm{d}E'$; $n(E_0,E,t)$ and $N(E_0,E,t)$ are the corresponding functions for photons.

Using the probability distribution functions of the physical effects driving the shower evolution (Bremsstrahlung, Compton, $\mathrm{d}E/\mathrm{d}x$, pair production) one can write and solve [15, 16] 'evolution equations' correlating $C(E_0,E,t)$ and $N(E_0,E,t)$. In the so called 'approximation B' of Rossi, the energy loss of electrons by $\mathrm{d}E/\mathrm{d}x$ is taken as constant, and the pair production and bremsstrahlung cross-sections are approximated by their asymptotic expression.

As an illustration, Fig. 6.8 shows the number of electrons and positrons as a function of depth, in a shower initiated by an electron of energy E_0, and by a photon of energy E_0 in units of the "Rossi critical energy ε_0" (see Sect. 6.2.1). These distributions are integrated over E from 0 to the maximum possible. The area under the curves is to a good approximation equal to E_0/ε_0, in accordance with the physical meaning of ε_0. The two sets of curves also show that a photon initiated shower is shifted on average by about 1 X_0 to larger depths compared to an electron (or positron) initiated one.

The total track length $TTL = \int_0^\infty C(E_0,0,t)\mathrm{d}t$ the energy transferred to the calorimeter medium by $\mathrm{d}E/\mathrm{d}x$, the source of the calorimeter signal.

Results from Monte Carlo Simulations

While analytical descriptions are useful guidelines, many applications require the use of Monte-Carlo (MC) simulations reproducing step by step, in a statistical manner, the physical effects governing the shower formation. For several decades, the standard simulation code for electromagnetic cascades has been EGS4 [17]. A recent alternative is encoded in the Geant4 framework [18].

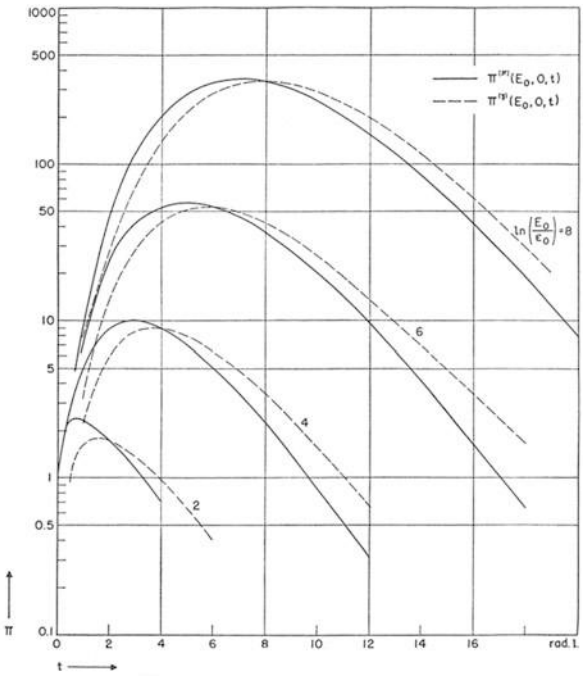


Fig. 6.8 Number of charged secondaries as a function of shower depth, for an electron initiated shower (full lines) and a photon initiated one (dashed lines), calculated analytically by Snyder [15, 16]. The numbers attached to each set of curves indicate $\ln(E_0/\varepsilon_0)$

As an illustration of the additional information obtained by this MC approach, Fig. 6.9 shows results of a 30 GeV electron shower simulation in iron ($E_c = 22$ MeV). The energy deposition per slab ($dt = 0.5X_0$) is shown as a histogram, with the fitted analytical function (see below) superimposed. This distribution is close, but not identical, to the distribution of electrons above a certain threshold (here taken as 1.5 MeV) crossing successive planes (right-hand scale): the energy deposition is slightly below the number of electrons at the beginning of the shower, and somewhat higher at the end. Multiple scattering (see below), affecting more the low energy shower tail, is one effect contributing to this discrepancy. The distribution of photons above the same threshold of 1.5 MeV is shifted to larger X_0 with respect to the electron distribution, reflecting the higher penetration power of photons already mentioned.

As a further illustration of the power of MC simulations, Fig. 6.10 displays longitudinal profiles of 10 GeV electron showers obtained by Geant4 simulation in lead, copper and aluminium. Since the dE/dx per X_0 is relatively more important in low Z material compared to high Z materials, one expects showers to penetrate more deeply in high Z materials, a fact born out by the simulations. Illustrating the energy dependence of shower parameters Fig. 6.11 displays shower energy deposition as a function of depth (shower profiles) for a range of incident electron energies (1 GeV to 1 TeV) in lead. The position of the shower maximum shows

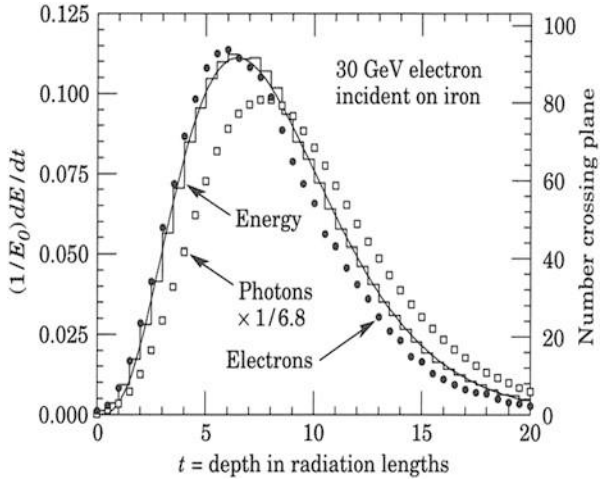


Fig. 6.9 EGS4 simulation of a 30 GeV electron-induced cascade in iron. The histogram is the fractional energy deposition per radiation length, and the curve is a gamma function fit to the distribution. The full (open) points represent the number of electrons (photons) with energy greater than 1.5 MeV crossing planes at $X_0/2$ intervals

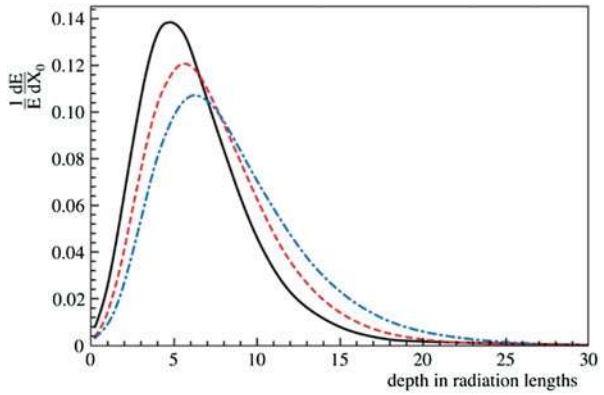


Fig. 6.10 Fractional energy deposition per longitudinal slice of 1 X_0 for 10 GeV electrons in aluminium (full line), copper (dashed) and lead (dash-dotted) (Geant4)

the expected logarithmic dependence on incident energy. In the parameterisation of shower profiles by Longo and Sestili [19].

$$F(E,t) = E_0\, b(bt)^{a-1}\, \mathrm{e}^{-bt}/\Gamma(a) \tag{6.15}$$

one finds accordingly $t_{\max} = (a-1)/b$, well fitted by $t_{\max} = \ln(y) + C_{\mathrm{i}}$, ($C_{\mathrm{i}} = 0.5$ for photons, -0.5 for electrons, and $y = E/E_{\mathrm{c}}$).

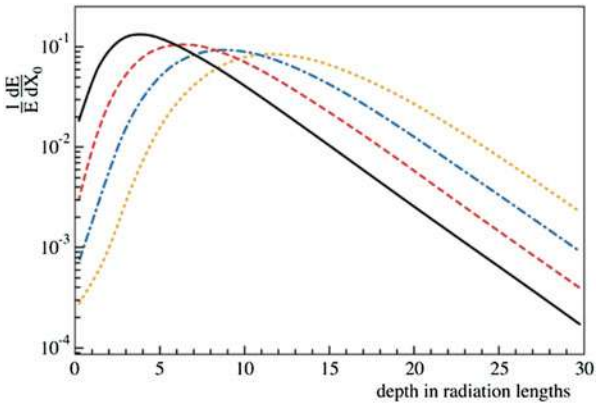


Fig. 6.11 Fractional energy deposition in lead, per longitudinal slice of 1 X_0, for electron induced showers of 1 GeV (full line), 10 GeV (dashed), 100 GeV (dash-dotted) and 1 TeV (dotted) (Geant4)

Finally, Fig. 6.12 illustrates the imbalance between electrons and positrons: in an electromagnetic shower, and rather material independent, about 75% of the energy deposited by charged particles is due to electrons, and 25% to positrons. This imbalance is due to the Compton and photoelectric effects which generate only electrons. It is more important towards the end of the shower.

Lateral Shower Development

Bremsstrahlung and pair creation on nuclei take place without appreciable momentum transfer to the (heavy) nuclei. Bremsstrahlung on electrons of the medium and Compton scattering involve however some momentum transfer. For example, in the Compton interaction of a 2 MeV (0.5 MeV) photon, 6% (16%) of the scattered photons are emitted with an angle larger than 90° with respect to the initial photon direction z. Another important effect contributing to the transverse spread in a cascade is multiple scattering of electrons and positrons.

After a displacement of length l along z, in a medium of radiation length X_0, the projected rms angular deviation along the transverse directions x and y, of an electron of momentum p is:

$$\theta_{x,y} = \frac{E_s}{\sqrt{2}} \frac{1}{p\beta c} \sqrt{l/X_0} \tag{6.16}$$

and the lateral displacement is

$$\delta_{x,y} = \frac{\theta_{x,y} l}{\sqrt{3}} \tag{6.17}$$

with $E_s = m_e c^2 \sqrt{(4\pi/\alpha)} = 21.2$ MeV. The lateral displacement contributes directly to the transverse shower broadening. If, after a step of length l, the electron emits

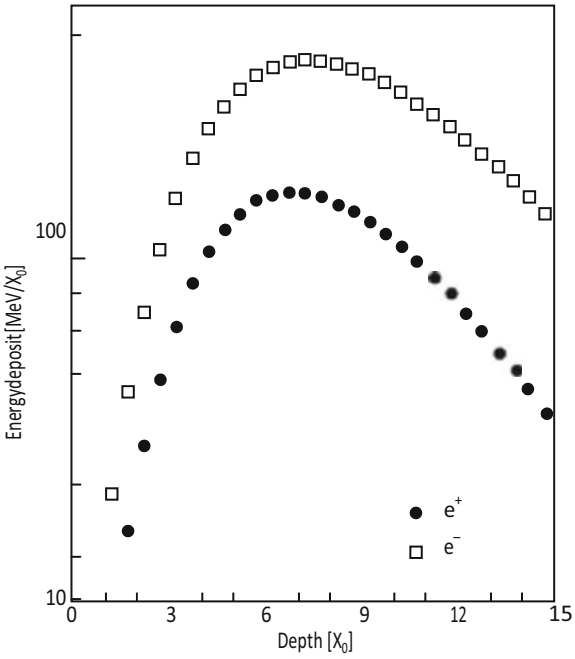


Fig. 6.12 Energy deposited in longitudinal slices of 1 X_0 by electrons (open symbols) and positrons (closed symbols) in a 10 GeV electron shower developing in lead (EGS4)

a bremsstrahlung photon, the emission will take place along the direction of the electron after l, thus at some angle (rms $\theta_{x,y}$ in both directions) with respect to the initial electron. Since the photon travels on average a considerable distance before materializing (9/7 X_0 if the photon is above a few MeV, significantly more at lower energy, see Fig. 6.6), the angular deviation of the electron gives a second, large contribution to the shower broadening.

In order to quantify the transverse shower spread, it is customary to use as parameter the Molière radius defined as:

$$\rho_{\mathrm{M}} = E_{\mathrm{s}}\, X_0 / E_{\mathrm{c}}, \tag{6.18}$$

where ρ_{M} equals $\sqrt{6}$ times the transverse displacement of an electron of energy E_{c}, after a path (without radiation nor energy loss) of 1 X_0. The most relevant physical meaning of ρ_{M} comes from Monte-Carlo simulations which show that about 87% (96%) of the energy deposited by electrons/positrons in a shower is contained in a cylinder of radius 1 (2) ρ_{M}.

Going back to the expressions of X_0 and E_{c}, it can be seen that their ratio is proportional to A/Z, and thus ρ_{M} is rather independent from the nuclear species, and is essentially governed by the material density. Calculations of ρ_{M}, for some pure materials and mixtures are reported in Table 6.1.

Table 6.1 Properties of calorimeter materials

Material	Z	Density [g cm^{-3}]	E_c [MeV]	X_0 [mm]	ρ_M [mm]	λ_{int} [mm]	$(dE/dx)_{mip}$ [MeV cm^{-1}]
C	6	2.27	83	188	48	381	3.95
Al	13	2.70	43	89	44	390	4.36
Fe	26	7.87	22	17.6	16.9	168	11.4
Cu	29	8.96	20	14.3	15.2	151	12.6
Sn	50	7.31	12	12.1	21.6	223	9.24
W	74	19.3	8.0	3.5	9.3	96	22.1
Pb	82	11.3	7.4	5.6	16.0	170	12.7
^{238}U	92	18.95	6.8	3.2	10.0	105	20.5
Concrete	–	2.5	55	107	41	400	4.28
Glass	–	2.23	51	127	53	438	3.78
Marble	–	2.93	56	96	36	362	4.77
Si	14	2.33	41	93.6	48	455	3.88
Ge	32	5.32	17	23	29	264	7.29
Ar (liquid)	18	1.40	37	140	80	837	2.13
Kr (liquid)	36	2.41	18	47	55	607	3.23
Polystyrene	–	1.032	94	424	96	795	2.00
Plexiglas	–	1.18	86	344	85	708	2.28
Quartz	–	2.32	51	117	49	428	3.94
Lead-glass	–	4.06	15	25.1	35	330	5.45
Air 20°, 1 atm	–	0.0012	87	304 m	74 m	747 m	0.0022
Water	–	1.00	83	361	92	849	1.99

mip minimum-ionizing particle

Comparing as an illustration lead and copper, one observes that the transverse dimensions of showers expressed in mm are essentially the same (because the transverse profiles are almost identical expressed in ρ_M (Fig. 6.14) and the ρ_M's are similar), while the shower in copper is (in mm) a factor 2.5 longer (because X_0 (copper) = 14.3 mm against 5.6 mm for lead).

On the other hand, despite being much shorter (in mm), the shower in lead contains about 2.5 more electrons (of lower energy in average) than the shower in copper, in the inverse proportion to their respective critical energies (7.4 MeV for lead against 20 MeV for Copper).

The lateral spread of showers is on average narrow at the beginning, where the shower content is still dominated by particles of energy much larger than E_c. In the low-energy tail the shower broadens. Monte Carlo simulations allow studying profiles at various depths. This is illustrated in Fig. 6.13 which shows the 90% containment radius as a function of the shower depth and in Fig. 6.14 which shows the radial profile of showers in three different materials. The broader width in the first 2 or 3 X_0 can be associated with backscattering (albedo) from the shower, which competes with the narrow core of the shower in its very early part. There is almost

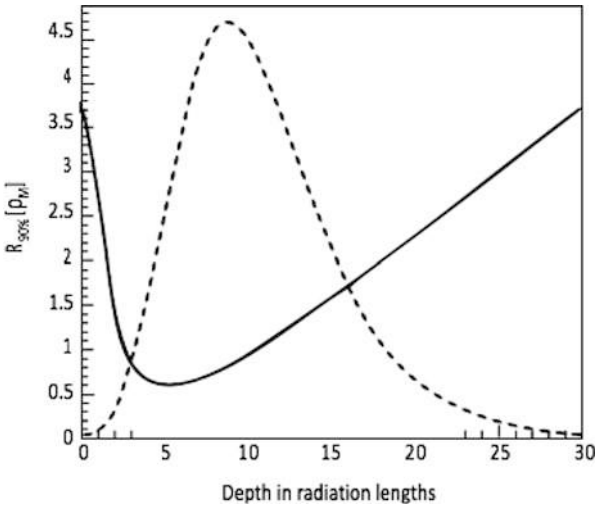


Fig. 6.13 90% containment radius $R_{90\%}$ (full line), in Molière radius ρ_M as a function of shower depth, for 100 GeV electron showers developing in lead. For comparison the longitudinal energy deposition is also shown (dashed line, arbitrary scale) (Geant4)

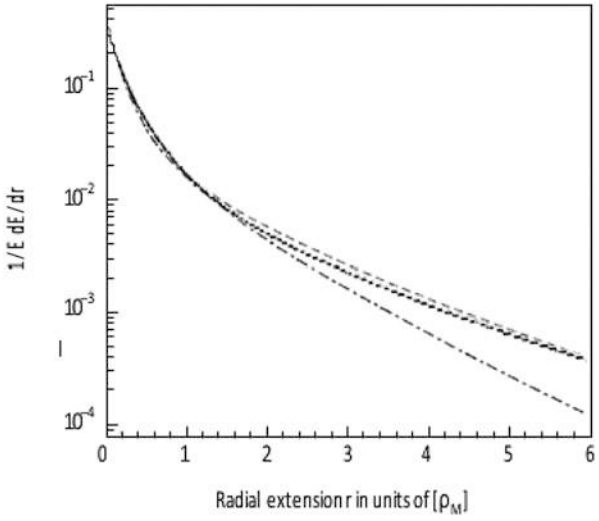


Fig. 6.14 Fractional energy deposition in cylindrical layers of thickness 0.1 ρ_M, coaxial with the incident particle direction, for 100 GeV electron-induced showers in aluminium (dotted line), copper (dashed line) and lead (dash-dotted) (Geant4)

no dependence of shower transverse profiles (integrated over depth) as a function of initial electron energy.

6.2.3 Homogeneous Calorimeters

For reasons explained later, large calorimeter systems are often 'sampling' calorimeters. These calorimeters are built as a stack of passive layers, in general of high Z material for electromagnetic calorimeters, alternating with layers of a sensitive medium responding to ('sampling' the) electrons/positrons of the shower, produced mostly in the passive layers.

A homogeneous calorimeter is built only from the sensitive medium. Provided all other conditions are satisfied (full containment of the shower, efficient collection and processing of the signal) homogeneous calorimeters give the best energy resolution, because sampling calorimeters are limited by 'sampling fluctuations' (see Sect. 6.2.4). It is instructive to study first the limitations in the "ideal" conditions of homogeneous calorimeters.

We first discuss low-energy applications, where the absorption does not involve showering. As an illustration, Fig. 6.15 shows the extremely narrow lines observed [20] when exposing a Germanium (Li-doped) crystal to a γ source of ^{108m}Ag and ^{110m}Ag. The resolution, at the level of one part in a thousand, is far better than obtained with NaI, a frequently used scintillating crystal (see below). Several

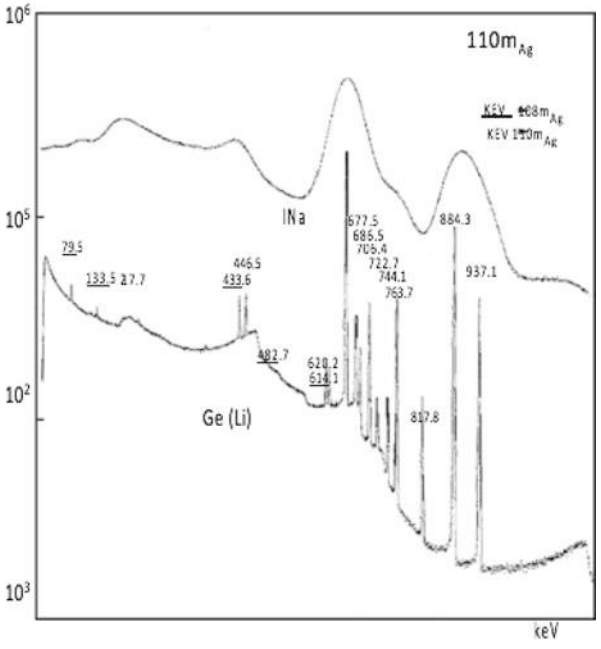


Fig. 6.15 Pulse height spectra recorded using a sodium iodide scintillator and a Ge (Li) detector. The source is a gamma radiation from the decay of ^{108m}Ag and ^{110m}Ag. Energies of peaks are labelled in keV

quantitative studies of the energy resolution of high purity Ge crystals, operated at low temperature (77 K) for γ spectroscopy have been made. A rather comprehensive discussion is given in [21]. After subtraction of the electronics noise, the width of the higher energy lines (above 0.5 MeV) is narrower than calculated assuming statistical independence of the created electron-hole pairs (~2.9 eV are needed to create such a pair). The reason for this was first understood by Fano [22]. Fundamentally it is due to the fact that the pairs created are not statistically independent, but are correlated by the constraint that the total energy loss must precisely be equal to the energy of the incident photon (in the limit of a device in which all energy losses lead to a detected signal, in a proportional way, the line width vanishes).

Calling σ the rms of the energy ε used to create an electron-hole pair, the actual resolution should be $\sigma/(\varepsilon\sqrt{Np})$, smaller than $1/\sqrt{Np}$ by a factor $\sqrt{F}$, where $F = (\sigma/\varepsilon)^2$ is the Fano factor. Monte-Carlo simulations [23] reproduce the phenomenon and give $F \sim 0.1$ for semiconductor devices, in reasonable agreement with measurements [21].

When two energy loss mechanisms compete, e.g. ionization and scintillation, the total energy constrain remains, but with a binomial sharing between the two mechanisms. It is thus expected that summing up the two contributions, assumed to be read out independently, will lead to an improved energy resolution (it should be remembered however that a certain fraction of the energy lost in the medium goes to heat.

This was first demonstrated with a liquid argon gridded cell exposed to La ions with an energy of 1.2 GeV/nucleon traversing the cell [24]. In this set-up both scintillation photons and electrons from electron-ion pairs were detected (see Sect. 6.3.3 for the collection mechanism). More recently, detailed studies of scintillation and ionization yields were made in liquid xenon using 662-keV γ-rays from a ^{137}Cs source [25]. With decreasing voltage applied to the sensitive liquid Xe volume, the scintillation signal increases while the ionization one decreases, as expected from recombination of electrons-ions giving rise to additional photons. The spectra obtained with scintillation alone, ionization alone, and their sum are shown in Fig. 6.16, together with the correlation between the two signals.

The ratio between scintillation and ionization depends also on the nature and energy of the particle making the deposit. Low energy nuclear recoils are highly ionizing, giving rise to more recombination and thus an increased light over charge ratio.

As discussed in Sect. 6.3.1, noble liquid detectors (using either argon or xenon) have been developed in the last decade which allowed pushing the limits of dark matter searches. They rely heavily on the existence of two correlated signals (ionization and scintillation) for a given energy deposit, exploiting in particular the ratio between the two to distinguish nuclear recoils from photon or muon background (see Sect. 6.7.2).

When the energy loss per unit length becomes very high (i.e. for low values of β and/or high values of the electric charge Ze for ions) saturation effects are observed in liquid ionization detectors, and also in scintillators. Empirically, the

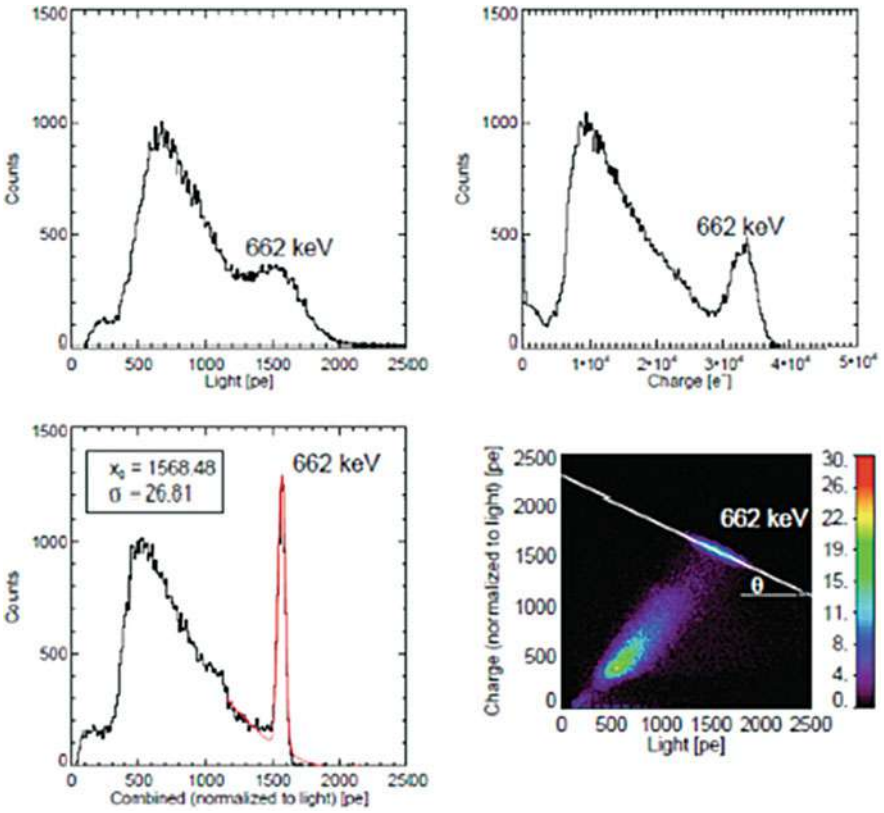


Fig. 6.16 Correlation between scintillation and ionization signals [25]. Scintillation alone (top-left), ionization alone (top-right), sum of both (bottom-left), 2-D correlation between scintillation and ionization (bottom-right)

effective scintillation (ionization) signal dL/dx (dI/dx) can be parameterized with "Birks law" [26]:

$$dL/dx = L_0.dE/dx/(1. + k_B\ dE/dx), \tag{6.19}$$

in which L_0 is the luminescence at low specific ionization density. The effect in plastic scintillators, for which k_B ~0.01 g.cm^{-2} MeV^{-1}, results in suppression ("quenching") of the light emission by the high density of ionized and excited molecules. Deviations from Birk's law have been observed for high Z ions [27].

In liquid ionization detectors the effect is associated with electron-ion recombination. It depends upon the electric field, in magnitude and direction with respect to the ionizing track. A typical value in liquid argon is k_B ~0.04 g.cm^{-2} MeV^{-1} for an electric field in a direction perpendicular to the track of 1 kV/cm, with k_B being inversely proportional to E, for E < 1 kV/cm [28].

Saturation effects are not relevant for electron or photon induced showers (at least below few TeVs) because the track density remains comparatively low (however, depending on the technique used for sampling calorimeters, internal amplification—like in calorimeters with gaseous readout—may saturate for high track density). Saturation effects do affect hadronic showers because of slow, highly ionizing fragments from nuclear break-up and slow proton recoils.

The—in general excellent—energy resolution of homogeneous calorimeters used for electromagnetic showers is affected by several instrumental effects. One of the most fundamental ones, the existence of a threshold energy E_{th} above which an electron of the shower does produce a signal will be illustrated in Sect. 6.3.2 when dealing with Cherenkov based electromagnetic calorimeters. Other effects include:

- longitudinal and transverse shower containment
- efficiency of light collection
- photoelectron statistics
- electron carrier attachment (impurities)
- space charge effects, . . .

These effects will be considered when dealing with examples where they are particularly relevant. The closer a detector approaches the intrinsic resolution—like for Ge crystals—the more important are the above limitations. In practice, large calorimeter systems for high energy showers based on homogeneous semiconductor crystals are unaffordable. Scintillating crystals and pure noble liquids are the best compromise between performance and cost, but do suffer from other limitations, as illustrated in examples given below.

6.2.4 Sampling Calorimeters and Sampling Fluctuations

In the simplest geometry, a sampling calorimeter consists of plates of dense, passive material alternating with layers of sensitive material.

For electromagnetic showers, passive materials with low critical energy (thus high Z) are used, thus maximizing the number of electrons and positrons in a shower to be sampled by the active layers. In practice, lead is most frequently used. Uranium has also been used to optimize the response towards hadrons (Sect. 6.2.7), and tungsten has been used in cases where compactness is a premium.

The thickness t of the passive layers (in units of X_0) determines the sampling frequency, i.e. the number of times a high energy electron or photon shower is 'sampled'. Intuitively, the thinner the passive layer (i.e. the higher the sampling frequency), the better the resolution should be. The thickness u of the active layer is usually characterized by the *sampling fraction* f_S which is the ratio of dE/dx of a

minimum ionizing particle in the active layer to the sum of $\mathrm{d}E/\mathrm{d}x$ in the active and passive layers:

$$f_S = u\ \mathrm{d}E/\mathrm{d}x_{\mathrm{active}}/\left(u\ \mathrm{d}E/\mathrm{d}x_{\mathrm{active}} + t\ \mathrm{d}E/\mathrm{d}x_{\mathrm{passive}}\right)\left[u, t \text{ in g cm}^{-2}, \mathrm{d}E/\mathrm{d}x \text{ in MeV/g cm}^{-2}\right]. \tag{6.20}$$

This 'sampling' of the energy results in a loss of information and hence in additional 'sampling fluctuations'. An approximation [29, 30] for these fluctuations in electromagnetic calorimeters can be derived using the total track length (*TTL*) of a shower initiated by an electron or photon of energy E. The signal is approximated by the number N_x of e^+ or e^- traversing the active signal planes, spaced by a distance $(t + u)$. This number N_x of crossings is

$$N_x = TTL/(t+u) = E/(\varepsilon_0\ (t+u)) = E/\Delta E,$$

ΔE being the energy loss in a unit cell of thickness $(t + u)$. Assuming statistical independence of the crossings, the fluctuations in N_x represent the 'sampling fluctuations' $\sigma(E)_{\mathrm{samp}}$,

$$\begin{aligned}\sigma(E)_{\mathrm{samp}}/E &= \sigma\left(N_x\right)/N_x = 1/\sqrt{N_x} = \sqrt{\{\Delta E\ (\mathrm{GeV})\ /E\ (\mathrm{GeV})\}} \\ &= 0.032\sqrt{\{\Delta \mathrm{E}\ (\mathrm{MeV})\ /\mathrm{E}\ (\mathrm{GeV})\}} = a/\sqrt{E}.\end{aligned} \tag{6.21}$$

The detector dependent constant a is the 'sampling term' of the energy resolution (see also below). For illustration, for a lead/scintillator calorimeter with 1.4 mm lead plates, interleaved with 2 mm scintillator planes, $\Delta E = 2.2$ MeV, one estimates $a \sim 5\%$ for 1 GeV electromagnetic showers. This represents a lower limit (the experimental value is closer to 7 to 8%), as threshold effects in signal emission and angular spread of electrons around the shower axis worsen the resolution [29]. In addition, a large fraction of the shower particles are produced as e^+e^- pairs, reducing the number of statistically independence crossings N_x.

The sampling fraction f_S has practical consequences, considering the actual signal produced by the calorimeter. If f_S is too small, the signal is small and may be affected by electronics noise and possibly other technical limitations due to the chosen readout technique (see below).

The dominant part of the calorimeter signal is actually not produced by minimum ionizing particles, but rather by the low-energy electrons and positrons crossing the signal planes. Defining the fractional response f_R of a given layer "i" as the ratio of energies lost in the active layer and of the sum of active plus passive layers one has

$$f_R^i = E^i_{\mathrm{active}}/(E^i_{\mathrm{active}} + E^i_{\mathrm{passive}}) \tag{6.22}$$

with the constraint that $\Sigma^i\ (E^{\,i}{}_{\mathrm{active}} + E^{\,i}{}_{\mathrm{passive}}) = E$.

Experimentally one finds that f_R (taking all layers together) is significantly smaller than f_S [31]. The ratio f_R/f_S, usually called '*e/mip*' for obvious reasons, can

be as low as 0.6 when the Z of the passive material (lead) is much larger than the Z of the active one (plastic scintillator, liquid argon). This effect, well reproduced by Monte-Carlo simulations, is to some extent due to the "transition effect" between the passive and active material, but also due to the fact that a significant fraction of electrons produced in the high Z passive material by pair production or Compton scattering do not have enough energy to exit this layer and are thus not sampled. This same effect induces a depth dependence of *e/mip*, which decreases by few percent towards the end of the shower.

Taking into account an energy independent contribution from electronics noise b, and a minimum asymptotic value of the relative energy resolution c (constant term, due for example to inhomogeneities in materials, imperfection of calibrations, ...) the energy resolution of a sampling calorimeter is in general written as[1]

$$\Delta E/E = a/\sqrt{E} \oplus b/E \oplus c \tag{6.23}$$

Experimentally it has been observed that the same relation holds also for homogeneous calorimeters, in general with smaller 'sampling terms' a, although their origin is not coming from sampling fluctuations, but from other limitations (see Sect. 6.2.3).

6.2.5 *Physics of the Hadronic Cascade*

By analogy with electromagnetic showers, the energy degradation of high-energy hadrons proceeds through an increasing number of (mostly) strong interactions with the calorimeter material. However, the complex hadronic and nuclear processes produce a multitude of effects that determine the performance of practical instruments, making hadronic calorimeters more complicated instruments to optimize and resulting in a significantly worse intrinsic resolution compared to the electromagnetic one. Experimental studies by many groups helped to unravel these effects and permitted the design of high-performance hadron calorimeters.

The hadronic interaction produces two classes of secondary processes. First, energetic secondary hadrons are produced with momenta typically a fair fraction of the primary hadron momentum, i.e. at the GeV scale. Second, in hadronic collisions with the material nuclei, a significant part of the primary energy is consumed by nuclear processes such as excitation, nucleon evaporation, spallation, etc., generating particles with energies characteristic of the nuclear MeV scale.

The complexity of the physics is illustrated in Fig. 6.17, which shows the energy spectra of the major shower components (weighted by their track length in the shower) averaged over many cascades, induced by 100 GeV protons in lead. These spectra are dominated by electrons, positrons, photons, and neutrons at low energy.

[1] In a formula like (6.23), $a \oplus b$ means $\sqrt{(a^2 + b^2)}$.

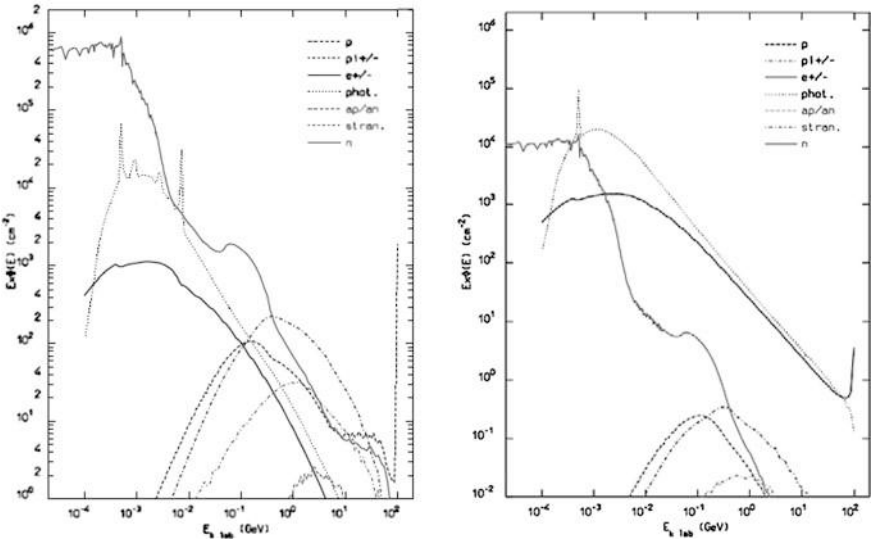


Fig. 6.17 Particle spectra produced in the hadronic cascade initiated by 100 GeV protons absorbed in lead (left). The energetic component is dominated by pions, whereas the soft spectra are composed of photons and neutrons. The ordinate is in 'lethargic' units and represents the particle track length, differential in log E. The integral of each curve gives the relative fluence of the particle [32]. On the right, same figure for 100 GeV electrons in lead, showing the much simpler structure, dominated by electrons and photons (hadrons are down by more than a factor 100)

The structure in the photon spectrum at approximately 8 MeV reflects a (n,γ) reaction and is a fingerprint of nuclear physics; the line at 511 keV results from e^+e^- annihilation photons. These low-energy spectra encapsulate all the information relevant to the hadronic energy measurement. Deciphering this message becomes the story of hadronic calorimetry.

The energetic component contains protons, neutrons, charged pions and photons from neutral pion decays. Due to the charge independence of hadronic interactions, on average approximately one third of the pions produced will be π^0s, $f_{\pi 0} \approx 1/3$. These neutral pions will decay to two photons, $\pi^0 \rightarrow \gamma\gamma$, before reinteracting hadronically and will induce an electromagnetic cascade, proceeding along its own laws of electromagnetic interactions (see Sect. 6.2.2). This physics process acts like a 'one-way diode', transferring energy from the hadronic part to the electromagnetic component, which will not contribute further to hadronic processes.

As the number of energetic hadronic interactions increases with increasing incident energy, so will the fraction of the electromagnetic cascade. This simple picture of the hadronic showering process leads to a power law dependence of the two components [33, 34]; naively, the electromagnetic component is $F_{\mathrm{em}} = 1 - (1 - f_{\pi 0})^n$, n denoting the number of shower generations induced by a particle with energy E. For the hadronic fraction F_{h} one finds in a more realistic evaluation $F_{\mathrm{h}} = (E/E_0)^k$. The parameter k expresses the energy dependence and is related to the average

multiplicity m of a collision, with $k = \ln (1 - f_{\pi 0})/\ln m$. The parameter E_0 denotes the average energy necessary for the production of a pion, approximately $E_0 \approx 2$ GeV; with the multiplicity $m \approx 6$–7 of hadrons produced in a hadronic collision k is ≈ -0.2. Values of F_{h} are of order 0.5 (0.3) for 100 (1000) GeV showers. As the energy of the incident hadron increases, it is doomed to dissipate its energy in a flash of photons. Were one to extrapolate this power law to the highest particles energies detected calorimetrically, $E \leq 10^{20}$ eV more than 98% of the hadronic energy would be converted to electromagnetic energy!

The low-energy nuclear part of the hadronic cascade has very different properties, but carries the dominant part of the energy in the hadronic sector. In the energetic hadron collisions with the nuclei of the calorimeter material, their nucleons will be struck initiating an 'intra-nuclear' cascade. In the subsequent steps, the intermediate nucleus will de-excite, in general through a spallation reaction, evaporating a considerable number of nucleons, accompanied by few MeV γ-emission. The binding energy of these nucleons released in these collisions is taken from the energy of the incident hadron. The number of these low-energy neutrons is large: ~ 20 neutron/GeV in lead. The fraction of the total associated binding energy depends on the incident energy and may be as high as ~20–40%. These neutrons will ultimately be captured by the target nuclei, resulting in delayed nuclear photon emission (at the ~ μs timescale). The energy lost to binding energy is therefore, in general, not detected ('invisible') in practical calorimeters.

In Fig. 6.18 the energy dependence of the electromagnetic, fast hadron and nuclear components is shown. The response of a calorimeter is determined by the sum of the responses to these different components which react with the passive and

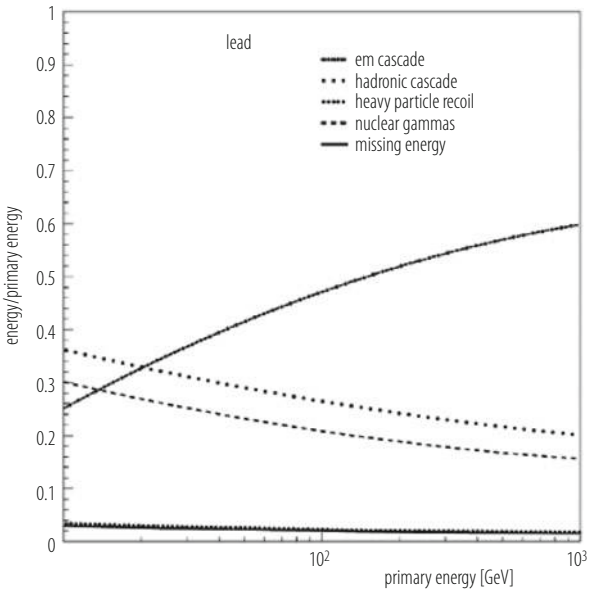


Fig. 6.18 Characteristic components of proton-initiated cascades in lead. With increasing energy the em component increases [32]

active parts of the calorimeter in their specific ways (see Sect. 6.2.7). Contributions from neutrons and photons from nuclear reactions, which have consequences for the performance of these instruments, are also shown in Fig. 6.18. The total energy carried by photons from nuclear reactions is substantial: only a fraction, however, will be recorded in practical instruments, as most of these photons are emitted with a considerable time delay (~1 μs). The event-by-event fluctuations in the invisible energy dominate the fluctuations in the detector signal, and hence the energy resolution. The road to high-performance hadronic calorimetry has been opened by understanding how to compensate for these invisible energy fluctuations [35].

6.2.6 *Hadronic Shower Profile*

The total cross section for hadrons is only weakly energy dependent in the range of few to several hundred GeV, relevant for calorimetry. For protons, the total pp. cross section σ_{tot} is approximately 39 mb. For pion-proton collisions $\sigma_{\text{tot}}(\pi\text{p}) = 2/3\ \sigma_{\text{tot}}(\text{pp})$ is naively expected, i.e. 26 mb, compared to the measured value of $\sigma_{\text{tot}}(\pi^{+}\text{p}) \approx 23$ mb. For hadronic calorimetry the inelastic cross sections, $\sigma_{\text{inel}}(\text{pA})$ or $\sigma_{\text{inel}}(\pi\text{A})$, determine the value of the corresponding interaction length, $\lambda_{\text{int}} = A/N_{\text{A}}\sigma_{\text{inel}}(\text{hadron, A})$. On geometrical grounds $\sigma_{\text{inel}}(\text{hadron, A})$ is expected to scale as $A^{2/3}\sigma_{\text{inel}}(\text{hadron, p})$, close to the measured approximate scaling $A^{0.71}\sigma_{\text{inel}}(\text{hadron, p})$ and therefore $\lambda_{\text{int}} \approx A^{0.29}/\{N_{\text{A}}\sigma_{\text{inel}}(\text{hadron, p})\}$ [g cm^{-2}].

This characteristic length λ_{int} is the mean free path of high energy hadrons between hadronic collisions and sets the scale for the longitudinal hadronic shower profile. The probability $P(z)$ for a hadron traversing a distance z without undergoing an interaction is therefore $P(z) = \exp.\ (-z/\lambda_{\text{int}})$. The equivalence with the characteristic distance X_0 for the electromagnetic cascade is evident. In analogy to the parameterization of electromagnetic showers the longitudinal profile of hadronic showers can be parameterized in the form.

$$\mathrm{d}E/\mathrm{d}x = c\left\{w\{x/X_0\}^{\alpha-1}\exp\left(-bx/X_0\right) + (1-w)\,(x/\lambda)^{\alpha-1}\exp\left(-dx/\lambda\right)\right\}. \tag{6.24}$$

The overall normalization is given by c; α, b, d, w are free parameters and x denotes the distance from the shower origin [36].

Longitudinal pion-induced shower profiles are shown in Fig. 6.19 for different energies together with the analytical shower fits. The longitudinal energy deposit rises to a maximum, followed by a slow decrease due to the predominantly low-energy, neutron-rich part of the cascade. Proton-induced showers show a slightly different longitudinal shape due to the differences in the first few initial collisions. Shower profiles in different materials, when expressed as a function of λ_{int} exhibit approximate scaling in λ_{int}, in analogy to approximate scaling of electromagnetic

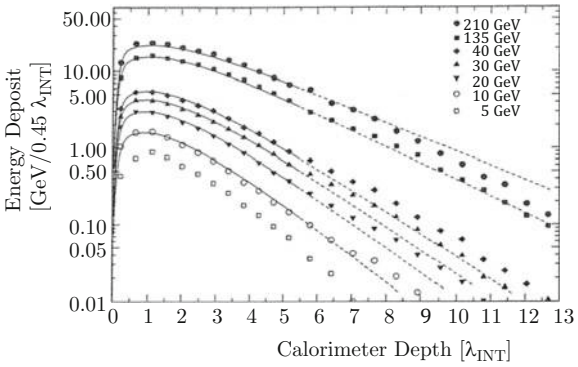


Fig. 6.19 Measured longitudinal shower distributions for pions at three energies together with the shower parameterization [37]

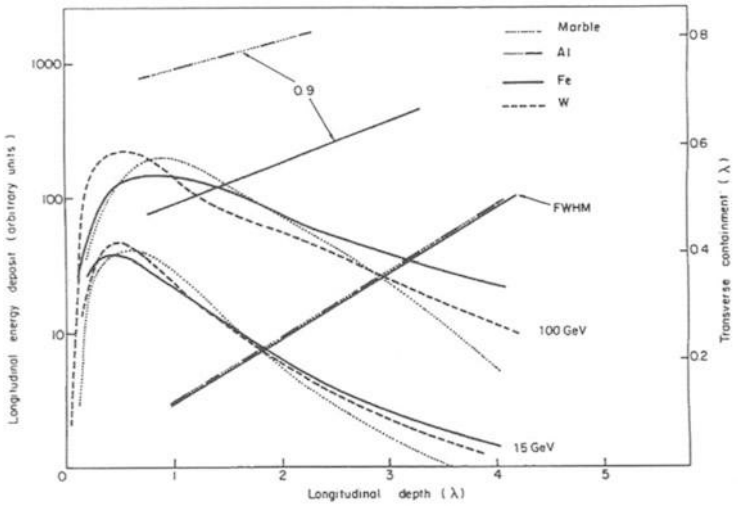


Fig. 6.20 Longitudinal shower development induced by hadrons in different materials, showing approximate scaling in λ. The shower distributions are measured with respect to the face of the calorimeter (left ordinate). The transverse distributions as a function of shower depth show scaling in λ for the narrow core. The 90% containment radius is much larger and does not scale with λ (right ordinate) [30]

showers in X_0, see Fig. 6.20. Also shown are the transverse shower distributions: the relatively narrow core is dominated by the high-energy (mostly electromagnetic) component. The tails in the radial distributions are due to the soft, neutron-rich, component. In Fig. 6.21 the fractional containment as a function of energy is shown, exhibiting approximately the expected logarithmic energy dependence for a given containment [38, 39].

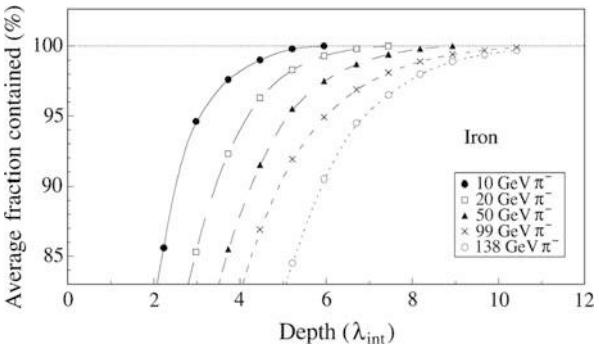


Fig. 6.21 Measured average fractional containment in iron of infinite transverse dimension as a function of thickness and various pion energies [38, 39]

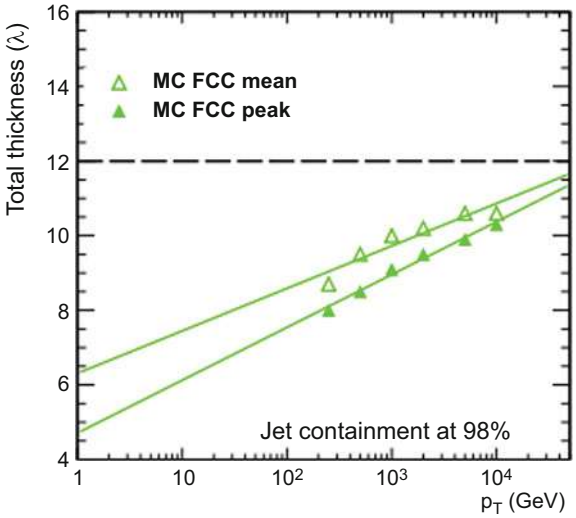


Fig. 6.22 Total thickness, expressed in λ, to contain up to 98% of a jet as a function of the jet transverse momentum. Mean and peak refer to different statistical measures of containment [40]

These results indicate that for 98% containment at the 100 GeV scale a calorimeter depth of 9 λ_{int} is required. At the LHC, where single particles energies in the multi-hundred GeV and jets in the multi-TeV range have to be well measured, the hadrons are typically measured in 10 λ_{int}. For the next jump in collider energy, as is presently studied e.g. for "Future Circular Collider, FCC", particle and jet energies are approximately a factor 10 higher. For adequate containment, i.e. at the 98% level, calorimeter systems with ~12 λ_{int} will be required, see Fig. 6.22 [40]

6.2.7 Energy Resolution of Hadron Calorimeters

The average properties of the hadronic cascade are a reflection of the intrinsic event-by-event fluctuations which determine the energy resolution. Most importantly, fluctuations in the hadronic component are correlated with the number of spallation neutrons and (delayed) nuclear photons and hence with the energy consumed to overcome the binding energy; these particles from the nuclear reactions will contribute differently (in general less) to the measurable signal.

Let η_e be the efficiency for observing a signal $E^e{}_{vis}$ (visible energy) from an electromagnetic shower, i.e., $E^e{}_{vis} = \eta_e\ E(em)$; let η_h be the corresponding efficiency for purely hadronic energy to give a measurable signal in an instrument. Decomposing a hadron-induced shower into the em fraction F_{em} and a purely hadronic part F_h the measured, 'visible' energy $E^\pi{}_{vis}$ for a pion-induced shower is.

$$E^\pi{}_{vis} = \eta_e F_{em} E + \eta_h F_h E = \eta_e (F_{em} + \eta_h/\eta_e F_h)\ E, \tag{6.25}$$

where E is the incident pion energy. The ratio of observable signals induced by electromagnetic and hadronic showers, usually denoted 'e/π', is therefore

$$E^\pi{}_{vis}/E^e{}_{vis} = (e/\pi)^{-1} = F_{em} + \eta_h/\eta_e F_h = 1 + (\eta_h/\eta_e - 1)\ F_h. \tag{6.26}$$

In general $\eta_e \neq \eta_h$: in this case, the average response of a hadron calorimeter as a function of energy will not be linear because F_h decreases with incident energy. More subtly, for $\eta_h \neq \eta_e$, event-by-event fluctuations in the F_h and F_{em} components produce event-by-event signal fluctuations and impact the energy resolution of such instruments. The relative response 'e/π 'turns out to be the most important yardstick for gauging the performance of a hadronic calorimeter.

A convenient (albeit non-trivial) reference scale for the calorimeter response is the signal from minimum-ionizing particles (*mip*) which in practice might be an energetic through going muon, rescaled to the energy loss of a mip. Let *e/mip* be the signal produced by an electron relative to a mip. Assume the case of a mip depositing e.g. α GeV in a given calorimeter. If an electron depositing β GeV produces a signal β/α, the instrument is characterized by a ratio $e/mip = 1$. Similarly, the relative response to the purely hadronic component of the hadron shower is $\eta_h F_h E/mip$, or *h/mip* which can be decomposed into $h/mip = (f_{ion}\ ion/mip + f_n\ n/mip + f_\gamma\ \gamma/mip)$, with f_{ion}, f_n, f_γ denoting the average fractions of ionizing particles, neutrons and nuclear photons.

Practical hadron calorimeters are usually built as sampling devices; the energy sampled in the active layers, f_S (Eq. 6.20), is typically a small fraction, a few percent or less, of the total incident energy. The energetic hadrons lose relatively little energy ($\leq$10%) through ionization before being degraded to such low energies that nuclear processes dominate. Therefore, the response of the calorimeter will be

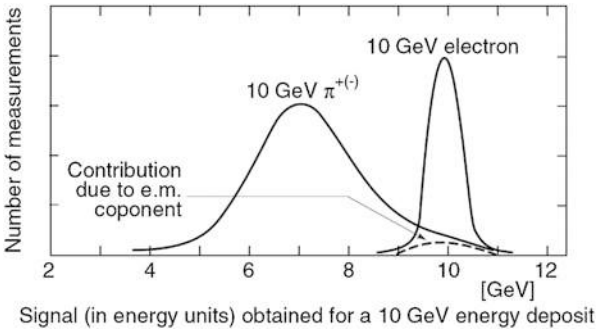


Fig. 6.23 Conceptual response of a calorimeter to electrons and hadrons. The curves are for a 'typical' sampling calorimeter with electromagnetic resolution of $\sigma/E = 0.1/\sqrt{E}(\text{GeV})$, with hadronic resolution of $\sigma/E = 0.5/\sqrt{E}(\text{GeV})$ and $e/\pi = 1.4$. The hadron-induced cascade fluctuates between almost completely electro-magnetic and almost completely hadronic energy deposit, broadening the response and producing non-Gaussian tails

strongly influenced by the values of *n/mip* and γ*/mip* in both the absorber and the readout materials.

This simple analysis already provides the following qualitative conclusions for instruments with $e/\pi \neq 1$, as shown conceptually in Fig. 6.23:

- fluctuations in $F_{\pi 0}$ are a major contribution to the energy resolution;
- the average value (F_{em}) increases with energy: such calorimeters have a non-linear energy response to hadrons;
- these fluctuations are non-Gaussian and therefore the energy resolution scales weaker than $1/\sqrt{E}$.

This understanding of the impact of shower fluctuations suggests to 'tune' the e/π response of a calorimeter in the quest for achieving e/π = 1, and thus optimizing the performance [41, 42].

It is instructive to analyze n/mip, because of the richness and intricacies of n-induced nuclear reactions and the very large number of neutrons with $E_n < 20$ MeV. In addition to elastic scattering a variety of processes take place in high-Z materials such as (n, n'), (n, 2n), (n, 3n), (n, fission). The ultimate fate of neutrons with energies $E_n < 1$–2 MeV is dominated by elastic scattering; cross-sections are large (~ barns) and mean free paths short (a few centimetres); the energy loss is ~1/A (target) and hence small. Once thermalized, a neutron will be captured, accompanied by γ-emission.

This abundance of neutrons gives a privileged role to hydrogen, which may be present in the readout material. In an n-p scatter, on average, half of the neutron kinetic energy is transferred. The recoil proton, if produced in the active material, contributes directly to the calorimeter signal, i.e., is not sampled like a mip (a 1 MeV proton has a range of ~20 μm in scintillator). The second important n-reaction is the production of excitation photons through the (n,n',γ) reaction [42].

This difference in neutron response between high-Z absorbers and hydrogen-containing readout materials has an important consequence. Consider the contributions of *n/mip* as a function of the sampling fraction f_S. The mip signal will be inversely proportional to the thickness of the absorber plates, whereas the signal from proton recoils will not be affected by changing f_S: the *n/mip* signal will increase with decreasing f_S. Changing the sampling fraction allows to alter, to 'tune' e/π. Tuning of the ratio R_d = passive material [mm]/active material [mm] is a powerful tool for acting on e/π [41]. This approach works well for high-Z absorbers with a relatively large fission cross section, accompanied by multiple neutron emission. Optimized ratios tend to imply for practical scintillator thicknesses rather thick absorbers with concomitant significant sampling fluctuations and reduced signals.

How tightly are the various *fluctuating* contributions to the invisible energy correlated with the *average* behaviour, as measured by e/π? A quantitative answer needs rather complete shower and signal simulations and confirmation by measurement. Two examples are shown in Fig. 6.24. One observes a significant reduction in the fluctuations and an intrinsic hadronic energy resolution of $\sigma/E \approx 0.2/\sqrt{E(\text{GeV})}$ for instruments with $e/\pi \approx 1$ [39, 41, 42]. The intrinsic

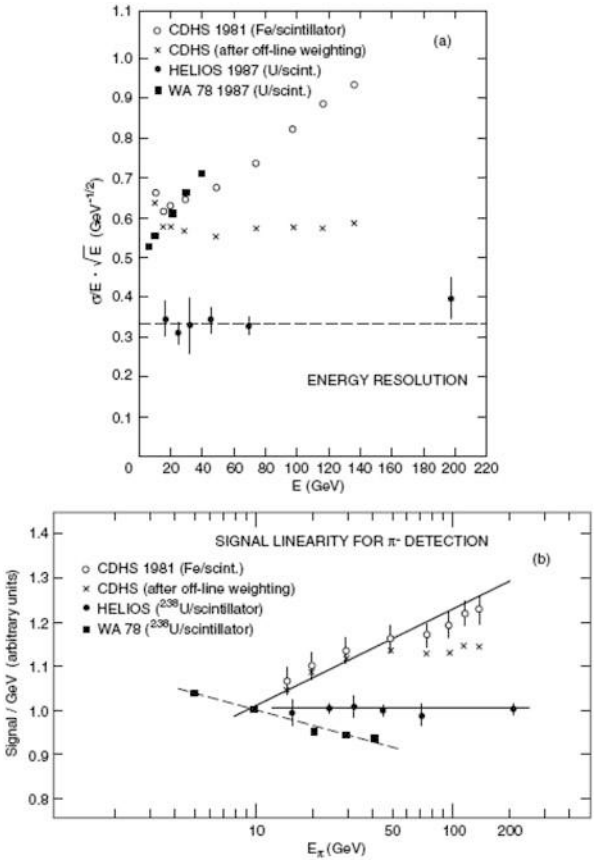


Fig. 6.24 Experimental observation of the consequences of $e/\pi \neq 1$. Shown is the measured pion response in under-compensating, compensating and over-compensating calorimeters; (**a**) energy resolution $\sigma/E\ \sqrt{E}$ as a function of the pion energy, showing deviations from scaling for non-compensating devices. (**b**) Signal per GeV as a function of pion energy, exhibiting signal non-linearity for non-compensating detectors [41]

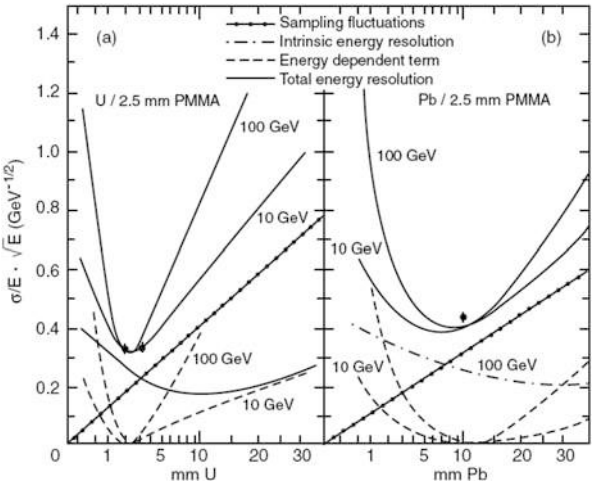


Fig. 6.25 Contributions to and total energy resolution of 10 and 100 GeV hadrons in scintillator calorimeters as a function of thickness of (**a**) uranium plates and (**b**) lead plates. The scintillator thickness is 2.5 mm in both cases. The dots in the curves are measured resolution values of actual calorimeters [42]

hadron resolution of a lead-scintillator sampling calorimeter may even be as good as $\sigma/E < \approx 0.13/\sqrt{E}(\text{GeV})$ [43].

Detectors achieving compensation for the loss of non-detectable ('invisible') energy, i.e., $e/\pi = 1$, are called 'compensated' calorimeters.

There are several further negative consequences if $e/\pi \neq 1$ in addition to reduced resolution. The energy resolution which no longer scales with $1/\sqrt{E}$, is usually parameterized as $\sigma/E = a_1/\sqrt{\text{E}} \oplus a_2$, where a 'constant' term a_2 is added quadratically, even though physics arguments suggest $a_2 = a_2(E)$. Since the fraction of π^0-production $F_{\pi 0}$ increases with energy, such calorimeters have a non-linear energy response. Furthermore, given that the average hadronic fraction F_h are different for pions ($F_\text{h}(\pi)$) and protons (neutrons) ($F_\text{h}(\text{p})$), typically $F_\text{h}(\pi) \sim 0.85 F_\text{h}(\text{p})$, the response in calorimeters with $e/\pi \neq 1$ depends on the hadron species [42].

The effects of e/π have been observed [41] (Fig. 6.24) and evaluated quantitatively [42]. Measurements and Monte Carlo simulations of the response of various calorimeter configurations are shown in Figs. 6.25 and 6.26.

Besides achieving "intrinsic compensation" with $e/\pi = 1$, effective compensation can be achieved by recognizing event by event independently the em fraction F_em and the hadronic fraction F_h, respectively. In instruments with a fine-grained longitudinal and lateral subdivision the different em and hadronic shower shapes provide an approximately independent determination of the two components and the basis for their off-line weighting, resulting in an effective $e/\pi = 1$ (see Sect. 6.7.5). Alternatively, the em component and the hadronic component in the shower may be measured independently with a dual readout: one active medium is only sensitive to Cherenkov radiation, predominantly caused by the em component, while the charged particles are measured e.g. with a scintillator, see Sect. 6.3.3.

To complete the analysis of the contributions to the energy resolution we need to consider sampling fluctuations, assuming fully contained showers and no degradation due to energy leakage. For electro-magnetic calorimeters a simple expla-

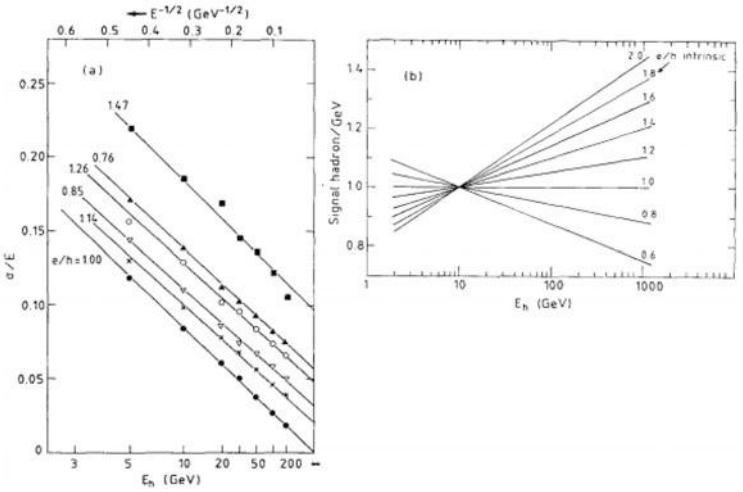


Fig. 6.26 Monte Carlo simulation of the effects of $e/\pi \neq 1$ on energy resolution (**a**) and linearity (**b**) of hadron calorimeters [42]

nation and an empirical parameterization holds (Eq. 6.21): $\sigma_{\text{samp}}(\text{em})/E = c(\text{em}) \cdot (\Delta E(\text{MeV})/E(\text{GeV}))^{1/2}$, where ΔE is the energy lost in one sampling cell and $c(\text{em}) \approx 0.05$ to 0.06 for typical absorber and readout combinations.

Similar arguments apply for the hadronic cascade; empirically, one has observed [30, 43] that.

$$\sigma_{\text{samp}}\,(\text{h})\,/E = c\,(\text{h}) \cdot (\Delta E\,(\text{MeV})\,/E\,(\text{GeV}))^{1/2} \text{ with } c\,(\text{h}) \approx 0.10. \tag{6.27}$$

For high-performance hadron calorimetry sampling fluctuations cannot be neglected.

The foundations of modern, optimized hadron calorimetry can be summarized as follows:

- the key performance parameter is $e/\pi = 1$, which guarantees linearity, $E^{-1/2}$ scaling of the energy resolution, and best intrinsic resolution;
- by proper choice of type and thickness of active and passive materials the response can be tuned to obtain (or approach) $e/\pi \sim 1$;
- the intrinsic resolution in practical hadron calorimeters can be as good as $(\sigma/E) \cdot \sqrt{E} < \sim 0.2$;
- sampling fluctuations contribute at the level of $\sigma/E \approx 0.10\ (\Delta E(\text{MeV})/E(\text{GeV}))^{1/2}$.

6.2.8 *Muons in a Dense Material*

The velocity dependence of the average energy loss by collisions of singly charged particles (muons, pions, protons, ...) with electrons of the traversed medium differs slightly from formula (6.1) and is given by:

$$-\frac{dE}{dx} = k\frac{Z}{A}\frac{1}{\beta^2}\left[\ln\frac{2m_e c^2\gamma^2\beta^2}{I} - \beta^2 - \frac{\delta}{2}\right] \left(\mathrm{MeV}/\left(\mathrm{g/cm^{-2}}\right)\right) \tag{6.28}$$

where $\delta \approx \ln(\gamma)$ accounts for screening effects at high energy. As a function of energy of the incident particle the most probable value shows a slow increase (relativistic rise) followed by a plateau whose value depends on the density of the material. The energy loss reaches a minimum for $\gamma\beta \sim 3$, corresponding to muon energies of few hundred MeV.

At a given energy, the energy loss distribution of $-\mathrm{d}E/\mathrm{d}x$ in a slab of material has an asymmetric distribution around its most probable value, usually referred to as the "Landau-Vavilov" distribution [44, 45]. The muon energy loss in dense materials has been extensively studied [46]. Both, the absolute energy loss and the straggling function agree with measurements at the percent level [47] up to several hundred GeV.

For muon energies above ~100 GeV, bremsstrahlung, pair production and deep inelastic scattering start to contribute, generating tails in the energy distribution ('catastrophic energy loss') [48, 49]. As an illustration, the average contribution of these processes for muons in iron up to 100 TeV is shown in Fig. 6.27. Very roughly speaking a muon behaves as an electron with a critical energy scaled as $\approx (m_\mu/m_e)^2$. However, unlike for electrons or positrons, pair production is larger than bremsstrahlung.

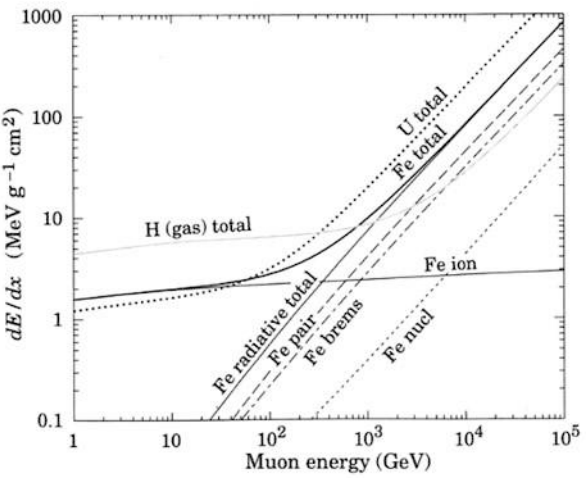


Fig. 6.27 Contributions to the energy loss of muons in iron, as a function of the muon incident energy. The total energy loss in hydrogen gas and uranium is also shown

Momentum correction to muon momenta can be applied, in setups where muons traverse a calorimeter before entering the muon spectrometer. For muons above ~10 GeV/c there is a good correlation between the total energy loss of muons in a calorimeter with the energy loss recorded in the active medium.

This is valuable, particularly for 'catastrophic' muon energy loss. Event-by-event correction for the muon energy loss is therefore useful in the hundred GeV momentum range for muon spectrometers behind the calorimeter with few percent momentum resolution [39].

Energy calibration and monitoring is frequently and conveniently done with muons. Exposing a calorimeter to a beam of electrons with well-known energy sets the 'electron-energy scale'.

In sampling calorimeters muons deposing a given energy produce in general more signal than electrons having deposited the same energy: $e/\mu < 1$. While establishing an absolute energy scale with muons requires very careful MC cross-checks, it is very convenient to use muons as a monitor of the calorimeter response as a function of time during data taking and as intercalibration tool between different parts of a calorimeter set-up [50]. The use of muons allows to transfer the absolute energy calibration established in a test beam to the experimental facility and to follow the energy calibration in situ using muons from physics channels. However, given the large dynamic range of energy measurements in many experiments, e.g. at the LHC and the smallness of the muon signal, complimentary calibration methods are necessary to achieve the required accuracy, see Sect. 6.3.6.

6.2.9 Monte Carlo Simulation of Calorimeter Response

Modern calorimetry would not have been possible without extensive shower simulations.

The first significant use of such techniques aimed to understand electromagnetic calorimeters. For example, electromagnetic codes were used in the optimization of NaI detectors in the pioneering work of Hofstädter, Hughes and collaborators [51]. One code, EGS4, has become the de facto standard for electromagnetic shower simulation [17]. Early hadronic cascade simulations were motivated by experimental work in cosmic-ray physics [52] and sampling calorimetry [53]. However, it were the codes developed by the Oak Ridge group [54], with their extensive modelling of nuclear physics, neutron transport, spallation and fission, which are indissociable from the development of modern hadron calorimetry [35].

Modern, high precision calorimetry and related applications have imposed a new level of stringent quality requirements on simulation:

- in many applications, electromagnetic effects have to be understood at the 0.1% level, hadronic effects at the 1% level;
- 'unorthodox' calorimeter geometries (Sect. 6.7) have to be optimized with simulation tools providing sophisticated interfaces to shower codes:

– in modern calorimeter facilities the energy deposits are usually distributed over several systems of different geometries and materials. Simulation codes are pushed to their limits in translating the recorded signal into a 1% precision energy measurement;
– at LHC and in particular in the study of the UHE Cosmic Ray Frontier simulation codes are used to extrapolate measured detector response by one to eight (!) orders of magnitude;
– particle physics MC codes are applied to areas outside particle physics, such as of radiation shielding, nuclear waste incineration and medical radiation treatment.

First, we will describe the general approach to these simulation issues before addressing some specific points. Regular conferences on this subject provide a good overview [55].

Electromagnetic Shower Simulation
For decades EGS4 [17] has been the standard to simulate electromagnetic phenomena. A modern extended incarnation has been developed by the GEANT4 Collaboration [18]. It includes the full panoply of radiation effects, including photons from scintillation, Cherenkov and Transition radiation up to electromagnetic phenomena relevant at 10 PeV.

Hadronic Shower Simulation
The simulation must cover the physics and the corresponding cross-sections from thermal energies (neutrons) up to (in principle) the 10^{20} eV frontier, requiring many different physics models; program suites, 'toolkits', such as GEANT4 [18], provide the user with choices of physics interaction models to select the physics interactions and particle types appropriate to a given experimental situation.

At high energies (~15 GeV to ~100 TeV)—in addition to measured cross sections—models describing the hadron physics are used, such as the 'Quark Gluon String' model [18], Fritiof or Dual Parton Models [56]. Such models are coupled to descriptions of the fragmentation and de-excitation of the damaged nucleus. At the highest energies other models, such as 'relativistic Quark Molecular Dynamics' models are being developed [57].

In the intermediate energy range (<10 GeV) Bertini-style cascade models [58] are employed to describe the intra-nuclear cascade phenomena. These models use measured cross-sections and angular distributions.

For the very low energy (<20 MeV) domain neutron transport codes have been developed, using experimental cross-sections.

The different energy regimes covered by these models are connected with parametric descriptions, in which cross-sections are parameterized and extrapolated over the full range of hadronic shower energies. Well-known examples are Geisha [59] and to a certain extent GCalor (or GEANTCalor) [60].

Applications: Illustrative Examples
We present comparisons of simulation with experiment to illustrate the quality of shower modelling.

(i) Energy Calibration and Reconstruction

Many physics programmes at the modern colliders (HERA, Fermilab, LHC) require energy measurements at the limit of the instrumental resolution and with ~1% accuracy. The calorimeters are frequently composed of different electromagnetic and hadronic instruments, made from different materials and sampling topologies.

Establishing the absolute energy scale in the reconstruction of particles (and jets) needs a major effort to understand the detector, from an instrumental and technical point. It requires a tight interplay between measurements and simulations. Energy calibration and reconstruction, proceeds in several steps. Customarily, a calorimeter (segment) is exposed to electrons, setting the 'electromagnetic' energy scale. For hadrons a 'weighing' has to be applied to each cell, such that.

$$E_{\mathrm{i}}\,(\mathrm{true}) = w_{\mathrm{i}} E_{\mathrm{i}}\,(\mathrm{reconstructed}) \;\text{ with } w_{\mathrm{i}} = \langle E_{\mathrm{i}}\,(\mathrm{true})\,/E_{\mathrm{i}}\,(\mathrm{reconstructed})\rangle\,.$$

E_{i}(true) expresses the total energy deposited. This can be a rather large correction, particularly in non-compensating calorimeters. In a further step, details of the energy reconstruction algorithm ('clustering') are simulated to evaluate the energy outside the cluster, usually chosen smaller than the true shower extent. In practical calorimeters, non-sensitive regions ('dead material', DM) are unavoidable leading to frequently sizeable corrections evaluated by MC.

Establishing the energy scale for jets is the most complex calibration task. Jets are calibrated with a series of simulation-based corrections and in situ techniques. In situ techniques exploit the transverse momentum balance between a jet and a reference object such as a photon, Z boson or multijet system for jets with $20 < p_T < 2000$ GeV, using both data and simulation. In this way an uncertainty in the jet energy scale approaching 1% is obtained for high-p_T-jets with $100 < p_T < 500$ GeV/ c. An uncertainty of about 4.5% is found for low-p_T jets ($p_T < 20$ GeV/ c), dominated by uncertainties in the corrections for multiple proton-proton interactions (pile-up), see Fig. 6.28 [61].

(ii) Particle Flow Analysis in Calorimeter Systems at Present and Future Colliders

An important recent development is an ambitious analysis strategy for reconstructing the jet energy in calorimeters, the "Particle Flow" concept. It aims at identifying and reconstructing individually each particle arising from the collision (proton-proton, electron-positron, ...) by combining the information from all the subdetectors. The resulting particle-flow event reconstruction leads to an improved performance for the reconstruction of jets and "Missing Transverse Energy" (MET). The algorithm also improves the identification of electrons, muons, and taus. While the concept has first been applied in the physics analysis at the LEP collider, it is presently heavly used by the LHC collaborations [62, 63]. The improvement can be dramatic, as shown in Fig. 6.29.

The benchmark performance for calorimeter systems (Sect. 6.7.6.2) for future colliders (International Linear Collider, ILC; Future Circular Collider, FCC) aims at a jet energy resolution of $\sigma(\mathrm{jet}) \sim 0.3/\sqrt{E(\mathrm{GeV})}$. This is motivated by the need to measure, e.g. W- and Z-decays into two jets with a mass resolution approaching

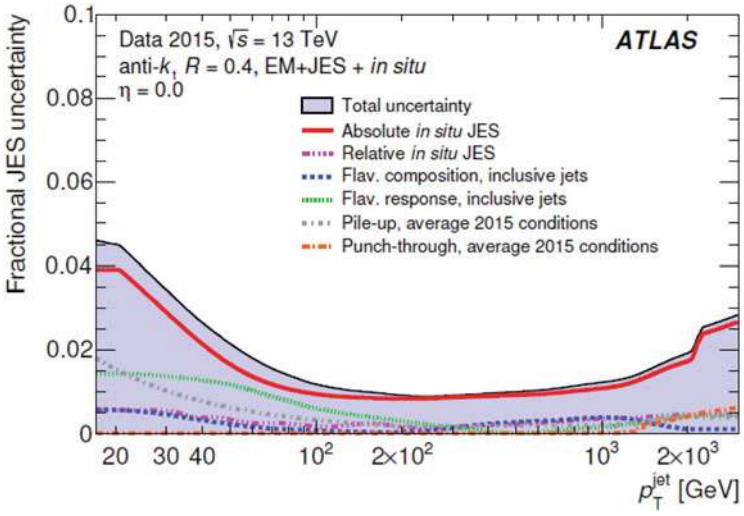


Fig. 6.28 Combined uncertainty in the jet energy scale (JES) of fully calibrated jets as a function of jet p_T in the central region of the ATLAS calorimeter system [61]

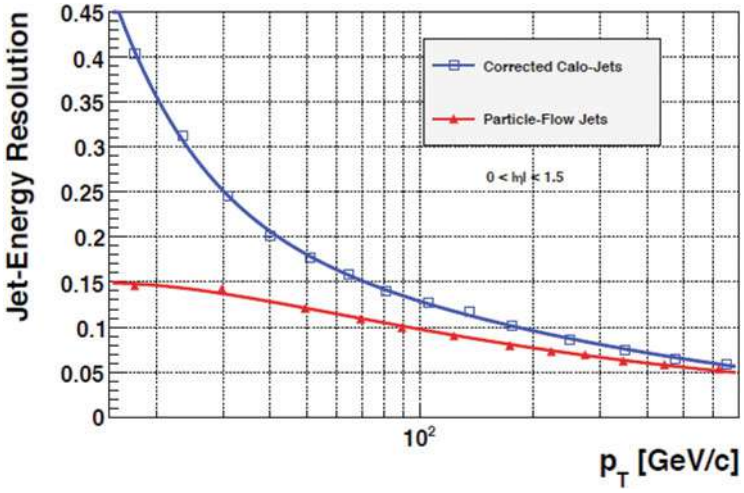


Fig. 6.29 Jet resolution for di-jets events in the CMS calorimeter reconstructed with the particle flow (red triangles) and the calorimeters (blue open squares) [63]

their natural width, i.e. with ~2 GeV (FWHM). Given that these jets are composed on average of ~60% hadrons, ~30% photons (the rest being shared by slow neutrons, neutrinos, muons, ...) a rather conventional resolution of $\sigma(\text{em}) \sim 0.15/\sqrt{E(\text{GeV})}$ and $\sigma(\text{hadronic}) \sim 0.5/\sqrt{E(\text{GeV})}$ would suffice, provided the individual energy deposits can be correctly associated with the individual particles measured in the charge particle spectrometer. This places a new level of performance requirements

on the calorimetry in terms of granularity, but also on the correct association of photonic and hadronic energy. Modeling has shown that this performance can be achieved in principle using the concept of 'Particle Flow Analysis'.[64, 65].

(iii) Ultra-High Energy Modelling
A particularly challenging application of these Monte Carlo techniques is extrapolation beyond present accelerator energies. The use of the Earth's atmosphere as a hadron calorimeter allows cosmic hadrons and nuclei up to and beyond 10^{20} eV to be probed. This requires 'dead-reckoning' of the detector response based on Monte Carlo techniques. Considerable faith in the extrapolation of the simulation models is needed in establishing the absolute energy scale. The estimate of the primary energy is based on measuring the shower shape: knowledge of F_{em}, the nucleon–nucleon cross-section, particle multiplicities, transverse momentum distributions, etc., all contribute to the estimate of the primary energy.

(iv) Low Energy Performance and Radiation Background
In many applications, e.g. dosimetry, careful modelling of the physics down to the MeV scale is needed. Certain codes [66] have been carefully benchmarked showing agreement to better than 20%, remarkable, as the very low-energy modelling of nuclear physics processes is involved.

Faithful modelling is also necessary to estimate the radiation levels in the LHC experimental caverns. Such modelling [67], based on the FLUKA code, was the basis for a number of design criteria and choices for the ATLAS and CMS experiments.

(v) Medical Applications
In cancer treatment with particle beams the tumour is exposed to proton or light ion beams, such as He or C^{12}, with energies of a few hundred MeV/nucleon. The energy deposition of the beam inside the human body (here the $1/\beta^2$ part of $\mathrm{d}E/\mathrm{d}x$ is relevant) can be monitored by positron emission tomography (PET), the β^+ emitters being produced through nuclear fragmentation reactions of the beam ions with the tissue nuclei.

Both, the patient treatment plan and the interpretation of these images is evaluated with the same MC programs as used in particle physics. More generally, the improvement in radiation treatments achieved with proper (particle physics) quality simulation is very significant, a very important legacy of particle physics to society [68].

We conclude that

- modern calorimetry owes much to Monte Carlo modelling;
- as always, predictions have to be taken with circumspection, in particular the extrapolation to performance and energy regimes inaccessible to experimental checks. Caveat emptor.

6.3 Readout Methods in Calorimeters

6.3.1 Scintillation Light Collection and Conversion

Scintillator materials used in calorimetry are inorganic crystals, organic compounds and noble liquids. Dense inorganic crystals represent one of the best techniques for homogeneous electromagnetic calorimetry. These crystals are insulators with a normally empty conduction band. When energy is deposited in the crystal, an electron can jump into the conduction band and cascade to the valence band by intermediate acceptor levels, part of the energy being emitted as light. The emitted light needs to be in the wavelength range where good photodetectors are available, and the crystal must be transparent to this wavelength range. The lifetime of the light emission depends on the concentration of acceptor levels, and temperature. In general, different decay times are present in the light luminescence spectrum of a given crystal (see also Chap. 3).

A list of commonly used scintillators, with some of their characteristic properties is given in Table 6.2. Crystals for homogeneous calorimetry are usually shaped as bars, typically of ~25 X_0 length and ~ 1 × 1 ρ_M transverse size. In colliding beam detectors, the cylindrical geometry leads in general to the use of tapered bars, with the incident radiation impinging on the smaller face. The growth of good quality ingots, followed by sawing and polishing to the needed size and surface quality requires specialized tooling available in industry. Careful packaging of the crystal in appropriate material (Tywek or equivalent) and sometimes lateral masking are needed to minimize the response dependence on position, transversally and longitudinally. The light detector (photomultiplier, photodiode, ...) is optically coupled to the back face of the crystal. The overall light yield, including the area and quantum efficiency of the transducer, influences the achievable energy resolution. A light yield of 1 photoelectron per MeV implies that the energy resolution cannot be better than $\sigma(E)/E = 3\%/\sqrt{E}$ (GeV). The number of emitted photons per MeV is in general much larger, being for example $4{\cdot}10^4$ in NaI doped with Thallium, one of the best scintillating crystal in terms of light yield. $PbWO_4$ produces ~150 times less light than NaI, but is far superior in other aspects (density, radiation resistance). New (and expensive) materials, like LYSO (a compound of Lutetium) are being developed for applications requiring fast response and high light yield.

A photomultiplier is schematically sketched in Fig. 6.30. All elements are located in an evacuated glass envelope. At the photocathode an electron is extracted by the photo-electric effect. A voltage difference accelerates the electron towards the first dynode out of which several electrons are extracted by secondary emission. This process is repeated over ~10 dynodes up to the anode at the highest (~1000 to 2000 volts) positive potential. With a sufficiently large gain at the first dynode the fluctuation of the number of electrons in the final charge pulse is dominated by the Poisson fluctuation of the number of photo-electrons. Amplification factors of several thousands are typical. A careful design of the High Voltage divider chain is mandatory to avoid non-linear effects. With recently developed "super bi-alkali"

Table 6.2 Properties of scintillating crystals applied in particle physics experiments

	NAI(Tl)	CsI(Tl)	CsI	BaF_2	CeF_3	BGO	$PbWO_4$	LYSO
Density [g cm^{-3}]	3.67	4.51	4.51	4.89	6.16	7.13	8.3	7.1
Radiation length [cm]	2.59	1.85	1.85	2.06	1.68	1.12	0.89	1.16
Molière radius [cm]	4.8	3.5	3.5	3.4	2.6	2.3	2.0	2.07
Interaction length [cm]	41.4	37.0	37.0	29.9	26.2	21.8	18.0	20.3
dE/dx)mip [MeV cm^{-1}]	4.79	5.61	5.61	6.37	8.0	8.92	9.4	9.2
Refractive index [at λ_{peak}]	1.85	1.79	1.95	1.50	1.62	2.15	2.2	1.8
Hygroscopicity	Yes	Slight	Slight	No	No	No	No	No
Emission spectrum, λ_{peak}								
Slow component [nm]	410	560	420	300	340	480	510	
Fast component [nm]			310	220	300		510	420
Light yield rel. to NaI								
Slow component	100	45	5.6	21	6.6	9	0.3	
Fast component			2.3	2.7	2.0		0.4	75
Decay time [ns]								
Slow component	230	1300	35	630	30	300	50	
Fast component			6	0.9	9		10	35

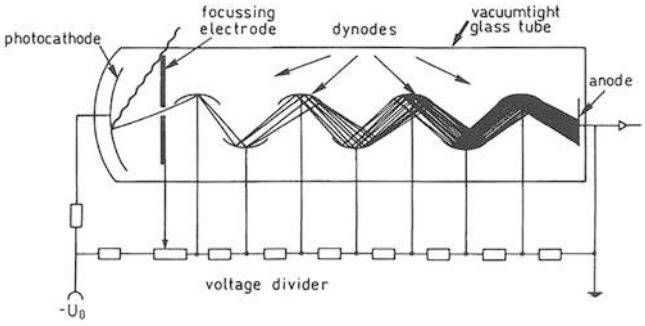


Fig. 6.30 Working principle of a photomultiplier. The electrode system is mounted in an evacuated glass tube

photocathodes (Cs-K) the quantum efficiency can reach more than 40% at 400 nm wavelength. For short wavelengths the efficiency is determined by the transparency of the entrance window. Quartz, CaF_2 or even LiF windows are necessary when efficiency in the near UV is required.

Because of their sensitivity to external magnetic fields, their rather large size and their cost, photomultipliers are nowadays being replaced by devices with less internal gain, followed by a high gain low-noise amplifier. Besides photo-triodes, the new devices are solid state based, like photodiodes or Avalanche Photo-Diodes (APD) [69]. Both offer good quantum efficiency, magnetic field insensitivity, moderate cost, small volume and—for APDs -a significant charge gain. The amplification is however accompanied by an "excess noise factor", of

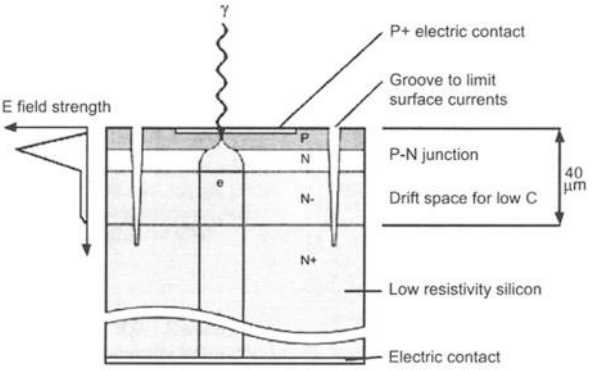


Fig. 6.31 Schematic diagram showing the structure of an avalanche photo-diode (APD)

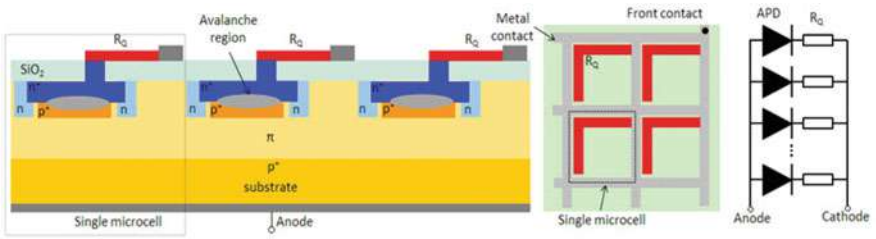


Fig. 6.32 Schematic diagram showing the structure of a Silicon photomultiplier (SiPM)

typically a factor 2 for a gain of ~50. This, together with the reduced size (and hence light collection) as compared to photocathodes can affect the energy resolution. The light detection and electron multiplication take place (see Fig. 6.31) in a thin layer (<40 μm) which lowers the sensitivity of APDs to minimum ionizing particles traversing the detector, as compared to simpler photodiodes.

The concept of APDs was extended to "Silicon Photomultipliers" by dividing the surface exposed to photons into small pixels, in a number large enough that each of them receives at most one photon.

Operating the device in the Geiger mode-i.e. with a very large gain-, and summing the current of a large number of pixels, one obtains effectively the equivalent of an analogue response to the number of incident photons, while each pixel operates in a binary mode.

Since the pioneering work [70], these devices have seen an extremely fast development [71]. A sketch of the layout of a SiPM is shown in Fig. 6.32.

Crystal calorimeters are the choice technology for precision electromagnetic calorimetry at medium energy machines like B-factories. CsI was used by Babar and Belle, and is used again for Belle II. The L3 experiment at LEP used BGO with success. However, the energy resolution reached for high energy electrons or photons (~50 GeV and above) was limited by the difficulty to calibrate a large

Table 6.3 Properties of noble liquids used in particle physics experiments

	LAr	LKr	LXe
Z	18	36	54
Boiling point [K]	87.3	119.8	165.0
Density in liquid phase [g cm^{-3}]	1.40	2.41	2.95
Radiation length [cm]	14.0	4.7	2.40
Molière radius [cm]	8.0	5.5	4.2
Nuclear interaction length for protons [cm]	84	61	57
Ionization properties			
Energy needed per electron-ion pair [eV]	24	17	15
Drift speed [mm/μs] at 10 kV/cm	5	3.8	2.6
Scintillation properties			
Emission spectrum, λ_{peak} [nm]	128	147	174
Decay time [ns]			
Fast component	5.0–6.3	2.0	2.2
Slow component	860–1090	80–91	27–34
Relative light yield in fast/slow component			
Fast component	8% (57%)	1%	5% (31%)
Slow component	92% (43%)	99%	95% (69%)
Refractive index at 170 nm	1.29	1.40	1.60

system (constant term of the energy resolution, see Eq. (6.23), of about 1% for the L3 BGO system) and not by the intrinsic resolution of the BGO crystals.

CMS and ALICE (for a part of its angular coverage) at the LHC decided to use $PbWO_4$. The most challenging case is CMS, given the very large size of the EM calorimeter, and the high radiation levels in the high luminosity collision points of the LHC, with nominally 500 fb $^{-1}$ of integrated luminosity at 14 TeV. More details are given in Sect. 6.7.3.

In some applications crystals are read on both ends, providing longitudinal information. However, so far it has not been possible to split the crystals longitudinally in independent segments without degrading the performances, a limitation for particle identification (see Sect. 6.4.3).

Noble liquids are also good, fast scintillators. Table 6.3 gives the properties of liquid argon, krypton and xenon already used in several practical cases for their scintillation properties.

In liquid argon about 4.10^4 photons are emitted per MeV deposited, a number very close to what is quoted for NaI. The light is however emitted in the far ultraviolet range, which complicates the conversion to electrical signals. Recent work [72] has shown that the scintillation light emitted by helium in the extreme vacuum ultraviolet range (~80 nm) can be used for particle detection, thanks to wavelength shifters (see below). The mechanism of scintillation in noble liquids involves the formation of excited diatomic molecules around the primary ions, which decay to free atoms by emitting radiation. In order to keep the emitted light associated with a well-defined region of space, thin reflecting boxes can be

introduced in the liquid volume. At present, one of the largest size detectors using light from noble liquids is the xenon calorimeter of the MEG experiment [73] (see also Sect. 6.7.1). As already mentioned in 6.2.3, the search for dark matter has triggered the development of several large size experiments using liquid xenon. These experiments [74] exploit both the scintillation and the ionization signal of the sought for nuclear recoils. Ionization electrons are preferentially transported to the surface of the liquid bath where, in a high electrical field region, they are extracted with high efficiency [74] and accelerated in the gas phase, giving in turn rise to (delayed) light emission. One example is described in Sect. 6.7.2.

Future long baseline neutrino experiments of very large size, like the DUNE [75] project at Fermilab envision liquid argon detectors of several tens of kilotons. DUNE will exploit both the scintillation and the ionization signals. In one of the read-out options, called "single-phase", the ionization signal is directly collected by a set of wires, each equipped with a readout chain, in order to have access to details of all secondary produced particles. The other option, "dual-phase", is close to what is described above for dark matter searches.

Liquid scintillators have been used abundantly in neutrino experiments, either in totally active large volume detectors, like Kamland and SNO, or as a large array of tubes filled with doped mineral oil.

The most recent example of the latter is NOvA [76] in which each tube is read out by means of a wavelength shifting fiber connected to a single pixel of an APD. The chapter on neutrino detectors provides further details.

Plastic scintillator plates, such as Polymethylmetacrylate (PMMA) doped with organic scintillator, have been used for electromagnetic and even more extensively for hadronic sampling calorimetry. The principal difficulty using this technology is the light extraction. The dimension of scintillator tiles of typically 10 cm $\times$ 10 cm size and 0.5 cm thickness would require light guides of typically 10 cm $\times$ 0.5 cm section in order to extract the light while preserving the emission phase space (respecting Liouville's theorem), a very difficult task in realistic detector layouts.

An elegant solution is the use of wavelength shifters [77, 78] in which due to their isotropic emission a constant fraction of the light is transported from the scintillating tile to a small rod, or even a plastic fibre separated from the tile by an air gap. The principle is shown in Fig. 6.33. Many calorimeter facilities at colliders were built following this principle, see also Sect. 6.7.

In a further development, detectors capable of accommodating smaller transverse granularities (like 5 cm $\times$ 5 cm) were proposed, like the "Shashlik" concept in which readout fibres cross the scintillating tile and the passive converter perpendicularly to their faces [79]. Originally considered in CMS, this scheme was later chosen by the LHCb experiment at the LHC for its electromagnetic calorimeter. A sketch of the arrangement of absorbers, scintillating tiles and fibers is shown in Fig. 6.34.

Even more ambitious was the "Spaghetti" calorimeter [80, 81] in which each calorimeter cell (typically 1 $\times$ 1 ρ_M transverse size and 25 X_0 deep) is built out of scintillating fibres embedded in a lead matrix, oriented parallel to the long side of the block. The electromagnetic calorimeter of the KLOE [82] experiment at the DAFNE

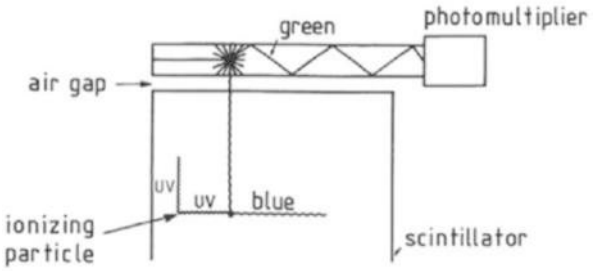


Fig. 6.33 Wavelength shifter readout of a scintillator

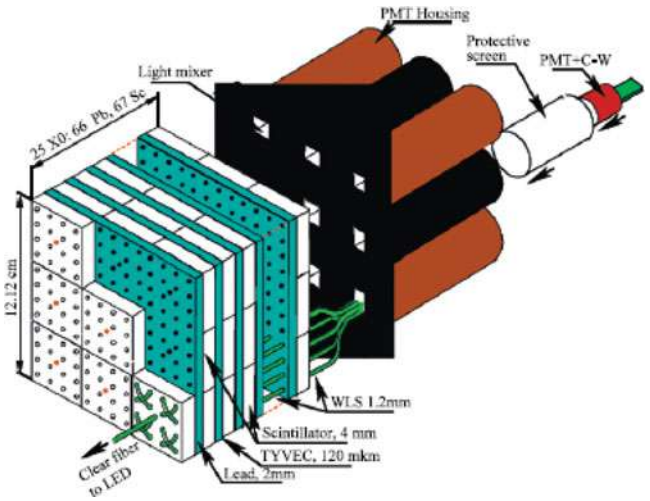


Fig. 6.34 The "shashlik" concept as realized in the LHCb Electromagnetic calorimeter

electron–positron collider in Frascati was built along these principles-although with a different geometry—and gave excellent results in the energy range of this machine.

6.3.2 Cherenkov Light Collection and Conversion

Although much less intense than scintillation light in good scintillators, Cherenkov radiation represents in some cases an interesting alternative. When a charged particle (electron or positron in the case of an electromagnetic shower) propagates in a transparent medium with a speed βc, larger than the speed of light c/n in this medium, an electromagnetic wave forms along a cone of half-angle $\theta_c = A\cos$

(1/βn) with respect to the incident particle direction, and with a number N of emitted photons in the visible range (400 to 700 nm) per unit length:

$$\mathrm{d}N/\mathrm{d}x = 490 \sin^2\theta_\mathrm{c} \left[\mathrm{cm}^{-1}\right]. \tag{6.29}$$

Lead glass, a dense material with a high index of refraction, has been used in several experiments (in particular OPAL [83] at LEP) with very similar geometries (tapered bars) as described above for scintillating crystals. The energy resolution is limited by the number of electrons and positrons in the shower above the Cherenkov threshold, resulting in a stochastic term $\sigma(E)/E$ of $> \sim 5$–$6\%/\sqrt{E}$, comparable to very good sampling calorimeters. Given the small number of photons, readout with photomultipliers is mandatory. As for crystals, longitudinal segmentation is in general not feasible. In several cases, "preshowers" of a few X_0 depth, instrumented with another higher granularity readout technique, have been used in front of lead glass arrays, in order to improve particle identification (see Sect. 6.4.3). Another limitation for large collider systems is the reduced response of lead glass to hadronic showers (a large fraction of the hadronic cascade is made of non-relativistic particles), inducing a performance limitation for hadronic calorimetry. However, the preponderance of Cherenkov-light production from electrons and positrons, i.e. the electromagnetic part of the hadronic shower, offers an interesting possibility. A hadronic sampling calorimeter instrumented with two sets of fibres—one set sensitive to Cherenkov-light only, the other set consisting of scintillating fibres, sensitive to all charged particles—can measure separately the electromagnetic component of the hadronic shower. This possibility is being studied in the dual-readout "DREAM" project. Test beam results are reported in Ref. [84].

Exploiting only the Cherenkov component, an hadronic calorimeter made of quartz fibers (parallel to the beam axis) embedded in an iron matrix has been chosen for the very forward calorimeter of the CMS experiment (for the pseudorapidity region up to 5). This choice was motivated by the high radiation resistance of quartz fibers, well adapted to this harsh environment [85].

Energy measurement with Cherenkov light produced in water was used with great success in very large detectors for nucleon decay and solar neutrino experiments, like Superkamiokande [86]. For the required detector volume of 50,000 tons water, the Cherenkov light was read out using large photomultipliers. In Superkamiokande, 50% of the outer surface of the detection volume is covered by 50 cm diameter phototubes. Electrons of 10 MeV are reconstructed with an energy resolution of about 15%. Their position in the detector volume is reconstructed with an accuracy of 70 cm and their direction with an accuracy of ~25 degrees. The detector also provides some discrimination between electrons (showering) and muons (single Cherenkov cone).

6.3.3 From Ionization to Electrical Signal in Dense Materials

One major avenue for calorimetry instrumentation is the measurement of the ionization charge produced in dense, active materials. In the presence of an applied electric field the charges move, inducing a current in readout electrodes proportional to the liberated charge and hence to the energy deposited by the showering particle. Electric charges are much easier to transport and to collect compared to light, which is the basic, decisive advantage of this concept.

This technique was introduced in the early 1970s [87] using liquefied argon as the active material. It has matured into one of the most widely used methods for calorimetry instrumentation, in particular, of sampling calorimeters. Noble liquid ionization calorimeters offer a number of attractive advantages, especially for instruments in the difficult environment of colliders. They are characterized by intrinsic stability and excellent uniformity of response (the only amplification is in the electronics chain which is fairly easy to calibrate), relative ease of a high segmentation and reasonable cost.

Other materials than argon are suitable for this method of detection, in particular the heavier noble liquids (Kr, Xe). In liquid helium and liquid neon, electrons are trapped in nano-scale cavities, and drift with characteristic speeds about a thousand times slower than electrons in other noble liquids. Solid neon was found to be usable at low rate [88]. Some saturated molecules like Tetramethylpentane (TMP), which is a liquid at room temperature, have also been tried. High purity at the ppb-level, required to avoid electron trapping, has limited their use compared to noble liquids, which however require cryogenic operation. The properties of noble liquids for ionization calorimetry are given in Table 6.3. Besides the value of $\mathrm{d}E/\mathrm{d}x$ and X_0 specific to the material, important parameters are the mean energy needed to create an electron-ion pair, the electron drift speed as a function of the electric field, and the dielectric constant, which affects the capacitance of a readout cell. Since the ions have a much smaller drift velocity compared to electrons, a track crossing a gap (and depositing charge uniformly) will give rise to a triangular current (see Fig. 6.35) given by Eq. (6.30) where $+Q_0$ and $-Q_0$ are the liberated charges, d the gap, and v the drift velocity of electrons. The resulting current is

$$I(t) = Qv/d \tag{6.30}$$

with $Q = Q_0(1 - vt/d)$. This formula is easily derived by remembering that a point charge q at a distance x from one of the parallel planar electrodes defining the gap of width d, induces a charge $-q(d - x)/d$ on this electrode, and $-xq/d$ on the other one.[2]

[2]In case of test cells with a grid at an intermediate potential in between the two electrodes, all charges of the grid-cathode region contribute with the same weight to the anode signal.

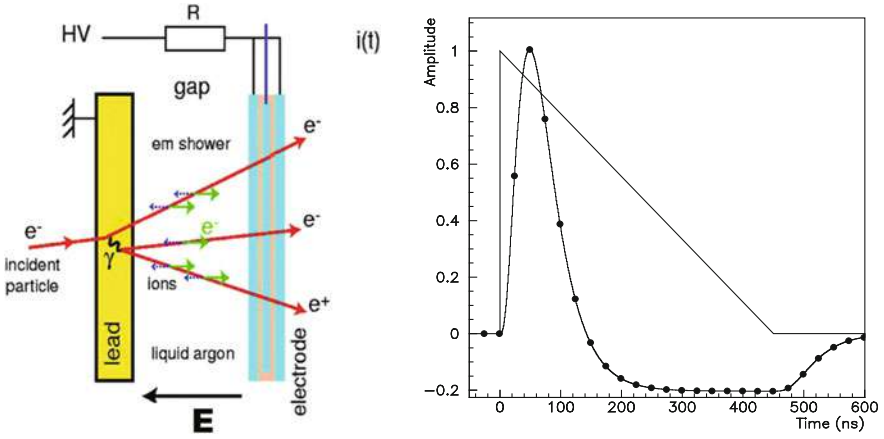


Fig. 6.35 Current induced by charges drifting in the sensitive gap of an ionization calorimeter. Left: charges drifting in the gap; right: current from drifting charges (triangle), and after CR-RC2 shaping. The dots every 25 ns represent times where the signal is being sampled (40 MHz sampling)

Depending on the rate of particles hitting a given cell, the readout can be an integrated charge readout (this charge is equal to $Q_0/2$ for uniform charge deposition in the gap) or a current readout. In the first case, the response is rather slow (~400 ns for a 2 mm gap in LAr). In the latter (Fig. 6.35) the response can be much faster (~40 ns rise time with a suitable CR-RC2 electronics filtering) but the signal to noise ratio is worse given that less "equivalent" charge is sampled, and the bandwidth of the electronics needs to be larger. At high speed (current readout) the limitation comes from the capacitance and inductance of the elementary readout cell, which must be kept appropriately small.

For LHC applications the optimization for high rate requires current readout with fast shaping, together with high granularity to limit pile-up of showers from consecutive events. While the electronics noise decreases when the electronics response becomes slower, the pileup noise generated by low energy particles from consecutive events increases. The shaping time is optimum when the two contributions are equal (see Fig. 6.36). One of the most ambitious realizations is the electromagnetic calorimeter of the ATLAS experiment at the LHC, which uses an 'accordion' geometry [89] to achieve the LHC performance specifications. This geometry provides full azimuthal symmetry without "cracks" between adjacent modules. The geometry, which includes three samplings in depth, is shown in Fig. 6.37. More details about the ATLAS calorimeter are given in Sect. 6.7.4.

The NA 48 collaboration at CERN developed a homogeneous noble liquid ionization calorimeter [90]. It had a cross-section of 2.5 m × 2.5 m, and was optimized for the study of neutral decays of high-energy neutral kaons. Liquid krypton was chosen as compromise between short radiation length (LXe would be preferable) and acceptable cost (the radiation length of argon is too large for fitting

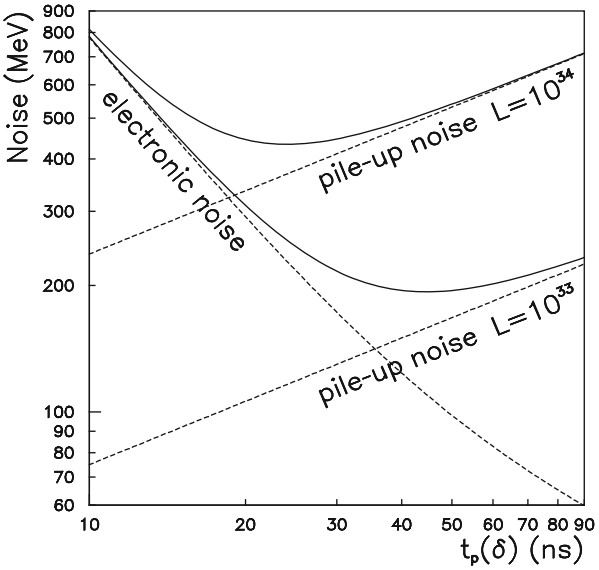


Fig. 6.36 Optimization of shaping time as a function of preamplifier noise and pile-up noise

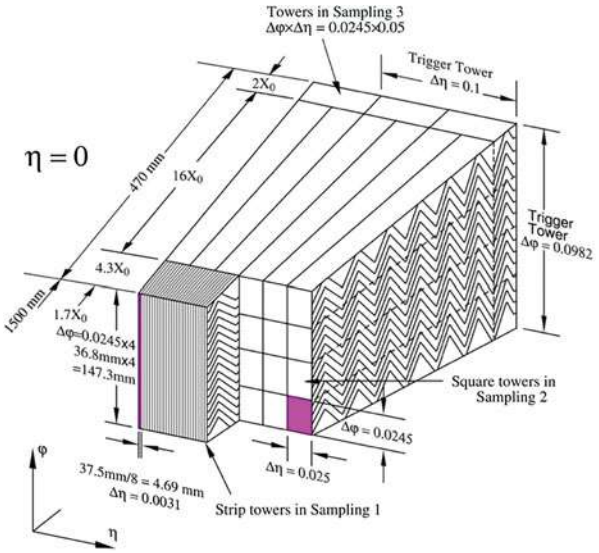


Fig. 6.37 Conceptual view of the 'accordion' geometry

a calorimeter able to contain high energy showers in an acceptable longitudinal space). Readout cells were defined by thin copper-beryllium ribbons stretched in the direction of the beam. The width of the bands (2 cm) and the gap (double gap of 2 × 1 cm) defined readout cells of 2 cm × 2 cm, smaller than the Molière radius

of krypton. In order to smooth the sampling of the shower, the bands were given a zigzag shape in depth by passing the ribbons through staggered glass-epoxy frames. The preamplifiers, connected to each signal band through a blocking capacitor, were located in the liquid for best performances. This calorimeter operated at a high voltage of 3 kV (0.3 kV/mm electric field), in a stable way during several years, with performances characterized by a stochastic term of $3.5\%/\sqrt{E}$, a signal peaking time of 80 ns, a noise per cell of 9 MeV (about 100 cells are needed to reconstruct with high accuracy an electromagnetic shower), a linearity better than 1 part in a thousand between 10 and 90 GeV, and an uniformity of response of 0.5%. Liquid krypton is also being used for the calorimeter (KEDR) of the VEPP2M collider at Novosibirsk [91].

Homogeneous noble liquid calorimeters with very high granularity readout can lead to very interesting imaging and energy measurement properties. One concept, inspired by gaseous tracking chambers (TPCs), was pioneered by the ICARUS collaboration [92, 93]. A more recent example is microBoone at Fermilab [94]. Detectors of this type with long drift distances (1 m or above) find their application in low rate experiments, such as neutrino experiments. The DUNE project, already mentioned, combines the readout of scintillation light and ionization.

A potentially attractive alternative to noble liquids is the use of silicon detectors. However, due to the high cost of silicon diode sensors, the silicon calorimeters operated so far have been restricted to places where the lack of space, and the limited volume, made the use of this technology mandatory. An example is given by SiCal [95], the luminosity calorimeter of the Aleph experiment at LEP. It consisted of a stack of 12 layers of silicon sensors interleaved with tungsten absorber plates, for a total thickness of ~24 X_0 in a longitudinal extension of only 150 mm. High resistivity, n-type (7 kΩcm, 300 μm thickness) Si was used for the 1.3 m^2 readout area, divided into 12,228 channels. The primary purpose of the detector was an absolute measurement of the luminosity using Bhabha scattering. The precision in the reconstructed position of showers (see Sect. 6.4.1) and the precision of the detector acceptance and alignment were essential for the measurement.

For the High-Luminosity LHC phase (HL-LHC) the CMS collaboration is embarking on an extremely ambitious replacement of the electromagnetic part of its end-cap calorimeters. Sampling calorimeters with Si-diode readout are being developed. The total Si readout area will be 600 m2 with a total of 6 million readout and 1 million trigger channels. Remarkably, intensive R&D has demonstrated that the Si detectors will withstand the radiation load [96]. This approach will be taken one step further for detector facilities at future colliders, such as a e^+e^- Linear Collider, with Centre of Mass energy up to several hundreds of GeV. Electromagnetic and hadronic calorimeters with extreme granularity and up to 100 million channels are being considered [97]. For such devices the use of Silicon sensors is one technology of choice. The cost of this option may be an obstacle, to be weighted against the potential performance advantages (see Sect. 6.5). In the forward direction, where the level of electromagnetic radiation from the beams is expected to be high, more radiation resistant sensors, like diamond, are being considered [98].

6.3.4 Gas Detectors

Charge collection in gases, usually followed by some degree of internal amplification, forms the basis of another important category of ionization sampling calorimetry. This method lends itself naturally to highly segmented construction and has profited from the diversified developments of gaseous position detectors (see Chap. 4). The relatively low costs of gaseous detectors favours their use in large area applications such as calorimeters for neutrino physics.

While gaseous ionization calorimetry offers several of the advantages found in ionization calorimetry with dense active materials, the low density of the gaseous readout planes—even compensated by internal charge amplification—limits the performance of such devices [29]. The low density has several disadvantages: Landau fluctuations of the energy deposit in the active gaseous layers can be comparable to the mean deposit and contribute to fluctuations at levels similar to sampling fluctuations; low-energy shower-electrons may multiple-scatter into the readout planes, where they may travel distances large compared to the gap thickness of the active layer, resulting in path-length fluctuations. These effects are relatively unimportant in dense materials, but may reach the level of Landau fluctuations in gaseous readout. Soft particles in the shower will spiral in strong magnetic fields, further increasing these path-length fluctuations. The absolute level of gas amplification depends on external operating conditions (pressure, temperature, gas composition) and is therefore difficult to control precisely. Variations of gas amplification also contribute to worsening the resolution.

An illustration is the electromagnetic calorimeter of the Aleph experiment at LEP [99]. The barrel part of the calorimeter consisted of 12 identical modules surrounding the central tracking system (a Time Projection Chamber), immersed in a solenoidal magnetic field of 1.5 T. The modules had 45 lead/wire-chamber layers for a total of 22 X_0. The cathodes of the readout chambers were segmented into pads of ~30 × 30 mm, providing energy and position information for each shower. The calorimeter was operated with a xenon-CO_2 mixture to increase the density of the active medium, thus reducing pathlength fluctuations. Wires connected to the pads of each layer were brought to module edges, where they were grouped into towers pointing to the vertex. The towers were segmented in three layers in depth of 4, 9 and 9 X_0, respectively. The connections of individual pads to the module edges resulted in a large inductance and therefore limited the rise-time of the readout signals (in the μs range). This was acceptable at LEP given the low event rates. This calorimeter, segmented in 74,000 towers, had an energy resolution of $\sigma(E)/E = 0.18/\sqrt{E} \oplus 1.9\%$, with E expressed in GeV (due to internal amplification, the electronics noise term was negligible).

One of the weak points of this technique is the non-linearity of response. Test beam studies showed that the energy E_{raw} recorded for electromagnetic showers needed to be corrected by:

$$E_{\text{corr}} = E_{\text{raw}} \left(1 + 0.00078\ E_{\text{raw}}\,(\text{GeV})\right),$$

implying a 7.8% correction at 100 GeV. Such non-linearities affect in particular high energy jets in which several showers may be superimposed, thus affecting the result in a way difficult to correct.

While this technique was still adequate at LEP, gas calorimeters were not considered for the LHC. With a very small cell size allowing a binary readout, they may find some application in hadronic calorimetry for the ILC (see for example [100]). An exception at the LHC concerns the very forward region in which, due to the high density of energy deposits, gas ionization chambers (ie without any amplification) are being used for specific purposes, including beam loss monitoring and luminosity measurements [101].

6.3.5 High Rate Effects and Radiation Damage

High particle rates and associated backgrounds impact both on the performance and the useful operating time of calorimeters. Radiation damage needs to be considered for the active readout material and signal processing electronics. Particle rates drive the choice of the calorimeter technology and construction.

Calorimeters with gaseous readout are particularly vulnerable to the high radiation environment due to the ageing effects associated with internal gas amplification, as discussed in Chap. 4.

Such radiation damage is essentially absent in noble liquids making this technology one of the most intrinsically radiation-hard techniques used to date. However, care has to be taken to select adequately radiation resistant components, including electronics, to limit deterioration of the performance (e.g. due to out-gassing). Particularly vulnerable are plastic insulators used in multilayer electrodes or in signal cables. Among the insulators highly resistant to radiation and suitable for calorimeter construction are polyimide (like Kapton) and PEEK. A fundamental limitation of noble liquid calorimeters are space charge effects due to the low drift speed of the positive ions (typically in the range of few cm/s at a nominal electric field around 1 kV/mm). At high incident rates these ions form locally a charged domain which effectively shields the electrons in the gaps from the externally applied field, reducing the drift velocity and thus the signal. These space charge effects are inversely proportional to the square of the detector gaps [102]. For this reason the forward calorimeters [103] of the ATLAS experiment feature gaps down to 250 μm.

Scintillators suffer from the formation of colour centres, which absorb part of the emitted light. The qualification of $PbWO_4$ as a candidate for the CMS crystal calorimeter required a world-wide R&D programme to study the radiation damage effects and to develop methods of crystal growth improving the radiation hardness. Several impurities were identified, which affect transparency in the useful wavelength range (above 350 nm). The best radiation resistance was obtained for crystals grown in Pb/W stoechiometric conditions, with the addition of a small quantity (~100 ppm) of Nb and Y [104]. These crystals showed a light loss of

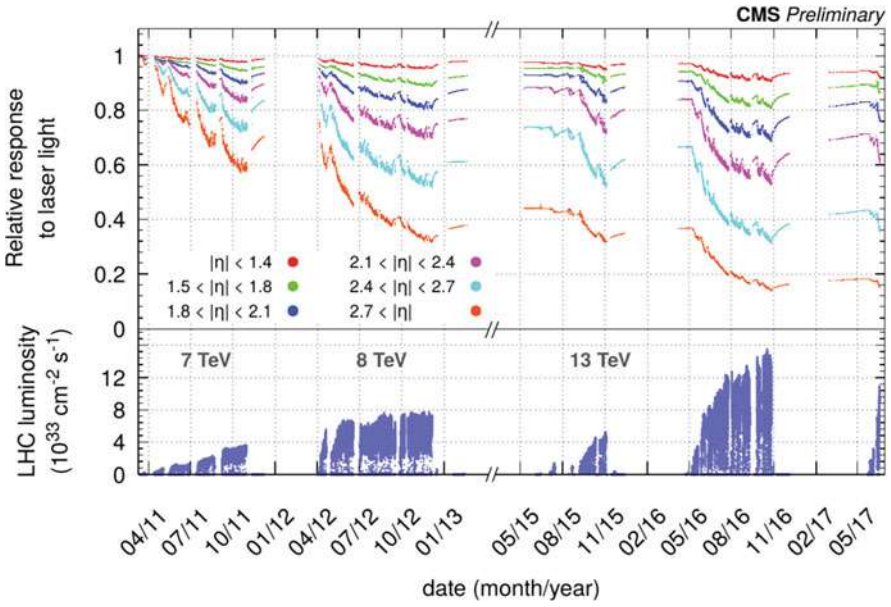


Fig. 6.38 Relative response of the CMS crystal calorimeter to laser light as a function of time, during the initial 5 years of LHC data taking

~3% after an exposure to ~10 Gy in ~10 h, corresponding to the radiation dose accumulated in calorimeters at LHC nominal luminosity during a typical operating period of 20 h. These colour centres show annealing with a recovery time of ~10 h (see also Sect. 3.1.1). After some years of data taking at the LHC, with instantaneous luminosities up to twice the nominal (i.e. 2 10^{34} cm^{-2} s^{-1}) and close to 100 fb^{-1} of accumulated data at 13 TeV in the centre of mass, there is enough experience to judge the crystal behaviour, conveniently followed using laser pulses sent in turn to each crystal. At central rapidities the light loss remains small, due to effective annealing between data taking periods. Some permanent damage accumulates in the more forward region. This is illustrated in Fig. 6.38 [105].

Radiation effects on the light transducers (APD) give an additional contribution to the electronics noise, still rather minor after the integrated luminosity quoted above.

As anticipated, the response of the ATLAS liquid argon calorimeter remains stable during LHC running. Using the position of the Z^0 mass peak reconstructed from electron-positron pairs, a variation of less than 0.05% over the whole 8 TeV data taking period of the "run-I" in 2012 is observed. The peak position is also independent of the mean number μ of collisions per crossing, ie there are no significant rate effects [106] at least up to μ of order 30.

6.3.6 Calibration and Monitoring of Calorimeter Response

Modern calorimetry operates frequently at the 1% accuracy level and requires therefore appropriate calibration methods. An extraordinary effort went into the development and deployment of adequate calibration techniques for the LHC calorimeters. In general, the following tasks have to be performed:

- establishing the absolute scale of response of a calorimeter, averaged over an entire data set
- assessing the uniformity and linearity of response
- monitoring the response as a function of time, locally and globally, in order to correct for time dependent effects, rate effects, aging.

A few examples are discussed below to illustrate each of these tasks.

Energy Scale

(i) *Low energy domain*: one large-scale example is the Superkamiokande experiment, dedicated to low-energy neutrino interactions. After a careful calibration of the gain of each of the phototubes, and an assessment of the water transparency (absorption length greater than 100 m), the absolute energy calibration was made using two radiation sources for cross-checks:

- the beam of an electron Linac operated in-situ above the liquid volume was sent through an evacuated beam pipe into several places of the detector volume recording the corresponding light signals. The Linac was operated at energies between 5 and 20 MeV. The absolute energy scale of the beam was known to better than 1%;
- ^{16}N radioactive nuclei were produced in situ from ^{16}O nuclei of the water volume using a neutron generator. The decay products to $^{16}O^{*}$ (beta emission with an endpoint energy of 4.3 MeV in coincidence with a 6.13 MeV photon) were then recorded during a few lifetimes of ^{16}N (7.13 s). The two methods agreed to better than 0.6% rms.

(ii) *Medium energy domain*: one example is the Babar experiment at SLAC, which used a CsI crystal electromagnetic calorimeter and employed three calibration sources to cover the full energy range:

- at low energy, the 6.13 MeV photons of ^{16}N decays were used (see Superkamiokande above). At this energy, the resolution of the calorimeter was found to be $5 \pm 0.8\%$.
- at high energy (~10 GeV) the Bhabha scattering was used. With a luminosity of $3 \cdot 10^{33}$ cm^{-2} s^{-1} this reaction provided about 200 events per crystal in a 12 h run.
- finally the peak position of known neutral resonances decaying in two photons were used for further checks. Figure 6.39 shows the recorded $\gamma\gamma$ invariant mass spectrum. The π^0 peak was observed at the nominal mass of 135.1 MeV with a width of 6.9 MeV.

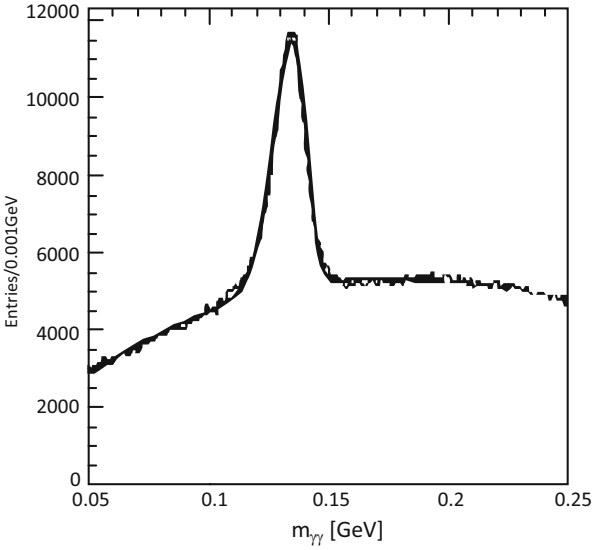


Fig. 6.39 Invariant mass of two photons in $B\bar{B}$ events recorded in Babar. The position of the π^0 peak provides the reference for the energy scale

– Bhabha scattering was used to calibrate the electromagnetic calorimeters of the four LEP experiments.

(iii) *High energy domain*: At the Tevatron the energy scale of the electromagnetic calorimeters was set using the precisely known mass of the Z^0 ($M_Z = 91{,}188 \pm 2$ MeV) decaying into e^+e^- pairs. The LHC experiments rely heavily on this approach given the high rate of Z^0 production: about 10 millions reconstructed Z^0 decays to e^+e^- were used by ATLAS and CMS to establish the energy scale of their electromagnetic calorimeter for the "run-I" at 7 and 8 TeV [106, 107]. The high-accuracy calibration of the electromagnetic calorimeter is essential for precision measurements (at the level of a few tens of MeVs) of the W mass [108] in the $e\nu$ decay mode, and for the measurement of the mass of the recently discovered Higgs boson, using decays in 2 photons, and in 4 leptons [109].

Uniformity and Linearity

With large enough statistics, the Z^0 mass constraint can be used to rescale in situ the response of an LHC calorimeter sector by sector and to improve its uniformity of response. ATLAS uses this method after dividing the calorimeter in about 30 slices in η. The residual non-uniformity is about 0.8% in the barrel region, being somewhat worse (up to about 3% locally) in the end-cap region [106].

If the amount of material in the magnetic spectrometer in front of the calorimeter is low enough, the relation between the energy measured in the calorimeter and the momentum measured in the spectrometer (E/p constraint) can be used to assess both

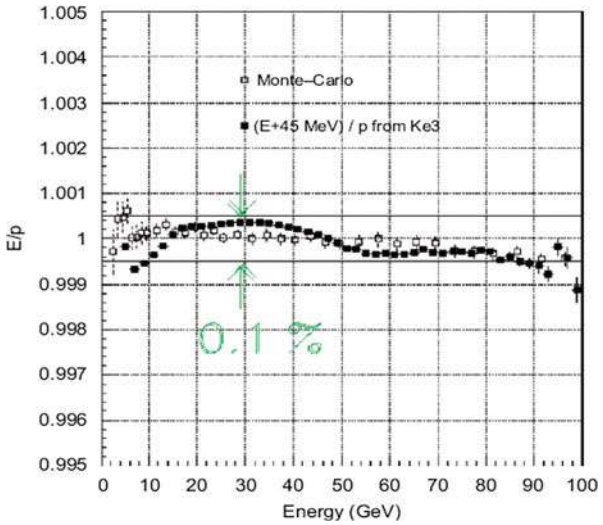


Fig. 6.40 Linearity of the NA48 homogeneous krypton calorimeter. The term added (45 MeV) corresponds to the average energy loss of electrons in the material preceding the sensitive volume

the uniformity and the linearity of response of the calorimeter. A correspondingly high precision mapping of the magnetic field in the spectrometer is of course needed. This technique was used with success in the NA48 experiment with a large sample of Ke_3 decays, demonstrating a linearity better than $\pm 5 \cdot 10^{-4}$ between 10 and 80 GeV, see Fig. 6.40. At the LHC the amount of material in the tracking volume is too large to get the best of this technique. Instead, the large sample of J/ψ decays in electron-positron pairs allows to assess the linearity of the electromagnetic calorimeters between ~5 GeV (high-p_T J/ψ are used in order to have a selective enough trigger) and ~ 50 GeV [107, 110]. An excellent linearity ($\pm$ $1 \cdot 10^{-3}$ between 20 and 180 GeV) was also demonstrated-locally-for ATLAS lead-liquid argon calorimeter modules exposed to a specially equipped beam line at CERN, used as a precision spectrometer (see Sect. 6.7.4).

Monitoring of Short Term Effects

In some cases the calorimeter response is subject to time dependent effects, on a time scale too short to allow for correction with the recorded physics data itself. External monitoring is in this case necessary. An example is the laser monitoring of the CMS crystal calorimeter designed to follow the light absorption and recovery as a function of the instantaneous luminosity, as discussed above, and shown in Fig. 6.38.

In many cases, the detector response depends on operating conditions. As an example, the energy response of the ATLAS liquid argon calorimeter depends on the temperature of the liquid bath with a coefficient of -2% per degree. Precision thermometers (Pt100 resistances) are used to follow the temperature with a precision better than 50 mK. Given the temperature stability observed no short-term correction

was required. In all precision experiments, the gain of the front-end electronics is monitored by injecting precision electrical pulses, allowing subsequent corrections to be made with a precision of 10^{-3} or better.

6.4 Auxiliary Measurements

The analysis of shower properties provides important additional information on position, angular direction and arrival time of the particles which initiated them. Shower shape analysis gives insight on the particle nature. The efforts lavished by the LHC collaborations on electron and muon identification and spectroscopy are eloquent testimony.

6.4.1 Position and Angular Measurements

Conceptually, two methods can be used to obtain spatial information: transverse and longitudinal granularity of the instrument on a scale smaller than the characteristic showers sizes gives position and direction by 'design'. Alternatively, if the readout volume is far larger than the shower dimensions, spatial information may be obtained by 'triangulation' using signals from several sensors distributed over the outer surface of the calorimeter volume.

The latter approach is used for calorimeters with large sensitive volume read out by photomultipliers distributed over their surface (e.g. Superkamiokande). The position is obtained by measuring the difference of light arrival times at the photomultipliers. With a timing resolution between 1 and 3 ns (depending on the pulse height) a position resolution of 70 cm is obtained for 10 MeV showers inside the sensitive volume.

In calorimeters with a more classical tower structure, the position of the incident particle is obtained by calculating the energy-weighted barycentre of energy deposition, using a cluster of cells around the local maximum energy deposition. Because of the finite size of the cells as compared to the Molière radius, the barycentre position is biased towards the centre of the cell with the largest energy deposition. This systematic bias can be corrected by fitting empirical functions. After applying this correction the position accuracy scales as $1/\sqrt{E}$ (decrease of shower fluctuations with increasing energy) convoluted with a constant and a noise term.

In the homogeneous NA48 krypton calorimeter (2×2 cm cells) a position resolution $\sigma_{x,y} = (4.2/\sqrt{E(\text{GeV})} \oplus 0.6)$ mm was measured, while the Babar CsI crystal calorimeter (4x4 cm crystals) gave slightly better results (3.2 mm/$\sqrt{E(\text{GeV})}$. This difference is explained by the smaller Molière radius of CsI (3.8 cm, against 5.5 cm for liquid krypton) and larger signal to noise ratio.

Segmented calorimeters, especially sampling calorimeters with ionization readout, allow lateral and longitudinal segmentation. With two or more samplings in

depth the direction of photon showers can then be estimated. As is shown in Fig. 6.13, the shower is particularly narrow and already well developed after~5 X_0; it is thus advantageous to sample it with high granularity over this depth. In ATLAS, with a cell size of ~5 mm the position of electron and photon showers is determined in the first ~5 X_0 (above~30 GeV) with an accuracy of about 300 μm, a critical asset for physics at the LHC. An important example is the discovery for the Higgs boson using the two-photon final state. The ATLAS electromagnetic calorimeter has three longitudinal samplings for measuring the direction of photons with an accuracy of about 50 mr/$\sqrt{E}$. This angular resolution is such that it makes a negligible contribution to the Higgs mass resolution [111], even if the interaction point cannot be identified among the numerous primary collision vertices at high luminosity. Search for new long-lived neutral particles decaying into photons (like gravitinos) also benefit from a high-resolution angular measurement.

6.4.2 Timing

The electromagnetic cascade develops at the sub-nanosecond timescale, allowing accurate timing measure-ements. This measurement allows identifying the bunch crossing associated to a particular event at colliders. Timing may be used to infer the shower position (see Sect. 6.4.1) or may discriminate between relativistic electromagnetic and slow particles, such as antineutrons.

In a segmented calorimeter the timing resolution is limited by fluctuations of the light path reflecting on edges of the tower, in case of light readout, or by electrical signal reflections at the ends of tower electrodes in case of ionization readout. Electronics noise and shower fluctuations introduce a further limitation, dominant at low and medium energies. While the energy in a tower can be obtained by sampling the signal at its maximum, the optimal time measurement requires additional signal processing. Constant fraction discriminators or digital treatment of multiple samplings of the signal (also beneficial for energy measurements) are frequently used. The shaping time of the electronics is a critical parameter in optimizing the timing accuracy.

As an example, the homogeneous NA48 krypton calorimeter showed a resolution of $\sigma = 0.5$ ns/$\sqrt{E}$, up to ~100 GeV. With the light readout in the "spaghetti" lead-fiber sampling calorimeter of KLOE [82] a spectacular resolution of 0.054 ns/$\sqrt{E} \oplus 0.14$ ns was obtained for photons between 50 and 300 MeV, allowing the shower barycentre along the spaghetti bar structure to be located with a precision of ~3 cm.

With a time resolution better than 100 ps, vertex localisation becomes possible, with an accuracy of a few cm. At the LHC, the rms spread of collision vertices along the beam axis is about 5 cm or ~180 ps. At high luminosity when 50 to 200 collisions per bunch crossing are observed, or envisaged (in the case of HL-LHC), a significantly better resolution is required in order to help in the vertex selection. Upgrade projects at HL-LHC are aiming at 30 ps, which seems the best

possible value with the technology available or under development. One of the most advanced projects is the High Granularity Calorimeter (HGCal) replacement of the crystal system in the endcaps of CMS [96]. In the dense core of the early part of the shower, the signal to noise ratio and the intrinsic shower fluctuations are such that a ~20 ps resolution has been obtained with Si diodes. A similar precision could possibly be reached for non-showering particles (mips) by using "low gain avalanche diodes" as developed and tested by several groups [112, 113].

For hadronic showers, the time development of the energetic component of the cascade is of the order of tenths of nanoseconds, whereas the thermal neutron capture may extend up to 1 μs. Nevertheless, typical time resolutions are found to be at the level of 1–2 ns/$\sqrt{E}$. As an example, with multiple digital sampling a time resolution of $\sigma = 1.5$ ns/$\sqrt{E}$ is measured in the ATLAS Tile Calorimeter [114]. The different time evolution of electromagnetic and hadronic showers offer interesting possibilities for improved shower treatment, a feature likely to be exploited at future facilities (see 6.7.6.2).

6.4.3 Electron and Photon Identification

Apart from certain final states easily identified, like Bhabha scattering at e^+e^- machines, electrons and photons are in general buried inside the copious production of hadrons or jets. This is particularly true at hadron colliders where the electron/hadron ratio ranges from 10^{-3} to 10^{-5}. Since electrons and photons are often signatures of interesting physics, their identification at the trigger and analysis level is crucial. The basic criterion for electromagnetic shower identification is the transverse and longitudinal shower shape, restricting em showers to the electromagnetic compartment, as opposed to hadrons and jets depositing energy in the full calorimeter. This condition is easy to implement, already at the trigger level. Comparing shower shape parameters in the electromagnetic compartment (width, length) to pre-programmed patterns provides the needed additional discrimination. Further discrimination is obtained by treating electrons and photons separately. An electron is signed by a charged track pointing to the shower barycentre, with a momentum p compatible with the calorimetric energy E. The rejection power of this E/p test is however compromised when the electron starts to shower in the tracking device in front of the calorimeter, distorting the momentum measurement and possibly the calorimetric measurement. The remaining background is dominated by π^0s overlapping with a charged pion. A photon is identified through the absence of a track pointing to its barycentre. At this stage the background for photons is often dominated by a π^0 decaying into close-by photons. Very fine granularity in the first ~5 X_0 is one approach to reject these π^0s. As an illustrative figure, simulations made for the ATLAS experiment, give a rejection factor of jets of about 3000 (for a photon acceptance of 80%), when studying the γ + jet final state as a possible background to the $\gamma\gamma$ reaction, with photon energies around 50 GeV [115]. For certain physics reactions an 'isolation criterion'-absence of tracks above a certain

p_T, nor calorimeter energy in a cone around the electromagnetic shower can be applied to sharpen photon or electron identification. This criterion does not apply e.g. for electrons resulting from heavy quark decays inside a heavy quark jet.

The Higgs boson discovery in the di-photon mode was a brilliant demonstration that the necessary jet rejection was achieved by both ATLAS and CMS experiments. At an invariant mass of the Higgs boson of about 125 GeV, the di-photon continuous background consists of about 75% prompt di-photons, 20% photon-jet background and about 5% jet-jet background.

Samples of electrons-positrons with an invariant mass around the Z^0 mass allow a clean measurement of the electron sample purity, as well as of the selection efficiency, using the "tag and probe" method, see Refs. [107, 110] for details.

6.4.4 Muon Identification

The registration of muons in calorimeters contributes to their identification, provides an important means of cross-calibration and in-situ monitoring of calorimeter cells and is used to improve the quality of the muon spectroscopy for instruments located behind the calorimeter.

Identification relies on the reconstruction of a penetrating, charged track behind the hadron calorimeter and possibly on the measurement of an energy deposit in the calorimeter cells along the path of the muon. Typical most probable energy deposits in an electromagnetic calorimeter (e.g. the CMS $PbWO_4$ calorimeter or the ATLAS Accordion) are of order 300 MeV, whereas in the hadronic calorimeters several GeVs are deposited. Such values are in general large compared to electronic noise and to energy deposits from particle background. In the ATLAS hadron calorimeter muons deposit more than ten times the energy from particle background due to average inelastic collisions, even in case of event pile-up at the highest collision rates.

Identification and triggering on muons based on calorimeter information is an essential complement to the main muon trigger using tracking chambers, for physics reactions producing low-p_T muons, e.g. tagging c- or b-jets, or detecting J/ψ or Y; production.

Muons are abundantly produced in pp. collisions (see Fig. 6.41). At low p_T the rate is dominated by 'punch-through' particles, i.e. hadrons, which have not interacted in the calorimeter. At high p_T prompt muons (in particular from W decay) become dominant. [116].

6.5 Jets and Missing Energy

Jet spectroscopy and the related signature of 'Missing Transverse Energy' (MET) have contributed to major discoveries (gluon, W-boson, top quark, ...). At LHC,

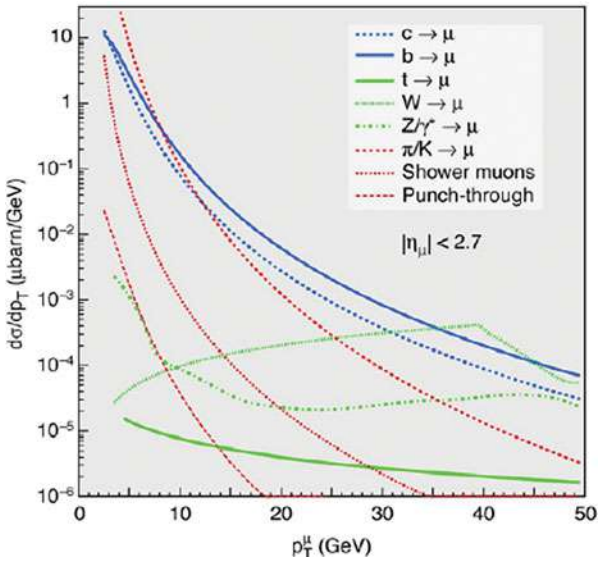


Fig. 6.41 Estimated muon spectra from various sources in the ATLAS Muon Spectrometer

MET is a key signature, e.g. for SUSY and/or dark matter searches. Very high-performance jet spectroscopy is also one of the principal design considerations for future Collider Detectors. The resolution and linearity of the jet energy reconstruction is the principle performance criterion.

The measured jet energy has to be related to the corresponding parton (quark, gluon) energy in a sequence of complex steps. Initial and final state gluon radiation and parton fragmentation affect the observable particle composition and momenta in the jet, limiting the 'intrinsic' parton energy resolution to order $\sigma(E_{\text{parton}})/E \approx 0.5/\sqrt{E_{\text{parton}}(\text{GeV})}$ [117]. Experimental factors—different response as a function of particle species and momentum, nonlinearities, insensitive detector areas, signal noise, magnetic field—require large corrections. Finally, jets are not uniquely defined objects. Different procedures are used to attribute a particle to a given jet. The choice of 'jet algorithms' influences the energy attributed to the jet, as do the additional particles in the 'underlying' event or particles from other collisions, recorded with the jet ('pile-up') [117, 118]. Two classes of jet algorithms have been widely used: The cone-algorithm draws a cone in the η-φ space with radius $R = \sqrt{}[(\Delta\varphi)^2 + (\Delta\eta)^2]$ around a 'seed', an energy deposit above a certain threshold, calculates the total transverse energy $E_{\text{T}} = \sum E_{\text{Tparticles}}$ and the E_{T} position and iterates around the new cone position until a stable result is obtained. This algorithm is sensitive to soft radiation effects; its well-defined jet-boundary however eases corrections due to the underlying event produced in the hadron collision. The k_{T}—algorithm clusters particles according to their relative transverse momenta over the η-φ space, controlled by a size parameter D. This algorithm is theoretically attractive, because in principle infrared and collinear

safe, but results in irregular jet boundaries and complicates the underlying event corrections. Recent work [119] has given rise to an improved version, the anti-k_T algorithm, which is safe against infrared and collinear divergences of QCD, and has regular boundaries. This algorithm is now the "default" of most LHC analyses using jets. Remarkably, despite the complexity and magnitude of the experimental corrections, modern analyses (and Monte Carlos) achieve experimental jet resolutions comparable to (sometimes even better than) the resolution measured for single hadrons: $\sigma(E_{\text{jet}})/E) \approx \alpha/\sqrt{\sum E_{\text{particles}}(\text{GeV})} \oplus \text{c}$, where $\sum E_{\text{particles}}$ represent the energy of the particles associated with the jet and where α is close to the stochastic and c close to the constant term measured for single hadrons [120, 121].

Within a jet, the electromagnetic part—coming mostly from π^0 decays—is better reconstructed than the charged hadrons—mostly $\pi^{\pm}$ and $K^{\pm}$ or long-lived neutral hadrons (K^0_L, n, $\Lambda, \ldots$). While the latter are only detected in the hadronic calorimeter, modern algorithms aim to "replace" charged hadrons reconstructed in the hadronic calorimeter by the associated charged track, whose momentum is better reconstructed than the calorimeter energy. While this individual replacement of particles requires complex algorithms, the procedure has been constantly improved, giving rise to "particle flow" algorithms (see Sect. 6.2.9) which are alternatives to jet reconstruction from calorimeters alone. CMS [122] in general prefers the more performant "particle flow" rather than calorimeter reconstruction. Particle flow is well suited for algorithms analyzing a substructure within jets in view, for example, of distinguishing between jets originating from a high p_T W or Z from quark or gluon jets [113].

The jet energy scale can be experimentally validated studying specific final states in which the jet is balanced by a well measured object, such as γ + jet(s) or Z + jet(s). Another powerful constraint is provided by W's decaying into two jets. A convenient source for identified Ws is the ttbar final state, abundantly produced at the LHC. In the p_T range from 30 GeV to 300 GeV, the linearity of the jet energy scale over the whole angular range is better than 2% in both experiments [123, 124].

The measurement of MET' is the only way to infer the production of neutrinos or weakly interacting SUSY-type particles. It is defined as the negative vector sum of the momentum of all reconstructed objects (leptons, photons, jets) in an event, projected onto the plane transverse to the collision direction. In general, a "soft term" is added corresponding to tracks or energy deposits not associated to the reconstructed objects. At high luminosity, in order to avoid unwanted contributions from pile-up, only tracks are considered, because of their unambiguous association with the corresponding collision vertex. Empirically, a MET resolution of $\sigma(E_{\text{missing}})/E \approx 0.7\alpha/\sqrt{\sum E_{\text{Tparticles}}(\text{GeV})}$ is observed (at low luminosity) for soft collisions with α expressing the stochastic term for single hadron resolution. Calorimetric systems with an acceptance of at least $|\eta|$ ~5 and very good 'hermeticity' are required to achieve this performance.

For events with high p_T jets, at high luminosity and after adequate corrections for the contribution of the underlying event, and of residual pile-up, the resolution is only weakly increasing with the number of collisions during the relevant bunch

crossing, and is comparable to the level of the single hadronic particle resolution [125, 126].

6.6 Triggering with Calorimeters

The ability of calorimeters to provide rapidly (order 100 ns) information on the energy distribution of the collisions products is one of the major assets of this technique. In the very rich trigger 'menu' of the LHC experiments all but muon physics is based on calorimetric triggers at the first trigger level L1. The calorimeter trigger provides a selectivity of ~10^{-3} and reduces the 40 MHz bunch collisions rate accordingly. A 'Sliding Window' technique is used to search for local energy topologies in the $\Delta\eta \times \Delta\varphi$ transverse energy distribution. The optimum window size depends on the particle type (photons, electrons or jets), on their threshold, the depth of the calorimeter included in the sum and possibly luminosity. More complex topologies requiring isolated energy clusters (e.g. triggering on isolated photons or electrons) are also used. The L1 trigger is implemented with dedicated hardware processors. The trigger decision time or "latency" of its response is fixed, and is typically a few μs. The information contained in all detectors is "pipelined" during this time, in such a way that no dead time is generated by the L1 trigger. In subsequent stages, called "high-level-trigger" (HLT) selection criteria and energy thresholds are sharpened with software-based algorithms. The treatment during these phases is asynchronous, and many processors (up to thousands) work in parallel. One of the main challenges with the trigger systems is to allow recording W and Z leptonic decays (i.e. with transverse momenta thresholds below ~30 GeV) for calibration purposes, and for electroweak physics, without saturating the bandwidth of the data acquisition systems. As luminosity increases, refinements are necessary to meet this requirement. MET and B-tagging are part of the overall menu of the HLT, in which of the order of one thousand different conditions are examined in parallel. Triggers on hadronic decay modes of τs, which rely on narrow hadronic jets in the calorimeters are also implemented in HLT. See Ref [127] as example for ATLAS.

In LHCb, which addresses heavy flavour physics in the pseudorapidity range between 2 and 5, the transverse momentum thresholds are much lower, typically 3 GeV for both the electron and the hadron trigger. Such low thresholds are made possible due to the lower luminosity operation of the experiment (typically 0.4 $10^{\mathbf{33}}$ $\mathrm{cm}^{-\mathbf{2}}$ $\mathrm{s}^{-\mathbf{1}}$) and the high data acquisition rate (up to 1 MHz). See Ref [128] for details.

6.7 Examples of Calorimeters and Calorimeter Facilities

The development of calorimetric facilities was and continues to be driven by the main directions of particle physics. Not surprisingly, as particle physics had its origin in cosmic ray studies, rather crude hadronic sampling calorimeters were successfully used to measure the energy spectrum of cosmic rays [52]. Electron scattering experiments provided the impetus for the development of homogeneous [129] and sampling [130] electromagnetic calorimeters. A major step in understanding and perfecting hadronic sampling calorimeters was made for the study of hadron scattering experiments, both with protons and neutrons [131]. The basic properties of these instruments were derived and Monte Carlo studies helped to optimize them [132]. The ISR provided the next motivation for a major development effort [35], providing the basis for the calorimeter facilities at Fermilab, HERA and LHC. In parallel, equally innovative calorimeter developments were and are initiated for astro-particle physics.

The recent series of CP-violation experiments in neutral kaon decay has pushed the requirements for electromagnetic calorimetry (Sect. 6.3.3). The LEP physics program emphasized charged particle spectroscopy and identification, with one notable exception, the L3 electromagnetic BGO crystal calorimeter (Sect. 6.3.1) and U/gas hadron calorimeter. For the Fermilab Collider program general purpose electromagnetic and hadronic calorimeter facilities were developed; facilities with new levels of performance were required for HERA, motivated by the need for precision jet spectroscopy (Sect. 6.7.5).

The LHC physics needs state-of-the-art electromagnetic and hadronic calorimetry, optimized for photons at the 100 GeV scale and for jets at the TeV scale, posing challenging system questions, answered in novel and unconventional ways (Sects. 6.7.3 and 6.7.6.1). The Future Collider physics programmes require further performance improvements, particularly concerning jet spectroscopy, exploiting at the same time the specific operation environment (Sects. 6.7.6.2).

6.7.1 The MEG Noble Liquid Homogeneous Calorimeter with Light Readout

The MEG experiment at PSI [73] is dedicated to the search for lepton flavour violation in muon decays. It aimed at a sensitivity for $\mu \to e\gamma$ decays of 10^{-13}. This requires an outstanding background rejection (for example of the reaction $\mu \to e\nu\nu\gamma$), requiring a calorimeter with an excellent energy resolution for ~50 MeV photons and a sub-ns response to cope with the high rate.

The half-cylinder shaped calorimeter is shown in Fig. 6.42. It contains 800 litres of liquid Xenon, and is read out by 846 PMTs, covering approximately 30% of the outside surface of the detector volume.

Fig. 6.42 The MEG homogeneous xenon calorimeter during assembly

The PMTs have K-Cs-Sb photocathodes and silica entrance windows transparent to the peak of light emission (175 nm) of liquid xenon.

The detector was optimized for events with a single photon shower in the volume. An interesting technical feature is the construction of the front wall cryostat using a honeycomb technique for better transparency to photons.

High purity (at the ppb level) of the liquid is necessary to prevent absorption of UV photons by contaminants like oxygen and water. The measured absorption length, more than 3 meters, is much longer than the typical light path from emission to the PMTs. The PMT signals are digitized at 2 GHz with a 12 bit accuracy using custom designed electronics.

The energy scale of the calorimeter is calibrated with photons (17.6 MeV) from the Li(p,γ)Be reaction obtained by sending protons from a Cockroft-Walton source to a Li target close to the calorimeter. In addition, photons from π^0 decays produced by π^- hitting a LiF target are also used, with one photon being measured in the Xe calorimeter, and the other one in an auxiliary NaI crystal matrix.

The relative energy resolution at 50 MeV is $\sigma(E)/E = 1.3\%$. The position resolution is ~6 mm and the timing resolution 64 ps. This excellent performance, made possible with this innovative technique, matched the demanding requirements of the experiment.

An upper limit branching ratio of muons decaying to eγ of $4.2 \times 10^{-\mathbf{13}}$ has been published in 2016 [133], based on the total statistics of 7 $10^{\mathbf{14}}$ muons stopped in the target. This is the best limit so far. A plan has been put forward and accepted to pursue the experiment with various improvements, and a higher flux of stopping

muons. The liquid Xenon calorimeter is kept, but the PMTs are replaced by VUV sensitive SiPMs, with a size of 12×12 mm^2, in order to improve the photon energy and position resolution. The prospect is to reach a sensitivity of $5\ 10^{-14}$ [134].

6.7.2 The Xenon 1T Experiment

Xenon1T is the largest and most recent detector of a generation of xenon detectors optimized for the detection of nuclear recoils of very low energy (below 100 keV), as could be produced by the scattering of a WIMP on nuclei (xenon in this case). Observation of such recoils, if they were to be produced, requires high accuracy of the energy measurement and very low background. The detector, operated as a dual phase TPC, is sketched in Fig. 6.43 [135]. The sensitive volume is a vertical cylinder of about 1 m diameter and 1 m height. As described in Sect. 6.2.3 both the primary scintillation signal and the ionization signal are exploited.

The ionization electrons are first drifted to the surface by an electric field generated by a set of Copper rings at a linearly decreasing potential from a grounded grid under the surface to bottom. The field intensity is about 12 kV/m. Right above the surface a somewhat higher field accelerates the primary ionization electrons in such a way that they in turn excite (providing secondary photons) and ionize the

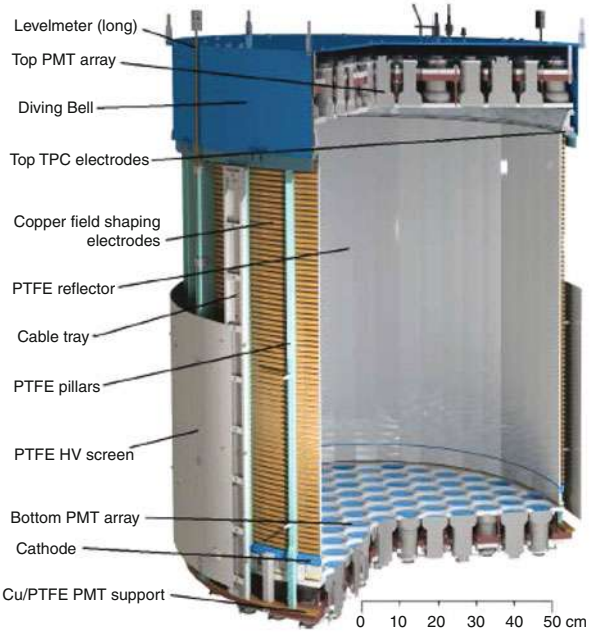


Fig. 6.43 Sketch of the Xenon-1T detector

surrounding gas. Both the primary and secondary photons are detected by a set of 248 VUV photomultipliers, with 78 mm diameter and 35% quantum efficiency at 175 nm, disposed in the liquid at the bottom of the vessel, and in the gas above the multiplication region. The light distribution in the top and in the bottom circles gives the position and lateral extension of the emitted signal. The time between the primary and secondary signals gives the vertical coordinate. All construction materials of the detector were selected for low radioactivity. The experiment is operated in the LNGS laboratory near the Gran-Sasso tunnel, shielded from cosmic background. It is furthermore enclosed in several layers of passive and active shielding. The remaining background is dominated by electron recoils from residual γ emitters, and nuclear recoils from residual neutron background. The former are very much suppressed by a requirement on the ratio of ionization over primary scintillation. The electron lifetime, which depends critically on the extreme liquid purity, and affects the magnitude of the ionization, is measured with photon to electron conversion signals generated in the liquid. A neutron generator is used to calibrate the energy response to recoils. The PMTs and electronics chain are calibrated with blue light pulses sent in fibers ending in the liquid volume. The dark count rate of the PMTs during the first science run was about 10 to 20 Hz. A first science run of about 30 days demonstrated that Xenon1T is the most sensitive device for WIMP masses above 10 GeV presently running. A science run of two years is planned. An enlarged version of the detector, Xenon-nT, with 8 tons fiducial volume is under construction. Its sensitivity should allow to approach the "neutrino floor" given by coherent scattering of solar neutrinos on nuclei.

6.7.3 The CMS Electromagnetic Crystal Calorimeter

The largest crystal calorimeter operated so far is the $PbW0_4$ calorimeter of the CMS experiment at the CERN LHC [110], clearly aimed at the Higgs $\rightarrow \gamma\gamma$ discovery. The calorimeter consists of a cylindrical barrel part (inner radius ~ 1.3 m) and two planar end-caps closing the cylinder at about 3 m from the proton-proton collision point (see Fig. 6.44). Each of the 61,200 barrel crystals is a tapered bar covering a $\Delta\phi \times \delta\eta$ solid angle of 0.018×0.018, and has a depth of 23 cm (24.7 X_0). In the end-caps, the calorimeter is preceded by a lead-Silicon strip preshower. Basic properties of $PbW0_4$ have been given in Sect. 6.3.1.

The calorimeter is located inside the hadronic calorimeter, which in turn is inside the 3.8 T superconducting solenoid. Barrel crystals are readout by APDs, while the end-cap crystals (somewhat bigger) are readout by phototriodes chosen for their better radiation resistance.

The front-end electronics processes signals corresponding to energy deposits of up to ~1.5 TeV (3.0 TeV) in the barrel (end-caps). The equivalent noise per crystal is ~30 MeV. This figure is likely to increase after high luminosity running, due to increased leakage current in the APDs.

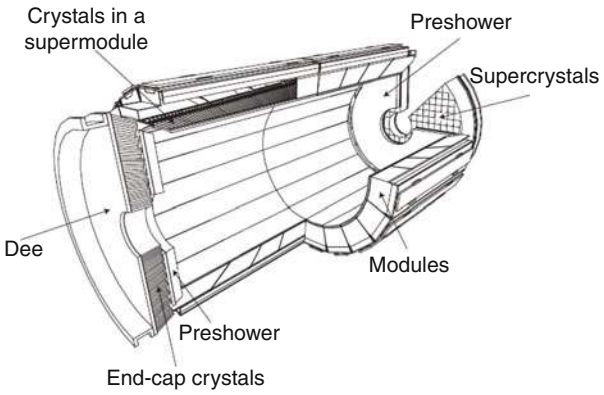


Fig. 6.44 Layout of the CMS electromagnetic calorimeter, showing the arrangement of crystals, with the preshower in front of the end-caps

Despite stringent quality controls during the crystal production, the particle response as observed in beam tests, showed an unavoidable crystal-to-crystal response dispersion of about 7% rms. Two calibration campaigns with beam test and cosmics were undertaken to establish the calibration constants for the initial LHC operation. Using various tools available at the LHC, like azimuthal uniformity of response, π^0, J/Ψ and Z^0 invariant mass constraints, all crystals were quickly intercalibrated to a precision around 1%. The laser pulse system monitors the short term response variations due to radiation effects.

The CMS crystal calorimeter successfully achieved its essential role for the experiment, both for triggering, as the source of identification and precise measurement of electrons and photons, and as input to particle flow. Among the most important results, based in particular on the calorimeter data, is the already mentioned discovery of the Higgs boson in 2012, revealed in the inclusive di-photon spectrum shown in Fig. 6.45.

6.7.4 The ATLAS Liquid Argon Electromagnetic Calorimeter

While ATLAS and CMS have almost identical physics programs, with the search for the Higgs boson as one of the main objectives, the two experiments have opted for a series of different detection techniques. The ATLAS electromagnetic calorimeter [103] uses a lead/liquid argon sampling technique, with an 'accordion' geometry, and is located outside of the inner solenoid. The liquid argon technique was chosen for its immunity to radiation, its intrinsic stability and linearity of response, and its relative ease of longitudinal and transverse segmentation. Its more modest intrinsic resolution is a limiting factor at medium and low energies.

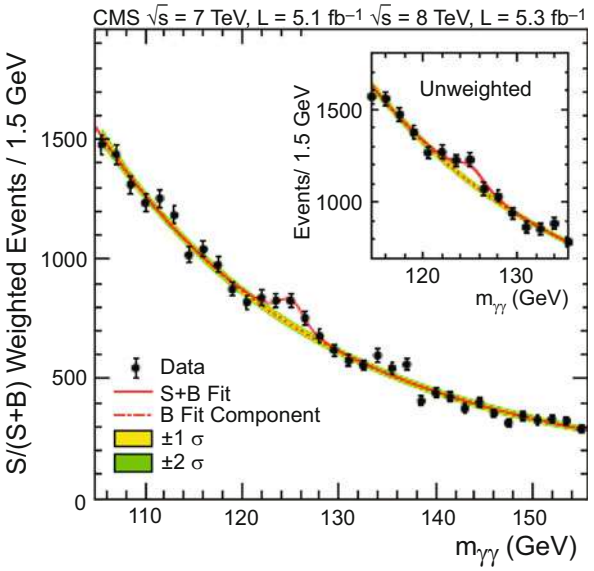


Fig. 6.45 Inclusive di-photon mass spectrum in CMS, from the Higgs discovery paper

The calorimeter features three segments in depth, the first one having an extremely fine segmentation in pseudorapidity (0.003) to allow separation between prompt photons and photons from π^0 decays up to $p_T \sim 70$ GeV/c, the interesting range for the Higgs boson search in the $\gamma\gamma$ decay mode.

The calorimeter is preceded by a presampler, located in the same cryostat, to correct for the loss of energy of electrons and converted photons in the inner detector material, in the solenoid and cryostat front walls (see Table 6.5). The barrel part, consisting of two cylinders, and the two end-cap wheels provide uniform azimuthal coverage despite being built of 16 (8) modules per cylinder (wheel) (Fig. 6.46).

The front-end electronics was optimized (Fig. 6.36) for best performance at the nominal LHC luminosity of 10^{34} cm^{-2} s^{-1}. The dynamic range is covered with three channels with gains in the ratio 1/9/81, digitized with 12 bit resolution. In this way quantization noise remains small compared to the noise level after the preamplifier (10 to 50 MeV depending on the sampling) up to the highest expected energy deposition per cell (~3 TeV). Trigger towers of size $\Delta\eta \times \Delta\varphi = 0.1 \times 0.1$ are built by analogue summing of signals at the front-end level, followed by digitization at 40 MHz with 10 bits ADCs (sensitivity of 1 GeV per count).

The uniformity of response within one module and the reproducibility from module to module were checked in a test beam. The overall dispersion of energy measurements in 3 barrel modules and 3 end-cap modules was respectively 0.43% and 0.62% [136]. The local energy resolution was found to be about 1% (rms) at 120 GeV [94], and is well described by $\sigma(E)/E = 10\%/\sqrt{E} \oplus 0.25/E \oplus 0.003$. The energy scale (Sect. 6.3.6) and the long range uniformity have been assessed in situ using the Z mass constraint. An overall "constant term" of about 0.8% in

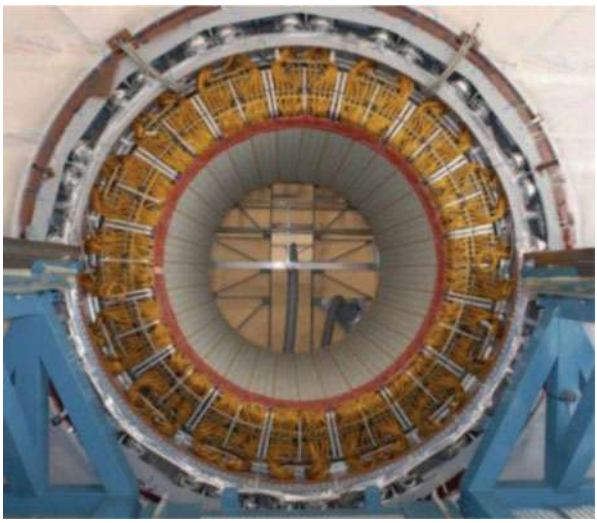


Fig. 6.46 Photograph taken during the assembly of the ATLAS electromagnetic barrel calorimeter. The pre-sampler sectors (in gray) are visible in front of the 16 calorimeter modules

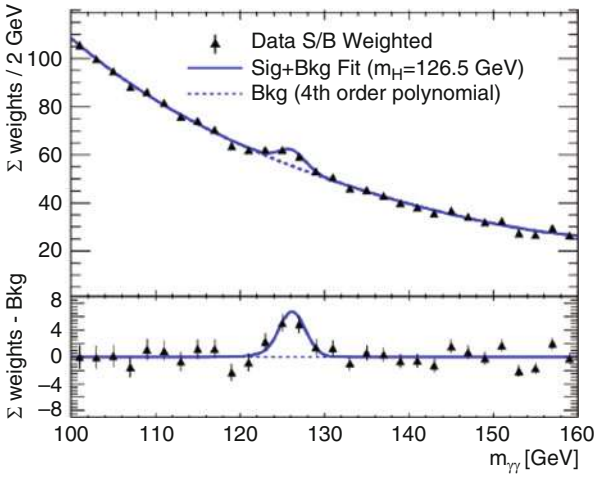


Fig. 6.47 Inclusive diphoton mass spectrum in ATLAS, from the Higgs discovery paper

the barrel and up to 3% in some pseudorapidity ranges of the end-caps covers the unavoidable dispersion in materials and in calibration, and the effects of material in front of the calorimeter not fully described in the simulation. Like for CMS, the electromagnetic calorimeter of ATLAS fulfilled successfully its task. Among the most important results, based in particular on the calorimeter data, is the already mentioned discovery of the Higgs boson in 2012. The corresponding inclusive diphoton spectrum is shown in Fig. 6.47. Also worth mentioning is the contribution of

the electron channel to the recent measurement of the W-mass, 80,370 ± 19 MeV in the muon and electron channels together [137].

6.7.5 The ZEUS Calorimeter at HERA

Research at the electron-proton collider HERA required precision jet spectroscopy at the 100 GeV level to study the underlying dynamics of e-quark collisions. Energy and position resolution for jets were at a premium.

The H1 Collaboration developed a calorimeter based on the LAr-Pb and LAr-Fe sampling technology. A certain level of 'off-line' compensation was achieved because hadron showers were measured longitudinally up to ten times and longitudinal shower-weighting could be applied [139].

The ZEUS Collaboration at HERA developed an intrinsically compensated calorimeter using the U-scintillator sampling technique [43, 138], modeled after the Axial Field Spectrometer facility [140]. The calorimeter is constructed in a modular form (Fig. 6.48), with units which are approximately 5 m long, 20 cm wide and more than 2 m deep. The ratio of the thickness of the ^{238}U plates (3.3 mm) to the scintillator plates (2.6 mm) was tuned to achieve $e/\pi = 1$, confirmed by measurements to be $e/\pi = 1.00 \pm 0.03$. The measured hadronic energy resolution, $\sigma\ (E)/E$ (hadrons) $= 0.35/\sqrt{E}$(GeV), is consistent with a sampling resolution of σ/E (sampling, hadrons) $\approx 0.29/\sqrt{E}$(GeV) and an intrinsic resolution of σ/E (intrinsic,

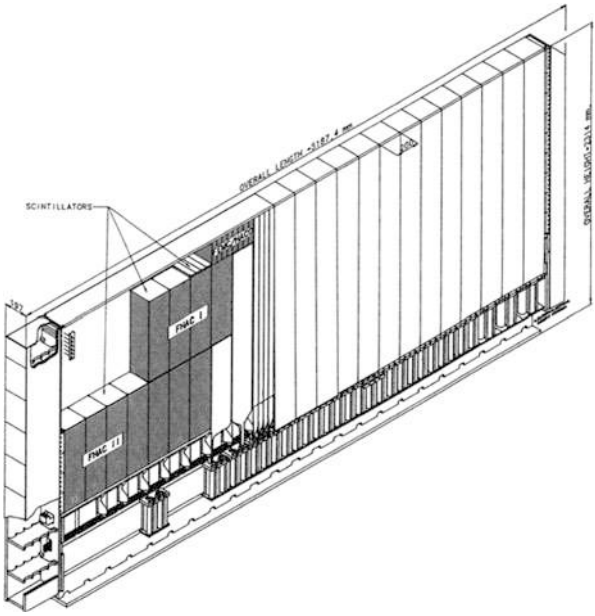


Fig. 6.48 View of a module of the ZEUS U-scintillator calorimeter. Wavelength-shifter readout is used to read cells of $5_{*}20$ cm^2 cross-section in the electro-magnetic compartment and of $20_{*}20$ cm^2 in the two subsequent hadronic compartments [138]

hadrons) $\approx 0.20/\sqrt{E}$(GeV). This sampling frequency is rather coarse for electrons resulting in an electron energy resolution σ/E (electrons) $= 0.18/\sqrt{E}$(GeV).

H1 and Zeus provided a detailed measurement of electron-nucleon scattering from which a new generation of parton distribution functions (PDFs) was derived. These functions have been used, and are still being used extensively for LHC physics analysis.

6.7.6 Facilities at the LHC and a Future Collider

The research programmes at the LHC and at a possible future Colliders impose a new level of performance requirements.

6.7.6.1 Facilities at LHC

The two general-purpose p-p experiments, ATLAS and CMS, have developed rather different approaches for the same physics research, promoted by different groups of physicists with their personal experience, background and taste, constrained by realities of funding. In both cases the extraordinary requirements on electromagnetic calorimetry imposed 'hybrid' solutions to allow independent optimization of electromagnetic and hadronic calorimetry. This 'independence' led ATLAS to choose two novel, unconventional detector geometries. The 'Accordion' calorimeter (see Sect. 6.7.4) is followed by a hadronic instrument with scintillator tile/WLS fibre readout. One of the 64 slices forming a complete and crack-less cylinder is shown in Fig. 6.49. The unconventional geometry of absorber plates and scintillating tiles oriented along the direction of the incident particle permits an economic construction and homogeneous sensitivity [141]. This geometry works because the preceding ~1.5 λ Accordion calorimeter provides enough hadronic shower development to permit good sampling in the Tile-geometry. This arrangement also greatly facilitates longitudinal and transverse segmentation hence permitting effective longitudinal weighting of the shower energies. Weighting leads to a resolution of the combined calorimetry system (accordion and Tile calorimeter) of $\sigma/E \approx (0.52/\sqrt{E} \oplus 1.6/E) \oplus 0.03$ and a good linearity of response [120]. A jet energy resolution of $\sigma(\text{jet})/E \approx 0.6/\sqrt{E}$(GeV) is estimated, adequate for LHC. The ATLAS Tile and Extended Tile calorimeter covers $|\eta| < 1.4$. For the forward ('endcap') regions ($1.4 < \eta < 3.2$) ATLAS had to adopt different solutions to cope with the even more ferocious radiation levels. An Accordion-type electromagnetic calorimeter precedes a Cu/Liquid Argon hadron calorimeter. In the very forward region ($3 < \eta < 5$) yet another novel geometry had to be invented: cylindrical readout elements with narrow LAr-gaps (0.25 to 0.35 mm) as sensitive medium are embedded in a tungsten absorber, sampling geometrically very tight showers at adequate readout speeds [120]. Figure 6.50 shows a cut-view through the ATLAS calorimeter facility.

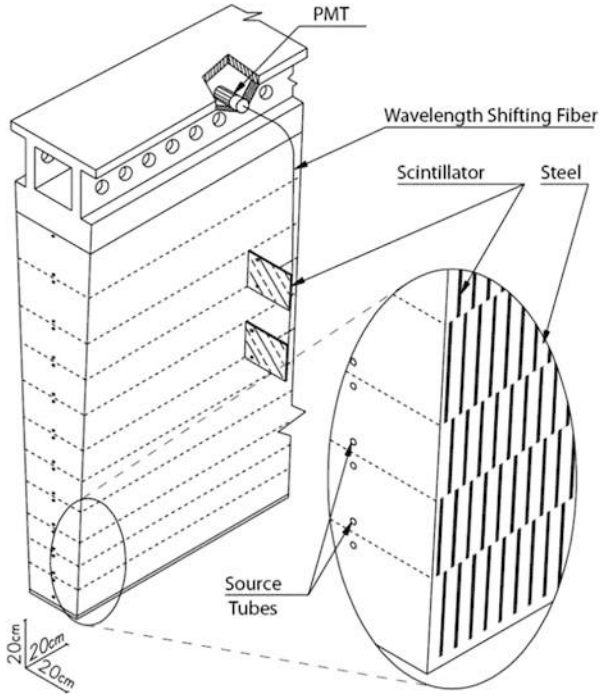


Fig. 6.49 View of one module of the ATLAS hadronic barrel calorimeter. Sixty-four such modules complete the cylindrical detector. Each of the longitudinally oriented scintillating tiles is instrumented with two wavelength-shifting fibers [141]

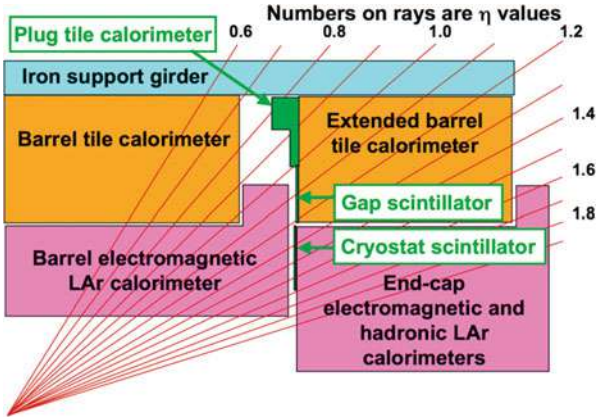


Fig. 6.50 Longitudinal quarter view of the Atlas calori-meter facility. The outer radius is at 4.2 m; it extends along the beam direction to ± 7 m. Auxiliary instru-mentation in the gap between the calorimeters allows energy correction for the non-instrumented zones [120]

Table 6.4 Parameters of the ATLAS and CMS electromagnetic calorimeter facilities

	ATLAS		CMS	
Technology	Lead/LAr accordion		$PbWO_4$ scintillating crystals	
	Barrel	**End-caps**	**Barrel**	**End-caps**
ηcoverage	0–1.475	1.4–3.2	0–1.48	1.48–3
Channels	110,208	63,744	61,200	14,648
Granularity ($\Delta\eta \times \Delta\varphi$)				
Pre-sampler	0.025 × 0.1	0.025 × 0.1	–	–
Strips/Si-preshower	0.003 × 0.1	0.003–0.006 × 0.1	–	32 × 32 Si-strips per 4 crystals
Main sampling	0.025 × 0.025	0.025 × 0.025	0.017 × 0.017	0.018 × 0.003 to 0.088 × 0.015
Back	0.05 × 0.025	0.05 × 0.025		
Depth				
Pre-sampler	10 mm	2 × 2 mm	–	–
Strips/Si-preshower	~4.3 X_0	~4.0 X_0	–	~3 X_0
Main sampling	~16 X_0	~20 X_0	~26 X_0	~25 X_0
Back	~2 X_0	~2 X_0	–	–
Energy resolution				
Stochastic term	10%	10–12%	3%	5.50%
Local constant term	0.20%	0.35%	0.50%	0.50%
Noise per cluster [MeV]	250	250	200	550

CMS calorimetry consists of the novel $PbWO_4$ electromagnetic calorimeter (Sects. 6.3.1 and 6.7.3) followed by a brass (70% Cu, 30% Zn) (50 mm thick) plate/scintillator tile calorimeter. The tiles are optically grouped into towers (0.087 × 0.087 in η-φ space in the barrel calorimeter) and read by hybrid photodetectors, all located in front of the 3.8 T superconducting solenoid. This favourable geometry, however, only allows for a total of ~7 λ, requiring a 'tail catcher' formed by scintillator tiles outside the coil in the first muon absorber layer [142]. Tables 6.4 and 6.5 summarizes the principal design parameters of the ATLAS and CMS Calorimeter Facilities.

6.7.6.2 Developments for Future Collider Calorimetry

The proposal for a future Linear e^+ e^- collider (LC) has triggered a worldwide R&D programme for the appropriate detector technologies [143]. One direction of present R&D addresses calorimetry optimized for its physics programme, emphasizing precision electromagnetic calorimetry and very high granularity for 'Particle Flow Analysis' (see Sect. 6.2.9).

Table 6.5 Parameters of the ATLAS and CMS hadronic calorimeter facilities

	ATLAS	CMS
Technolgy [ηhalf-coverage]		
Barrel/Ext. barrel	14 mm Fe/3 mm scint. [0–1.4]	50 mm brass/4 mm scint. [0–1.4]
End-caps	25 mm (front) − 50 mm (back) Cu/8.5 mm LAr [1.4–3.2]	80 mm brass/4 mm scint. [1.4–3.0]
Forward	Cu (front) − W (back)/0.25–0.50 LAr [3.2–4.9]	4.4 mm steel/0.6 mm quartz [3.0–5.0]
Channels		
Barrel/Ext. barrel	9852	2592
End-caps	5632	2592
Forward	3524	1728
Granularity ($\Delta\eta \times \Delta\varphi$)		
Barrel/Ext. barrel	0.1 × 0.1 to 0.2 × 0.1	0.087 × 0.087
End-caps	0.1 × 0.1 to 0.2 × 0.2	0.087 × 0.087 to 0.35 × 0.028
Forward	0.2 × 0.2	0.175 × 0.175
Longitudinal samplings		
Barrel/Ext. barrel	Three	One
End-caps	Four	Two
Forward	Three	Two
Absorption lengths		
Barrel/Ext. barrel	9.7–13.0	5.8–10.3
		10–14 (with coil/HO)
End-caps	9.7–12.5	9.0–10.0
Forward	9.5–10.5	9.8

One promising direction is being pursued by the DREAM Collaboration [144]. DREAM ('Dual REAdout Method') is a concept aiming at event-by-event separate detection of the electromagnetic component through Cherenkov light and the hadronic showers through scintillation light. Timing information might provide an additional handle to disentangle the various processes (e.g. delayed nuclear photon emission). The combined information could in principle allow complete reconstruction of the shower- and jet composition. The LC jet benchmark resolution of $\sigma/E \approx 0.30/\sqrt{E}$ might not remain a dream [84].

The CALICE (Calorimeters for the Linear Collider Experiment) Collaboration aims at the same performance: it makes the concept of Particle Flow Analysis an integral part of the design concept of the experimental facility aiming to separately measure the momenta of the charged component, photons in the electromagnetic and neutral hadrons (n, K^0) in the hadron calorimeter. The calorimeter is placed at a relatively large radius allowing the jets to open and charged and neutral particles to separate in the strong B-field. This strategy requires exceedingly high granularity (more than 10^8 channels) to measure the individual shower profiles [65].

Besides studies for a possible LC a vigorous programme has been intiated to undertand the physics potential and consequences for experimentation at a possible

"Future Circular Collider" (FCC). A center-of-mass energy for proton-proton collisions in the 100 TeV regime is envisaged, implying a collider circumference of about 100 km. The physics research determines the peak luminosity of about 3.10^{35} cm^{-2} s^{-1}. These key parameters shape the detector design and performance specifications, which are intensively studied [145]. The electromagnetic and hadronic calorimetry emphasizes very high granularity to cope with particle multiplicity and event pile-up, tight control of systematic effects (small constant term), very good linearity and—unsurprisingly—taming the ferocious radiation environment. The calorimeters are of the sampling type, because the stochastic term in the calorimeter performance is less an issue, given that the typical energy scales are in the TeV regime. Simulations show that rather conventional, LHC type calorimeter instrumentation will deliver the desired performance, without excluding novel developments with more "aggressive" technologies. LAr is the technology of choice, except for a possible scintillator option for the central hadron calorimetry. As an indication, the EM calorimeter could be a Pb/ LAr device, with cells sizes between 6_*6 mm^2 to 20_*20 mm^2 and an eightfold longitudinal subdivision. A possible geometry is shown in Fig. 6.51. Hadron calorimtry could be a scintillator/Pb/steel detector (in the central region), which would give e/h $\approx$ 1.1, resulting in the required good linearity and decent jet resolution, see Fig. 6.52.

While these concepts seem plausible, a closer look shows that the technical challenges are formidable ... fortunately, the LHC experience provided training, motivation and encouragement.

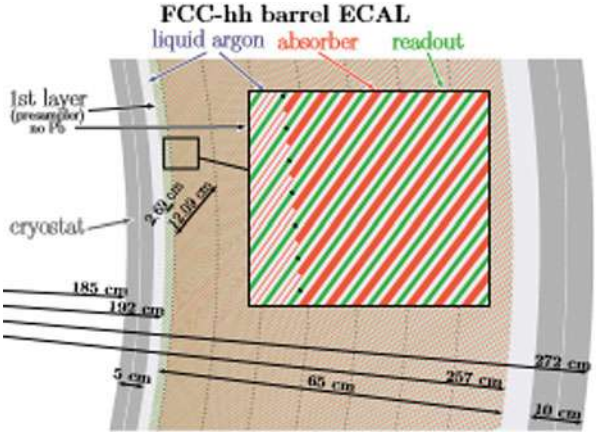


Fig. 6.51 Conceptual structure of an em calorimeter, showing the slanted absorber plates, LAr gaps and readout boards

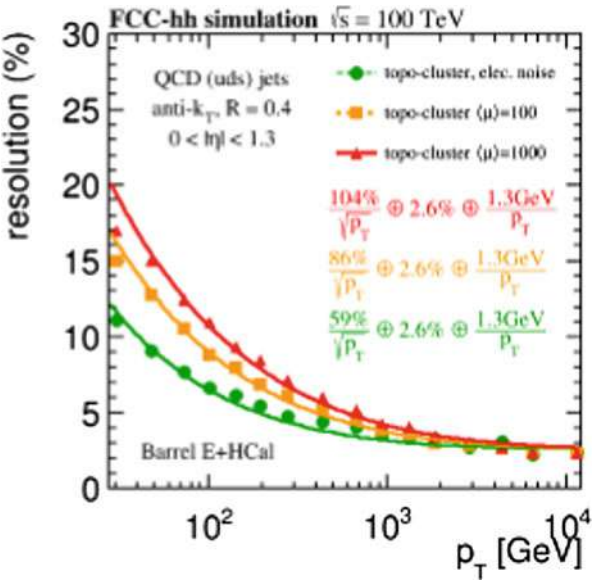


Fig. 6.52 Jet resolution for different hadron calorimeter configurations

6.8 Conclusions

During the past 40 years calorimetry has matured into a precision measurement technique, indispensable to modern particle physics experiments. The Higgs boson, cornerstone of our present understanding of matter, owes its discovery to calorimetry.

Understanding and modelling the physics processes at work in calorimetry at the 1% level has been achieved. Based on this understanding and helped by modern signal processing techniques, developments aim at characterizing the individual showers, at optimizing further particle identification and at reaching the intrinsic performance level for jet spectroscopy, needed for the next generation of precision and discovery experiments.

Acknowledgments We thank many colleagues for discussions and T. Carli for a careful reading of the manuscript.

References

1. M. Tanabashi et al., Particle Data Group, Phys. Rev. D98, 03001 (2018).
2. C. Grupen, Particle detectors, Cambridge University Press (1996).
3. W. Heitler, The Quantum Theory of Radiations (Oxford Clarendon Press) 1954.

4. H.A. Bethe, W. Heitler, Proc. R. Soc. A 146 (1934) 83.
5. Y.S. Tsai, Rev. Mod. Phys. 46 (1974) 815.
6. High Energy Particles, B. Rossi, Prentice-Hall (1952) revised (1961).
7. J.W. Motz, H.A. Olsen, H.W. Koch, Rev. Mod. Phys. 41 (1969) 581.
8. O. Klein, Y. Nishina. Z. Phys. 52 (1929) 853.
9. L.D. Landau, I.J. Pomeranchuk, Dokl. Akad. Nauk SSSR 92 (1953) 535, 735.
10. S. Klein, Rev. Mod. Phys. 71(1999) 1501.
11. R. Moore et al., Nucl. Instrum. Meth. B 119 (1996) 149.
12. D. Schildknecht, *Vector Meson Dominance*, arXiv:hep-ph/0511090.
13. T.H. Bauer, R.D. Spital, D.R. Yennie, Rev. Mod. Phys. 50, no. 2 (1978) 261.
14. H. Burkardt et al., Phys. Lett. B 206 (1988) 163.
15. B. Rossi, K. Greisen, Rev. Mod. Phys. 13 (1941) 240;
16. H.S. Snyder, Phys. Rev. 76 (1949) 1563.
17. W.R. Nelson, H. Hirayama, D.W.O. Rogers, *The EGS4 Code System*, SLAC-265 (1985).
18. S. Agostinelli et al., Nucl. Instrum. Meth. A 506 (2002) 250.
19. E. Longo, I. Sestili, Nucl. Instrum. Meth. 128 (1975) 283.
20. J.C. Philippot, IEEE Trans. Nucl. Sci. NS-17, no. 3 (1970).
21. A. Owens, Nucl. Instrum. Meth. A 238 (1985) 473.
22. U. Fano, Phys. Rev. 72 (1947) 1.
23. F. Gao et al., Nucl. Instrum. Meth. B 255 (2007) 286.
24. H.J. Crawford et al., Nucl. Instrum. Meth. A 256 (1987) 47.
25. E. Aprile et al. Phys.Rev. B76 (2007) 014115
26. J.B. Birks, Phys Rev. 86 (1952) 569.
27. A. Menchaca-Rocha et al., Nucl. Instrum. Meth. A 438 (1999) 322.
28. S. Amoruso et al., Nucl. Instrum. Meth. A 523 (2004) 275 and references quoted therein.
29. U. Amaldi, Physica Scripta 23 (1981) 409.
30. C.W. Fabjan, in: Experimental Techniques in High-Energy Physics, (ed. T. Ferbel), Addison Wesley, Menlo Park (1987).
31. R. Wigmans, Nucl. Instrum. Meth. A 259 (1987) 389.
32. A. Ferrari, private communication (2001).
33. T.A. Gabriel et al., Nucl. Instrum. Meth. A 338 (1994) 336.
34. D.E. Groom, in: Proc. seventh Intern. Conf. on Calorimetry in High Energy Physics, Tucson, AZ, (ed. E. Chen), World Scientific, Singapore (1998).
35. C.W. Fabjan, W.J. Willis, in: Proceedings of the Calorimeter Workshop, Batavia, (ed. M. Atac), FNAL, Batavia, Ill. (1975) 1.
36. R.K. Bock et al., Nucl. Instrum. Meth. 186 (1981) 533.
37. M.G. Catanesi et al., Nucl. Instrum. Meth. A 260 (1987) 43.
38. H. Abramowicz et al., Nucl. Instrum. Meth. 180 (1981) 429.
39. R. Wigmans, Calorimetry: Energy Measurement in Particle Physics, second ed.,Clarendon Press, Oxford (2017).
40. T. Carli et al., arXiv: 1604.01415v2 (2016).
41. C.W. Fabjan, R. Wigmans, Rep. Prog. Phys. 52 (1989) 1519.
42. R. Wigmans, Nucl. Instrum. Meth. A 265 (1988) 273.
43. G. Drews et al., Nucl. Instrum. Meth. A 290 (1990) 335.
44. L. Landau, J. Phys. (Moscow) 8 (1944) 201.
45. P.V. Vavilov, Zh. Eksp. Teor. Fiz. 32 (1957) 920.
46. W. Lohmann et al., CERN-85-03 (1985).
47. Z. Ajaltouni et al., Nucl. Instrum. Meth. A 388 (1997) 64.
48. A. Van Ginneken, Nucl. Instrum. Meth. A 362 (1995) 213.
49. D.E. Groom, N.V. Mokhov, S.I. Striganov, Atomic and Nuclear Data Tables 78, p. 183.
50. W. Lustermann, CALOR2008 (2001).
51. E.B. Hughes, IEEE Trans. Nucl. Sci. 19 (1972) 126.
52. V.S. Murzin, Prog. Elem. Part. Cosmic Ray Phys. 9 (1967) 247.
53. J. Ranft, Nucl. Instrum. Meth. 81 (1970) 29.

54. T.A. Gabriel, J.D. Amburger, Nucl. Instrum. Meth. 116 (1974) 33.
55. Hadronic Shower Simulation Workshop, Batavia, Ill., AIP Conf. Proc. 896 (2006).
56. G. Battistoni, in: Hadronic Shower Simulation Workshop, Batavia, Ill., AIP Conf. Proc. 896 (2006) 31.
57. J. Ranft, in: Hadronic Shower Simulation Workshop, Batavia, Ill., AIP Conf. Proc. 896 (2006) 102.
58. H.W. Bertini, P. Guthrie, Nucl. Phys. A 169 (1971).
59. H.C. Fesefeldt, Univ. Aachen, Report PITHA 85–02 (1985) .
60. C. Zeitnitz, T.A. Gabriel, The GEANT-Calor Interface User's Guide, ORNL, Oak Ridge (1996).
61. M. Aaboud et al., (ATLAS Collaboration),Phys. Rev. D 96 (2017) 077002.
62. M. Aaboud et al., (ATLAS collaboration), Eur. Phys. J. C 77 (2017) 466.
63. F. Beaudette (CMS Collaboration), (2014) arXiv:1401.8155 [hep-ex].
64. M. A. Thomson, in: Hadronic Shower Simulation Workshop, Batavia, Ill., AIP Conf. Proc. 896 (2006), pp. 215
65. C. Adloff et al., arXiv:0707.1245 and http://polywww.in2p3.fr/flc/calice.html
66. A. Fasso et al., Nucl. Instrum. Meth. A 332(1983)459.
67. E. Gschwendtner et al., Nucl. Instrum. Meth. A 482 (2002) 573.
68. F. Sommerer et al., Phys. Med. Bio. 51 (2006) 4385.
69. K. Deiters et al., Nucl. Instrum. Meth. A 453 (2000) 223.
70. B. Dolgoshein et al., Nucl. Instrum. Meth. A 563 (2006) 590.
71. E.Garutti et al., JINST 6 (2011) C10003.
72. N. McKinsey et al., Nucl. Instrum. Meth. A 516 (2004) 475.
73. R. Sawada, Nucl. Instrum. Meth. A 581 (2007) 522
74. B.N.V. Edwards et al. arXiv 1710. 11,032.
75. R.Acciari et al. arXiv:1601.02984.
76. D.S. Ayres et al. (NOvA collaboration), (2005) arXiv: hep-ex/0503053(2005).
77. R.L. Garwin, Rev. Sci. Instrum. 31 (1960) 1010.
78. B. Barish et al., IEEE trans. Nucl. Sci. NS-25 (1978) 532.
79. J. Badier et al., Nucl. Instrum. Meth. A 348 (1994) 74.
80. R. Wigmans, Nucl. Instrum. Meth. A 315 (1992) 299;
81. R.-D. Appuhn et al., Nucl. Instrum. Meth. A 386(1997) 397.
82. F. Ambrosino et al., Nucl. Instrum. Meth. A 598 (2009) 239.
83. OPAL Collaboration, Nucl. Instrum. Meth. A 305 (1991) 275.
84. S. Lee et al., Nucl. Instrum. Meth. A866 (2017) 76.
85. S. Chatrchyan et al., JINST 3 (2008) S08004.
86. S. Fukuda et al., Nucl. Instrum. Meth. A 501 (2003) 418.
87. W.J. Willis, V. Radeka, Nucl. Instrum. Meth. 120 (1974) 221.
88. V. Brisson et al., Nucl. Instrum. Meth. 215 (1983) 79.
89. D. Fournier, ECFA report 90–133(1990) 356, CERN 90/10.
90. V. Fanti et al., Nucl. Instrum. Meth. A 574 (2007) 433.
91. V.M.Aulchenko et al., Nucl. Instrum. Meth. A 394 (1997) 35.
92. S. Amerio et al., Nucl. Instrum. Meth. A 527 (2004) 329.
93. A. Ankowski et al., arXiv:0812.2373 (2008).
94. R.Acciarri et al., JINST 12 (2017) P02017.
95. D. Bederede et al., Nucl. Instrum. Meth. 365 (1995) 117.
96. CMS collaboration, Technical Design Report CMS-TDR-17-007 (2017)
97. J.C. Brient, Proc. XIII Conference on Calorimetery, CALOR2008, Pavia, 2008.
98. R.S. Wallny et al., Nucl. Instrum. Meth. A 582 (2007) 824.
99. D. Buskulic et al., Nucl. Instrum. Meth. A 360 (1995) 481.
100. I. Laktineh, Proc. TIPP09 Conf. (2009).
101. J.F.Beche et al., Proc. IEEE NSSMIC 2000.949101 (2000).
102. J. Rutherfoord, Nucl. Instrum. Meth. A 482 (2001) 156.
103. ATLAS Collaboration, Technical Design Report, CERN-LHCC/96–41.

104. A. Annenkov et al., Nucl. Instrum. Meth. A 426 (1999) 486.
105. Performance of the CMS precision electromagnetic calorimeter at the LHC run-II Z.Zhang, Talk at the CHEF Calorimeter Conference (Lyon-2017).
106. ATLAS Collaboration, Eur. Phys. J. C74 (2014) 10, 3071 and arXiv 1407.5063.
107. CMS Collaboration, JINST 8 (2013)9009.
108. T. Aaltonen et al., Phys. Rev. Lett. 99, 151,801 (2007).
109. ATLAS and CMS Collaborations, Phys. Rev. Lett. 114 (2015) 191803 and arXiv: 1503.07589.
110. V.Khachatryan et al. JINST 10 (2015) P06005 and arXiv 1502.02701.
111. ATLAS Detector and Physics Performance Technical Design Report, CERN-LHCC/99–14.
112. N.Cartiglia et al. NIM A850 (2017)83 and arXiv:1609.08681
113. ATLAS Collaboration (2017) arXiv 1708.04445.
114. R. Leitner et al., ATL-TILECAL-PUB-2007-002.
115. G. Aad et al., ATL-COM-PHYS-2008-243, Performance Chapter, p. 41.
116. ATLAS Muon Spectrometer Technical Design Report (1997) CERN/LHCC/97–22.
117. M. Bosman, ATL-CONF-2002-001.
118. S. Catani et al., Phys.Lett. B 285(1992)291.
119. M. Cacciari, G. Salam, G. Soyez, JHEP 0804:063 (2008).
120. G. Aad et al., ATLAS Collaboration, J. Instrum. 3 (2008) S08003.
121. S. Chatrchyan et al., CMS Collaboration, J. Instrum. 3 (2008) S08004.
122. CMS Collaboration, JINST 12 (2017) P10003 and arXiv 1706.04965.
123. CMS Collaboration, JINST 12 P02014 (2017) and arXiv 1607.03663.
124. ATLAS Collaboration, Phys.Rev. D96 072002 (2017) and arXiv 1703.09665.
125. CMS Collaboration, JINST 6 P09001 (2011) and arXiv:1106.5048.
126. ATLAS Collaboration, Eur. Phys. J. C77 (2017) 241.
127. ATLAS Collaboration, Eur. Phys. J. C77 (2017) 317 and arXiv:1611.09661.
128. T.Head for the LHCb Collaboration JINST 9 (2014) C09015.
129. R. Hofstadter, Phys. Rev. 74 (1948) 100.
130. R.R. Wilson, Nucl. Instrum. Meth. 1 (1957) 101.
131. J. Engler et al., Nucl. Instrum. Meth. 106 (1973) 189.
132. H. Schopper, Proc. Int. Conf. History of Original Ideas and Basic Discoveries in Particle Physics, Erice, 1994.
133. A.M. Baldini et al., Eur. Phys. J. C76 (2016) 434 and arXiv: 1605.05081.
134. M. Baldini et al. (MEG II Collaboration), PSI Proposal R-99-05.2) (2013) arXiv:1301.7225.
135. E.Aprile et al. Eur. Phys. J. C77 (2017) 881 and arXiv: 1705.01828.
136. M. Aharrouche et al., Nucl. Instrum. Meth. A 568 (2006) 601.
137. ATLAS Collaboration. Eur. Phys. J. C. 78(2018) 110, arxiv 1701.07240.
138. M. Derrick et al., Nucl. Instrum. Meth. A 309 (1991), 77.
139. B. Andrieu et al., Nucl. Instrum. Meth. A 336 (1993) 460.
140. T. Akesson et al., Nucl. Instrum. Meth. A 241(1985) 17.
141. ATLAS Collaboration, CERN/LHCC/96–042 (1996).
142. CMS Collaboration, CERN/ LHCC 97–31 (1997).
143. L. Linsen et al., CERN Yellow report CERN-2012-003 (2012).
144. N. Akchurin et al., Nucl. Instrum. Meth. A 584(2008)273.
145. A. Abada et al., FCC physics opportunities, Future Circular Collider Conceptual Design Report Vol. 1, Eur. Phys. J. C. 79, 474 (2019). https://doi.org/10.1140/epjc/s0052-019-6904-3

Chapter 7
Particle Identification: Time-of-Flight, Cherenkov and Transition Radiation Detectors

Roger Forty and Olav Ullaland

7.1 Introduction

Particle identification, PID, is of crucial importance in most experiments. The requirement can range from positive π/K identification in B-physics channels like $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ against a background from $B_s^0 \rightarrow D_s^- \pi^+$ which is ~15 times more abundant, to e/π separation at the level of $\sim 10^{-2}$ for momenta >1 GeV/c in order to effectively suppress a combinatorial background in channels like leptonic decays of heavy vector resonances.

The detectors should be non-destructive and should in addition introduce as little radiation length or interaction length as possible. We will in this chapter examine three experimental techniques which can be deployed for charged particle identification.

That is Time-of-Flight, Sect. 7.2, and Cherenkov detectors which measure the particle velocity relative to the speed of light in vacuum, $\beta = v/c$, Sect. 7.4, and transition radiation detectors which are sensitive to $\gamma = 1/\sqrt{(1-\beta^2)}$ of the charged particle, Sect. 7.5. These detectors cannot be stand-alone detectors for PID purposes. They all require that the momentum of the particle is defined by other means, see Sect. 4.3, and then

$$\frac{m^2}{p^2} = \frac{1}{\gamma^2 - 1} = \frac{1-\beta^2}{\beta^2} \tag{7.1}$$

allowing the mass m of the particle to be determined.
Only a limited amount of theory is included in this chapter as this is covered elsewhere in this book. The main emphasis will be on the working principles of these

R. Forty (✉) · O. Ullaland
CERN, Geneva, Switzerland
e-mail: Roger.Forty@cern.ch

C. W. Fabjan, H. Schopper (eds.), *Particle Physics Reference Library*,
https://doi.org/10.1007/978-3-030-35318-6_7

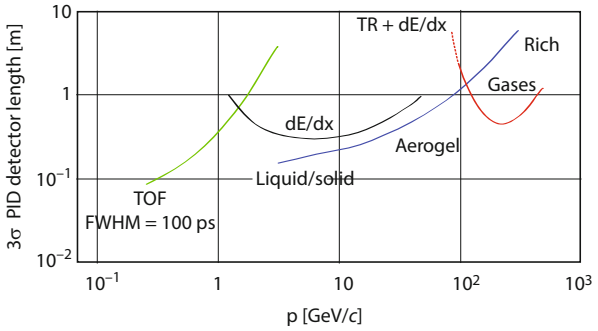


Fig. 7.1 Pion-kaon separation by different PID methods: the length of the detectors needed for 3 sigma separation. Adapted from [1]

detectors and how they are incorporated into compound experiments. A graphic representation of the different identification techniques can be seen in Fig. 7.1.

7.2 Time of Flight Measurements

The mass identification, m_i, of a momentum defined, p_i, charged particle is straight forward by measuring the flight time, t_i, over a path length, l. The mass, momentum, path length and flight time are related by:

$$m_i^2 = \frac{p_i^2}{l^2}\left[ct_i - l\right]\left[ct_i + l\right] \tag{7.2}$$

and the uncertainties by:

$$\left(\frac{\Delta m}{m}\right)^2 = \left(\frac{\Delta p}{p}\right)^2 + \gamma^4\left[\left(\frac{\Delta t}{t}\right)^2 + \left(\frac{\Delta l}{l}\right)^2\right] \tag{7.3}$$

There are essentially two sources of errors[1] in the measurement of time, t, in Eq. (7.3).

1. The limitation of the electronics to resolve short time intervals. A random time jitter in the pulse height at the detector and thereby a time slewing or time walk.
2. Variation of the transit time of the photons or the free electrons, and thereby the signal formation time, in the detector.

[1] *Irresolution* was proposed in [2]. Although a nice word, it did not catch on.

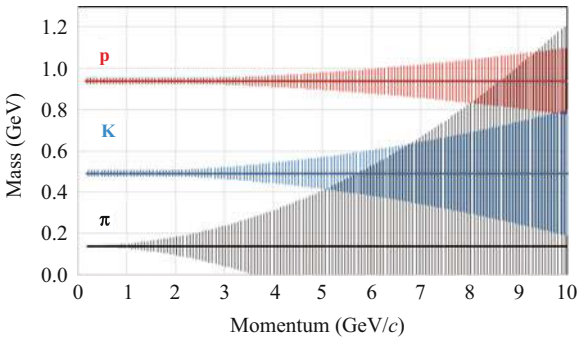


Fig. 7.2 Mass resolution as function of momentum for a Time of Flight, ToF, detector with $\Delta p/p = 4 \cdot 10^{-3}$, $l = 10\,\text{m}$, $\Delta l/l = 10^{-4}$ and $\Delta t = 50\,\text{ps}$

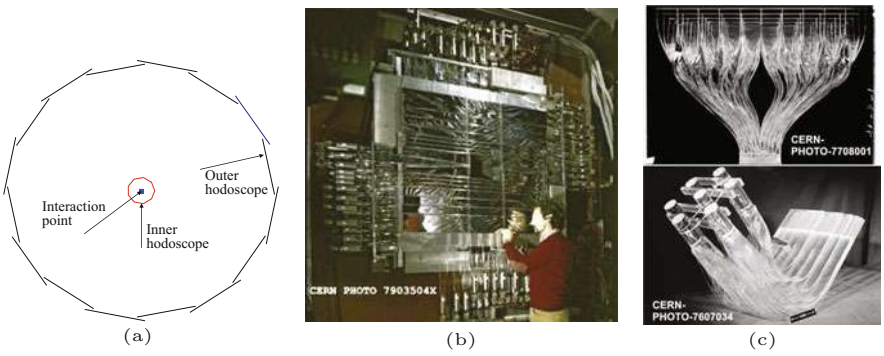


Fig. 7.3 (**a**) Simplistic sketch of a Time of Flight system. (**b**) Large scintillator hodoscope from CERN experiment NA1. (**c**) Light guides and scintillators

Particle misidentification will therefore occur when the time difference between two particles with the same momentum becomes comparable to the detector resolution. Figure 7.2 gives the mass resolution as function of momentum for π, K and proton.

Time of Flight detectors, ToF, have throughout been essential tools in physics experiments and have undergone impressive improvements in time resolution from micro-seconds to pico-seconds. The basis was worked out in [3]. A principle sketch is given in Fig. 7.3a in the Centre of Mass coordinate system. The interaction point is surrounded by a *time zero* hodoscope, the Inner hodoscope. Another hodoscope, the Outer hodoscope, is placed at a distance l from the first one. Assuming that there is a momentum measurement between the two, this is all that is needed to solve Eq. (7.2).

The Inner hodoscope is usually not required. In a colliding beam experiment, the RF structure can be adequate to give a sufficiently precise *time zero*. In events with

a large number of secondaries, one can use the feature that at least one particle will have a velocity $v \cong c$ and thereby use this one to define *time zero*.

The main work during the last years [4] has been in the improvement of the time resolution and, as the detectors have gradually increased in size, in the cost/m^2. The occupancy and radiation tolerance are playing a very important role for detectors that are proposed for the new high luminosity accelerators. We will here not explain the working principle of the detectors themselves. The reader is referred to Chap. 3. We will rather discuss the advantages and inconveniences of some of the most commonly used detector set-ups.

7.2.1 Scintillator Hodoscopes

A scintillator, read out in both ends by a photomultiplier, is the classic element of a Time of Flight hodoscope, Fig. 7.3b. The number of photons created is large. Plastic scintillators, as discussed in Chap. 3, have a density $\rho \simeq 1.03\,\mathrm{g/cm^3}$. About 10^4 photons/MeV are created with a mean wavelength of ∼400 nm and a time constant $\tau \sim 1.5\,\mathrm{ns}$. The number of emitted photons per time unit, N, will be approximately:

$$N = \frac{N_0}{\tau} \exp\left(-\frac{t}{\tau}\right) \tag{7.4}$$

N is the number of photons emitted at time t, N_0 is the total number of emitted photons and τ is the average lifetime. τ is characteristic to a specific scintillator material. A short decay time increases the maximum count rate and is therefore an important property for detection. Most inorganic scintillators have rather long decay times, $\tau \sim 100\,\mathrm{ns}$, but in some cases the decay constant can be very short. For example, $\tau = 1\,\mathrm{ns}$ for BaF_2.

The specific energy loss, Chap. 2, for a minimum ionizing particle, MIP, is given as:

$$\left(-\frac{dE}{dx}\right)_{\min} = 2.35 - 1.47\ln(Z) \quad \mathrm{MeV cm^2/g} \tag{7.5}$$

where Z is the atomic number.

$(dE/dx)_{\min}$ for a plastic scintillator is about $2\,\mathrm{MeV\,cm^2/g}$, or about $2 \cdot 10^4$ photons/cm are produced. This number of detectable photons will be greatly reduced due to the attenuation length of the material, the losses out from the material, quantum efficiency of the photon detector and the shaping time of the electronics. As the final number of photoelectrons is heavily dependent on the exact lay-out of the detector, it is very difficult to give a *typical* number. But, as a rule of thumb, approximately $2 \cdot 10^{-3}$ photoelectrons will be produced by the primary photon. This would give in the range of 40 photoelectrons/cm in a plastic

scintillator. Let N_D be the total number of detected photons. The time resolution is roughly proportional to $1/\sqrt{N_D}$. ToF detectors with high resolution, $\Delta t \sim 100$ ps, therefore use scintillator thickness of 2–3 cm. The material budget then becomes important.

The connection between the scintillator and the photon detector is a very important step in order to maximise the light collection efficiency of the system. These light concentrators are normally built around a Winston Cone [5] or a fishtail as in Fig. 7.3b. A Winston Cone is a non-imaging off-axis parabola of revolution which will maximise the collection of incoming rays. The ideal concentrator will achieve the highest possible concentration of radiant energy permitted by the second law of thermodynamics. This is equivalent to the general theorem of Liouville [6]. More specific in a case of a light guide, one can write:

$$n^2 - 1 \geq \left[\frac{d}{2r} + 1\right]^2 \tag{7.6}$$

where d is the light guide diameter and r is the bending radius. n is the refractive index relative to air. See Fig. 7.3c. Charged particles going through the light guides will give signal due to Cherenkov radiation and thereby give rise to an event correlated background.

A well designed scintillator for ToF must provide a good photon collection efficiency and a small time jitter. For fast timing one would normally rely on the first direct photon impact. This puts further constrains on the photon detector. The classic photon detector is the photomultiplier tube (PMT). Depending on the window geometry, dynode chain and HV configuration, the transient time spread is in the range of 1 ns. This can be reduced by instrumenting both ends of the scintillator and then use mean timing. This will also take care of the after-pulsing in the PMT. These are normally either due to ions in the residual gas in the PMT which drift back, strike the photo cathode and liberate new photoelectrons or light from the dynodes which hit the photo cathode. The first will give a signal about 100 ns after the event, while the latter signal comes after 30–60 ns. See Chap. 3 for more information. However, still to overcome the path length and transient time variation, the detector has to output a large amount of primary photons to achieve total time resolution in the range of 100 ps.

An example can be found in [7]. Mean timing and time slewing corrections are performed. Slew-correction time, T^{cor}, is defined as:

$$T^{\text{cor}} = T + \frac{A_0}{\sqrt{\text{ADC}}} \tag{7.7}$$

where the constant A_0 is normally evaluated for each PMT and ADC is the signal pulse height. They report a nearly constant time resolution of $\sigma \sim 55$ ps across a detector length of 15 cm.

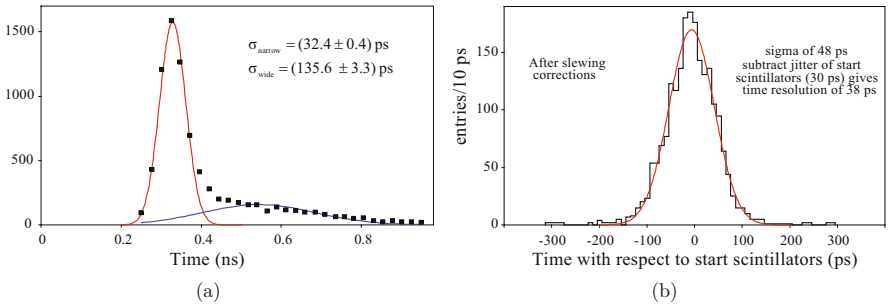


Fig. 7.4 (**a**) Single photoelectron timing resolution in Burle 64-pixel MCP-PMT 85012-501 with 10 μm hole diameter. Adapted from [10]. (**b**) Time distribution of MRPC after slewing corrections. Adapted from [11]

Other photon detectors are generally faster and with smaller time spread than the PMT. See Chap. 3 for a detailed description of these devices. Below are some listed from [8]:

- 100 μm diameter GaP SiPMT Avalanche Photo Diode operating in a Geiger mode with active quenching [9]. Single photoelectron regime: 25 ps
- Hamamatsu H-8500 Flat panel multi anode photo multiplier tube (MaPMT).[2] SLAC measurement [8] of single photon resolution: 140 ps
- Burle 85011 photo multiplier tube with micro channel plate (MCP-PMT).[3] SLAC measurement [8] of single photon resolution: <50 ps

A drawback with these detectors can be the non-Gaussian tails as shown on Fig. 7.4a.

7.2.2 Parallel Plate ToF Detectors

One of the main challenges in using gas based detectors, MWPC up to spark chambers as discussed in Chap. 3, is the time jitter caused by the spread in pulse heights due to the long Landau tail. This can to some extent be overcome by using many gaps and operating the detector in a regime where the pulse height is nearly independent of the primary ionisation. However, this can seriously diminish the rate capability of these detectors. Well adapted electronics will furthermore decrease the time walk.

[2] HAMAMATSU PHOTONICS K.K. 325-6, Sunayama-cho, Naka-ku, Hamamatsu City, Shizuoka Pref., 430-8587, Japan.

[3] BURLE INDUSTRIES, INC. 1000 New Holland Avenue, Lancaster, PA 17601-5688 U.S.A.

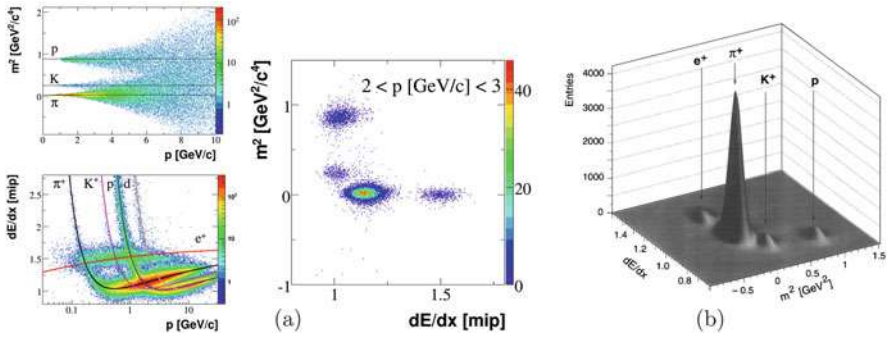


Fig. 7.5 (**a**) Particle identification in NA61. Reference [13]. (**b**) Particle identification at NA 49 by simultaneous dE/dx and TOF measurement in the momentum range 5 to 6 GeV/c for central Pb+Pb collisions. Reference [14]

Large area resistive plate chambers, see Chap. 3, are successfully used as time of flight detectors. An example is the $\sim$150 m^2, with $1.6 \cdot 10^5$ read-out channels, detector for ALICE [11]. Ten gaps of 250 μm width are made from 400 μm thick soda-lime glass with a gas composition of $C_2H_2F_4$:SF_6:C_4H_{10} = 0.90:0.05:0.05. The resistivity[4] of the glass is $\sim 10^{13}$ Ωcm. The detector is operated just below streamer mode. Tests indicate no change in performance up to 1 kHz/cm^2. As there are many gaps, the output charge distribution is a broad, but nearly Gaussian distributed with some Landau tail towards higher value. This will give rise to some time slewing. The time resolution is given as $\sigma < 40$ ps. See Fig. 7.4b.

7.3 The Power of Combined PID

The inherently simple ToF technique has greatly evolved over the years. The coming of the higher energy and/or higher intensity accelerators have required an ever better time and space resolution. Even though there has been great progress with small single pixel devices, progress with large systems has been slow. An overview of the current state of the art can be found in [12].

Combining different PID techniques, even with modest resolution, has been the preferred option for many experiments. An example of this powerful approach is shown in Fig. 7.5.

[4] It can be worth noting that materials which exhibit very large resistivity, might not be Ohmic, but rather ionic, and thereby show large variations depending on the applied current or voltage.

7.4 Cherenkov Radiation

The theory of Cherenkov radiation is discussed in Chap. 2. Further reading can be found in references [15–18]. We will here just recall some of the main features. The condition for emission of a Cherenkov photon is given by

$$\cos\Theta_C = \frac{1}{\beta \cdot \sqrt{\boldsymbol{\varepsilon(\lambda)}}} = \frac{1}{\beta \cdot n(\lambda)} \tag{7.8}$$

and the number of emitted photons by

$$\frac{d^2N}{dL\,d\lambda} = 2\pi\alpha Z^2 \frac{\sin^2\Theta_C}{\lambda^2}, \tag{7.9}$$

where Θ_C is the angle of the emitted photon relative to the particle trajectory, ε is the dielectric constant as function of the photon wavelength λ, L is the radiator length, $\alpha \sim 1/137$ is the fine structure constant, β is the particle velocity relative to the speed of light in vacuum, $\beta = v/c = pc/E$, and Z is the charge of the particle in units of electron charges. The refractive index, n, is given as $n^2 = \varepsilon$. The relationship between the photon wavelength and its angular frequency, ω, is given by $\lambda(\text{nm}) \sim 1240/\hbar\omega(\text{eV})$. A representation of the Cherenkov radiation domain is given in Fig. 7.6a.

From the discussion in Chap. 2 and Eq. (7.8) it is clear that ε has to be real and larger than 1 and that the speed of the charged particle must be larger than the phase velocity of the electromagnetic fields at the frequency ω in order to have emission of Cherenkov photons at that frequency.

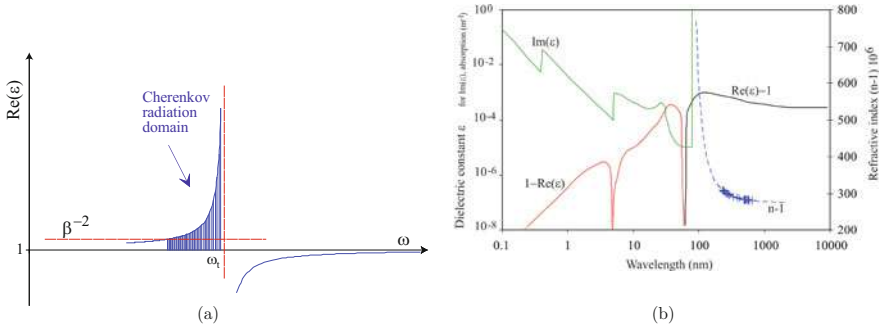


Fig. 7.6 (**a**) Simplistic representation of the real part of the dielectric constant, $\Re(\varepsilon)$, as function of the frequency, ω. (**b**) The dielectric constant, ε, and the refractive index, n, for argon at 0 °C and 101.3 kPa. ε data replotted from [19] and n from [20]

We see from the above that Cherenkov radiation is characterized by:

- Cherenkov radiation is a prompt signal.
- The existence of a threshold[5] in $\beta_{\min} = n^{-1}$
- The Cherenkov angle is depending on β.
- The number of Cherenkov photons emitted is depending on β.
- The number of photons emitted is depending on the square of the charge of the particle.

The properties described above of Cherenkov radiation can be used to measure the velocity of a charged particle traversing matter. Consider two charged particles with known momenta p and mass and velocity given by m_i and β_i. The mass difference can then be written as:

$$m_1^2 - m_2^2 = p^2 \cdot \frac{(\beta_1 - \beta_2)(\beta_1 + \beta_2)}{(\beta_1 \cdot \beta_2)^2} = n^2 p^2 \cdot (\cos^2 \Theta_1 - \cos^2 \Theta_2) \tag{7.10}$$

And if $n - 1$ is small

$$m_1^2 - m_2^2 = p^2 \cdot (\Theta_2 + \Theta_1)(\Theta_2 - \Theta_1) \tag{7.11}$$

The resolution in mass is thereby directly linked to the angular resolution of the detector. The main emphasis for all the Cherenkov detectors will be angular resolution.

The refractive index together with ε, for argon at 0 °C and 101.3 kPa, is given in Fig. 7.6b. The data for the refractive index of argon is well described by a single pole Sellmeier, see Eq. (7.16), representation:

$$(n - 1) \cdot 10^6 = \frac{0.05086}{73.82^{-2} - \lambda^{-2}} \tag{7.12}$$

with λ in nm. We observe that this pole is where $\Re(\varepsilon)$ goes from larger than 1 to smaller than 1. At about the same wavelength $\Im(\varepsilon)$ becomes important.

A Cherenkov light detector is therefore based on classical optics. The choice of radiator, and thereby the refractive index, is depending on the momentum range which has to be covered and the photon detector option. We will in the following discuss different radiator materials, Sect. 7.4.2, and the usage from Threshold, Sect. 7.4.3, to Ring Imaging Cherenkov detectors, Sect. 7.4.4. We will first take a closer look at the refractive index, Sect. 7.4.1.

[5] Due to diffraction broadening, Cherenkov photons can be emitted below threshold. We will not discuss that here.

Table 7.1 Atomic refraction constants from Ref. [22]

Atom		Atomic refraction
Carbon		2.418
Bromine		8.865
Chlorine		5.967
Fluorine		1.1
Hydrogen		1.1
Iodine		13.952
One double bond	=O	2.122
Two single bonds	−O−	1.643

7.4.1 Refractive Index

The dielectric constant is given by:

$$\varepsilon = 1 + 4\pi\chi = \frac{1 + \frac{8}{3}\pi \mathrm{N}\zeta}{1 - \frac{4}{3}\pi \mathrm{N}\zeta} \quad \text{from which} \quad \frac{4}{3}\pi \mathrm{N}\zeta = \frac{\varepsilon - 1}{\varepsilon + 2} \tag{7.13}$$

where χ is the susceptibility, N is the number of molecules per unit volume and ζ is the molecular polarizability.

A relation like this was first obtained by Mossotti in 1850, then by Lorenz in 1869, and refined by Clausius in 1879, and which is usually called the Clausius-Mossotti equation. Polarizable matter was modelled as an assembly of small conducting spheres in the early studies.[6] Maxwell's theory showed that the index of refraction of light, n, was related to ε by $n^2 = \varepsilon$, so that the formula could be applied to light as well as to static fields. H.A. Lorentz, in 1878, and L.V. Lorenz (1829–1891), in 1881, derived a similar formula on the basis of the electron theory in which n^2 replaced ε. This formula is called the Lorentz-Lorenz formula, and can be written in the following way:

$$n^2 = \frac{1 + 2\left(\rho \frac{R_M}{M_W}\right)}{1 - \left(\rho \frac{R_M}{M_W}\right)} \tag{7.14}$$

where R_M is the molar refraction, M_W is the molecular weight and ρ is the density.

The molar refraction may then be estimated from the chemical formula. Atomic refraction constants differ slightly in the literature, but the constants in Table 7.1 give reasonable results for many compounds.

The Lorentz-Lorenz equation, Eq. (7.14) together with Table 7.1, does not explicitly express the refractive index as a function of the photon energy.

[6] Strictly speaking, Clausius-Mossotti equation is only rigorously valid in the limit of zero density [21].

The most common way to represent the refractive index is in the form of a series with multiple poles

$$n-1=C\cdot\sum_i\frac{f_i}{\nu_i^2-\nu^2}\quad\text{with}\quad C=\frac{e^2A}{2\pi mc^2}=1.2098\cdot10^6 \tag{7.15}$$

where e and m are the charge and mass of the electron, A is Avogadro's number per cm^3 and $\nu(\mathrm{cm}^{-1})=10^7/\lambda(\mathrm{nm})$. f_i is the oscillator strength of the Eigen frequency ν. We will here mainly use the standard Sellmeier formula with one pole:

$$\frac{3}{2}\cdot\frac{n^2-1}{n^2+2}=\frac{a}{\lambda_0^{-2}-\lambda^{-2}}\simeq n-1\quad\text{for}\quad n-1\ll1 \tag{7.16}$$

$b=\lambda_0^{-2}$ will also be used. λ is in nm. a/b is the asymptotic value of n as $\lambda\to\infty$. A two pole Sellmeier representation might be required:

$$\frac{3}{2}\cdot\frac{n^2-1}{n^2+2}=\left[a_1\cdot\lambda^{-4}+a_2\cdot\lambda^{-2}+a_3\right]^{-1} \tag{7.17}$$

Clearly also other types of power series can be used to approximate the refractive index like in reference [23]. In this case the refractive index is approximated with the half empirical formula of a n-term Cauchy equation which is very similar to Eq. (7.17):

$$n-1=2\pi N_0\left[a_0+a_1\omega^2+a_2\omega^4+a_3\omega^6\right] \tag{7.18}$$

where ω is the frequency in atomic units. The $a_3\omega^6$ term has been added after the original series [24] was truncated at $a_2\omega^4$ and thereby was not very useful in the UV to VUV wavelength region.

The refractive index of a medium M which is a mixture of different molecules in the ratio $M=\sum_i[M_i/f_i]$ for $1=\sum_i f_i^{-1}$, is given by $n_M=\sum_i[n_i/f_i]$. We will illustrate this with a simple example. The refractive index of air and its constituents are well measured quantities, Fig. 7.7a.

The Sellmeier parameterisation for N_2, O_2, CO_2 and argon is given in Table 7.2. Note that whereas a single pole, Eq. (7.16), describes well N_2, CO_2 and argon, the data for O_2 is best described with a two pole, Eq. (7.17), representation. The parameters used to describe the data points for dry air in Fig. 7.7a are

$$\begin{aligned}n(\mathrm{air})&=0.7809\cdot n(\mathrm{N_2})+0.2095\cdot n(\mathrm{O_2})+0.0093\cdot n(\mathrm{Ar})+0.0003\cdot n(\mathrm{CO_2})\\&\simeq1+10^{-6}\cdot\left[\left(\lambda^{-2}-69.1^{-2}\right)\left(\lambda^{-2}+99.5^{-2}\right)\right]^{-1}\end{aligned} \tag{7.19}$$

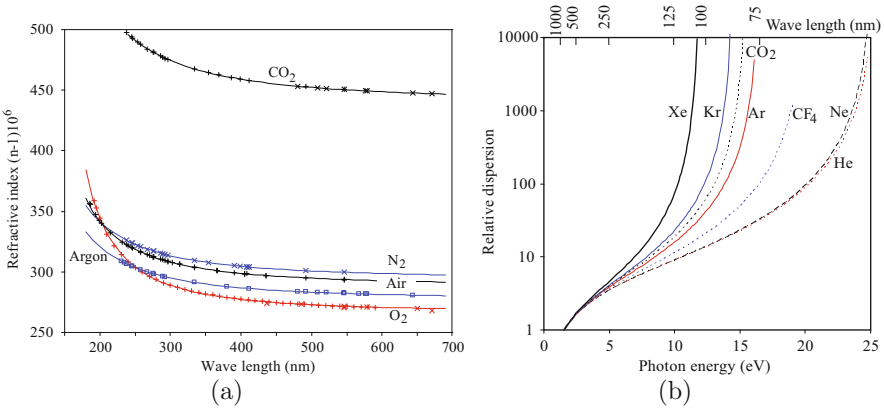


Fig. 7.7 (**a**) The refractive index of dry air, N_2, O_2, CO_2 and argon at 0 °C and 101.3 kPa [20]. (b) Dispersion dn/dE relative to the value at 800 nm, in some noble and n-atomic gases as function of the photon energy [20]

Table 7.2 Sellmeier fit, Eqs. (7.16) and (7.17), parameters for the gases at 0 °C and 101.3 kPa

Gas	A	B	a_1	a_2	a_3	λ_0
N_2	0.0532	0.000181				74.36
O_2			−54,955	−20.275	0.00376	122.90
CO_2	0.0687	0.000156				80.10
Ar	0.0509	0.000184				73.82

The pole, λ_0, in nm. O_2 has only one real pole in this representation

Although the last expression gives a good description of the refractive index for dry air at 0 °C, 101.3 kPa and for $\lambda \geq 130$ nm, the real pole at ∼69 nm has no physical meaning.

7.4.2 Cherenkov Radiators

Cherenkov radiators have to be reasonably optically transparent and with an appropriate refractive index. The scintillation and phosphorescence processes in the medium should be small. There is a wide variety to chose from, from transparent solids via liquids to gases. One can in addition change the refractive index by changing temperature and pressure of the medium.

The dispersion in a radiator can be written as

$$\frac{dn}{dE} \propto \frac{(n^2-1)^2}{n} \cdot E \quad \text{and for } (n-1) \ll 1 \text{it reduces to} \quad \frac{dn}{dE} \propto (n-1)^2 \cdot E \tag{7.20}$$

where E is the energy of the photon.

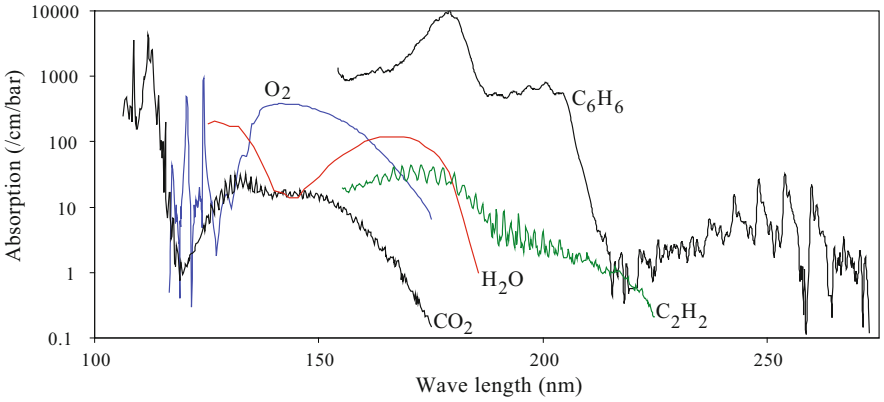


Fig. 7.8 Absorption as function of the photon wavelength [25]

The dispersion in some noble and n-atomic gases is plotted in Fig. 7.7b. He and Ne are very weakly dispersive in contrast to Kr and Xe. As can be seen from Fig. 7.7b, fluorocarbons are also weakly dispersive. If the definition of the Cherenkov angle is an important quantity for the detector, it is then clear that the dispersion has to be as small as possible over the photon detector efficiency window. The detector design will be a balance between number of photons and the spread in the Cherenkov angle.

The radiator medium becomes opaque when the imaginary part of the dielectric constant becomes important, Fig. 7.6b. Most media will in addition exhibit broad and strong absorption bands. Figure 7.8 shows the absorption in some commonly used Cherenkov media or trace impurities in them. For simple alkanes, C_NH_{2+2N}, the onset of photon absorption [25] can be approximated to:

$$\lambda_{CH}(\text{nm}) = 181 - \frac{226}{3(N+1)} \tag{7.21}$$

A similar approximation can be given for n-perfluorocarbons, C_NF_{2+2N},:

$$\lambda_{CF}(\text{nm}) = 175.6 - \frac{641}{3N+5.7} \tag{7.22}$$

It can be seen from these two expressions that n-perfluorocarbons are more transparent than alkanes. Alkanes are therefore good quenchers as used in MWPCs. Trace impurities are particularly difficult to eliminate especially when it is not clear which molecule is causing the absorption. The successful detector design should therefore not be sensitive to these bands.

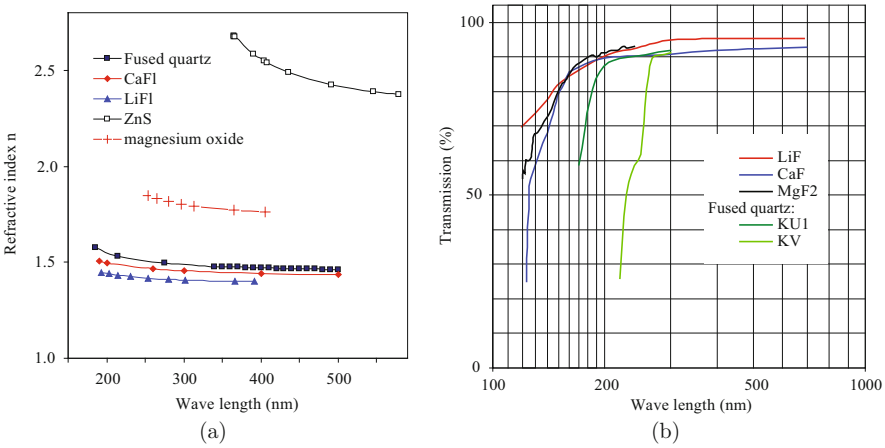


Fig. 7.9 (**a**) Refractive index for quartz and other special optical materials [26]. (**b**) Transmission in some commercially available quartz as function of wavelength. See footnote 7

7.4.2.1 Quartz Radiators

Quartz radiators are very popular for Cherenkov detectors operating in the low momentum range. The refractive index for some quartz and other optical materials is given in Fig. 7.9a. Figure 7.9b gives the transmission for some commercially[7] available quartz. By choosing a refractive index $n \sim 1.5$ and a photon detection window from 800 to 300 nm, the Cherenkov angle measurement between π and K becomes difficult for $p > 2\,\text{GeV}/c$ due to dispersion.

Quartz radiators in Cherenkov detectors are treated in two distinctly different ways. We see that for the $n \sim 1.5$ quartz, a π will pass the Cherenkov threshold at 125 MeV/c and at ∼280 MeV/c no light will escape the quartz due to total internal reflection for particles perpendicular onto the radiator. An elegant solution to this problem is shown in Fig. 7.10a.

The other option is to exploit the feature of internal reflections as a light guide for the Cherenkov photons. The working principle of a DIRC, Detection (of) Internally Reflected Cherenkov (light) [28], detector is sketched in Fig. 7.10b. The standoff region is designed to maximize the transfer efficiency between the radiator and the detector. If this region has the same index of refraction as the radiator, $n_1 \simeq n_2$, the transfer efficiency is maximized and the image will emerge without reflection or refraction at the end surface. Further improvements can be achieved by measuring the transfer time of the Cherenkov photons [29]. A large fraction of the uncertainties caused by the dispersion can then be eliminated.

[7] Data from:
Del Mar Ventures, 12595 Ruette Alliante No.148, San Diego, California 92130, US.
Crystran Ltd, 1 Broom Road Business Park, Poole, Dorset, England.

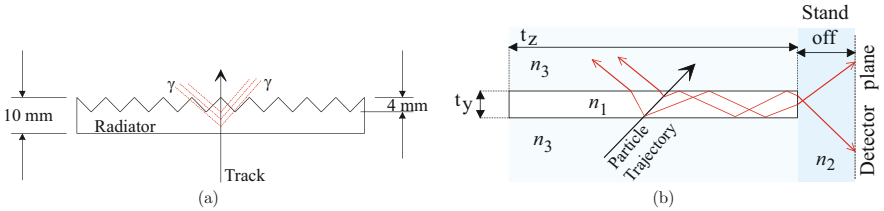


Fig. 7.10 (**a**) Sketch of the saw tooth quartz radiator for CLEO 3 [27]. (**b**) Schematic of the radiator bar for a DIRC [28] detector. Not to scale

Similar, but not identical, are the Time-of-Propagation, TOP [30] detector at the BELLE II experiment and the proposed detectors; TORCH [32] at LHCb and a DIRC [33] at the PANDA experiment.

The TOP consists of quartz radiator bars 270 cm long $\times$45 cm wide $\times$2 cm thick. See Fig. 7.11a. One end of the bar has a spherical mirror to reflect light back to the other end that has a small expansion prism. The prism is instrumented with 32 16-channel microchannel plate photomultiplier tubes (MCP-PMTs) readout with custom giga-sample per second waveform sampling electronics. The Cherenkov ring is imaged by the 512 MCP-PMT pixels with 5 mm pixel size and the time of arrival of each photon is measured with $<$50 ps timing resolution. The photon time of arrival is a sum of the time of flight of the charged particle to the quartz radiator and the time of propagation of the Cherenkov photons to the photodetectors. Results with test beam data is shown in Fig. 7.11b. Clear π/K separation can be observed. The detector will be ready for data taking in 2018.

7.4.2.2 Aerogel Radiators

The search for a stable Cherenkov radiator with a refractive index between gas and liquid started about the same time as the first Cherenkov detector became operational. The first successful was silica aerogel [34]. The Axial Field Spectrometer [35] at the CERN ISR was the first large experiment to use it. The principle fabrication reactions[8] are rather simple:

$$\begin{aligned} \mathrm{Si(OCH_3)_4 + 4H_2O \xrightarrow{NH_3} Si(OH)_4 + 4CH_3OH} \\ \mathrm{nSi(OH)_4 \longrightarrow (SiO_2)_n + 2nH_2O} \end{aligned}$$

The refractive index, n, as a function of the wavelength, λ, can be approximated by $n = 1 + k \cdot \rho$ for the density ρ in g/cm^3 and k is a function of λ. An example

[8] Tetramethyl orthosilicate is used in this example. Tetraethyl orthosilicate can also be used and is normally preferred as the byproduct is ethanol rather than methanol. Both tetramethyl orthosilicate and tetraethyl orthosilicate are highly reactive [36].

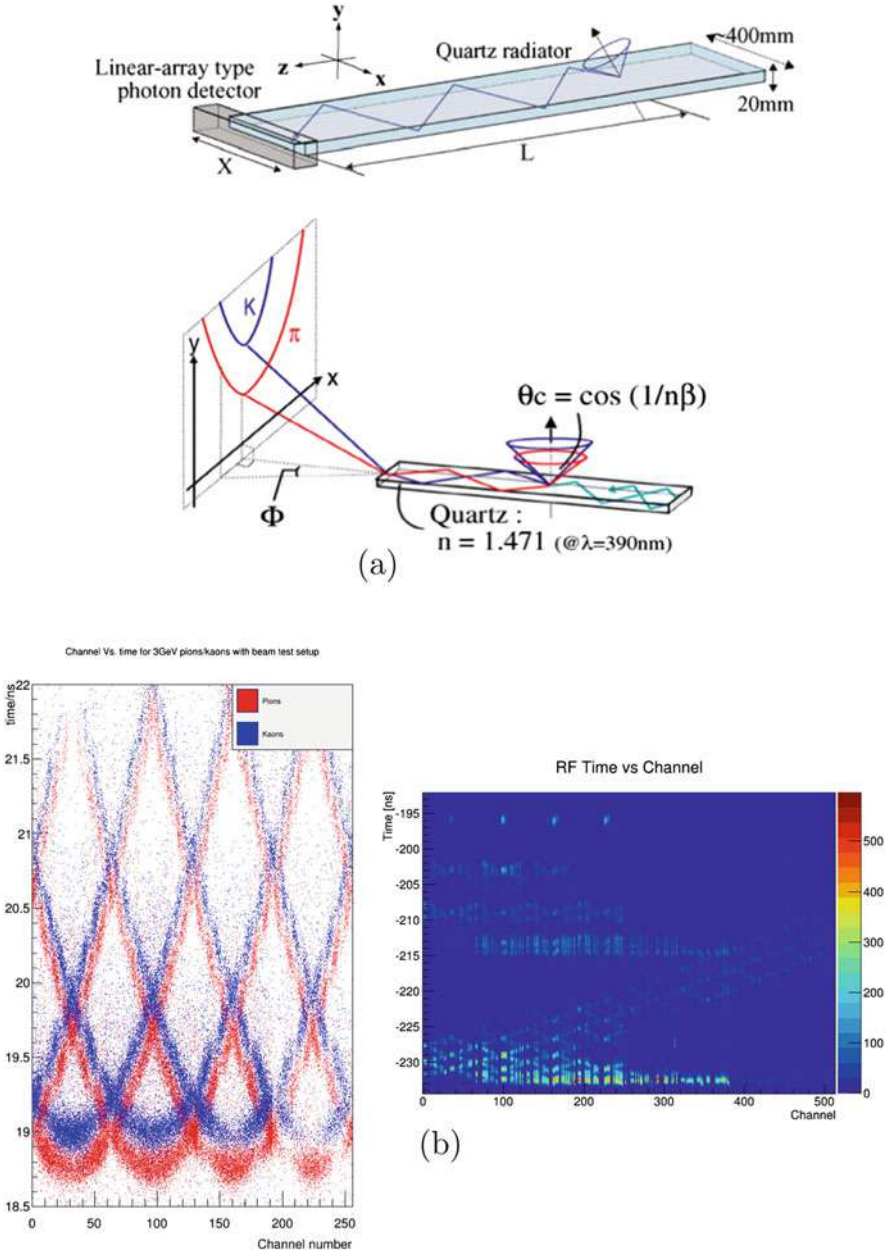


Fig. 7.11 (**a**) Schematic drawing of a TOP-counter. Reference [30]. (**b**) Test beam data. Cumulative distribution of measured time versus channel number for detected photons; the insert shows a zoom at short times, indicating the separation in time between the signals from pions and kaons. Reference [31]

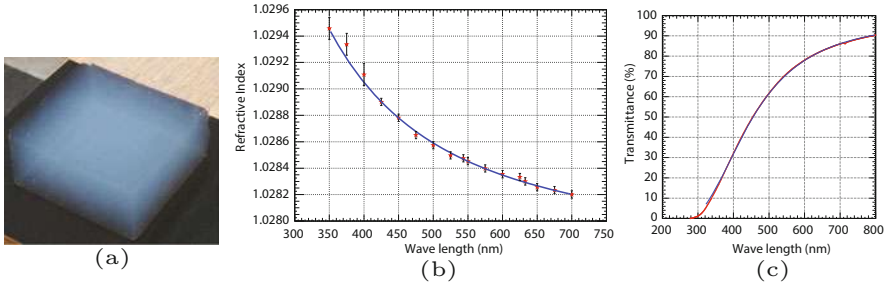


Fig. 7.12 (**a**) Aerogel tile. Courtesy of the LHCb Milano group. (**b**) Refractive index of aerogel as function of wavelength. Bellunato et al. [37] with permission. (**c**) Transmittance of 52.10 mm thick aerogel as function of wavelength. Perego [38] with permission

is shown in Fig. 7.12b. The data [37] is well described by a single pole Sellmeier equation:

$$n^2 - 1 = \frac{a_0 \lambda^2}{\lambda^2 - \lambda_0^2} \tag{7.23}$$

for $a_0 = 0.05639 \pm 0.00004$ and $\lambda_0 = (83.22 \pm 1.25)$ nm.

Assuming that aerogel is just a rarefied form of silica, a_0 and the density of the material are linked by:

$$\rho(\text{aerogel}) = \frac{a_0(\text{aerogel})}{a_0(\text{SiO}_2)} \frac{n^2(\text{SiO}_2) + 2}{n^2(\text{aerogel}) + 2} \rho(\text{SiO}_2) \tag{7.24}$$

which gives $\rho(\text{aerogel}) = (0.158 \pm 0.001)\,\text{g/cm}^3$, in reasonably good agreement with $\rho = (0.149 \pm 0.004)\,\text{g/cm}^3$ which was given by the manufacturer.

Two main types of aerogel are now available, hydrophobic[9] and hygroscopic.[10] Large homogeneous blocks of high optical quality are now readily available. The refractive index can be tuned between 1.008 and 1.1. By stacking aerogel blocks of different refractive indices, the total light output can be increased while minimizing the width of the Cherenkov ring. By modifying the reaction conditions of the sol-gel synthesis [39], it is possible to control the variations of n inside the aerogel tile and thereby create a monolithic block with well defined different layers of n.

[9] Advanced Technology Research Laboratory, 1048 Kadoma, Kadoma-shi, Osaka-fu, Japan 571.

[10] Boreskov Institute for Catalysis in collaboration with the Budker Institute of Nuclear Physics in Novosibirsk.

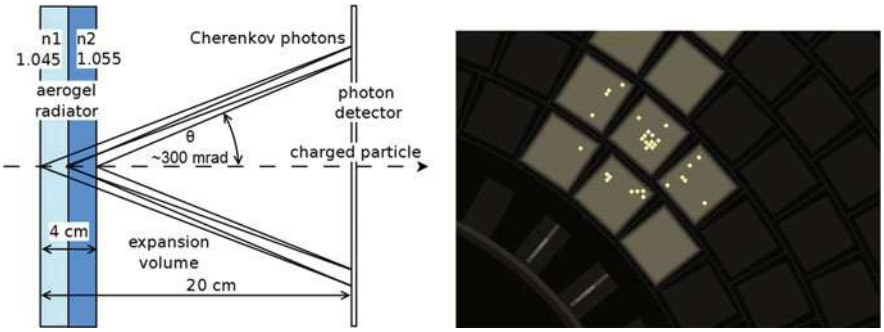


Fig. 7.13 (**a**) Proximity focusing RICH with two layers of the aerogel radiator: Cherenkov photons emitted in two aerogel tiles are detected on the same ring by the position sensitive photon detector, thus reducing the ring width. (**b**) Cosmic ray events registered by partially equipped detector. Reference [41]

The optical quality, light transmission T, see Fig. 7.12c, of aerogel is normally parameterized as:

$$T = T_0 \cdot \exp\left[-C \cdot \frac{t}{\lambda^4}\right] \tag{7.25}$$

where C is the clarity given in $\mu m^4/cm$ and t is the thickness in cm. T_0 describes the bulk properties of the aerogel and C the variation with the wavelength. The λ^4 term shows that the light attenuation, opacity κ,[11] is governed by Rayleigh scattering [40] which can be written as:

$$\kappa = \sigma_{\mathrm{Rlh}}(\lambda) \cdot N_0 \cdot t\,, \quad \text{where } \sigma_{\mathrm{Rlh}}(\lambda) \cong \frac{128\pi^5\zeta^2}{3\lambda^4} \cdot \frac{6+3\delta}{6-7\delta} \text{ for } \zeta = \frac{n-1}{2\pi N_0} \tag{7.26}$$

where N_0 is the number of particles per unit volume and δ is the polarization factor. δ is small and in the range from about 0.03 to 0.09.

The two-layer aerogel RICH detector of the Belle II spectrometer [41] will separate charged particles in the forward end-cap of the spectrometer inside a magnetic field of 1.5 T with a high separation capability in the momentum range from 0.5 to 3.5 GeV/c. See Fig. 7.13. The detector will be ready for data taking in 2018.

[11] Opacity is another term for the mass attenuation coefficient or, depending on context, mass absorption coefficient. κ_λ at a particular wavelength λ of the electromagnetic radiation.

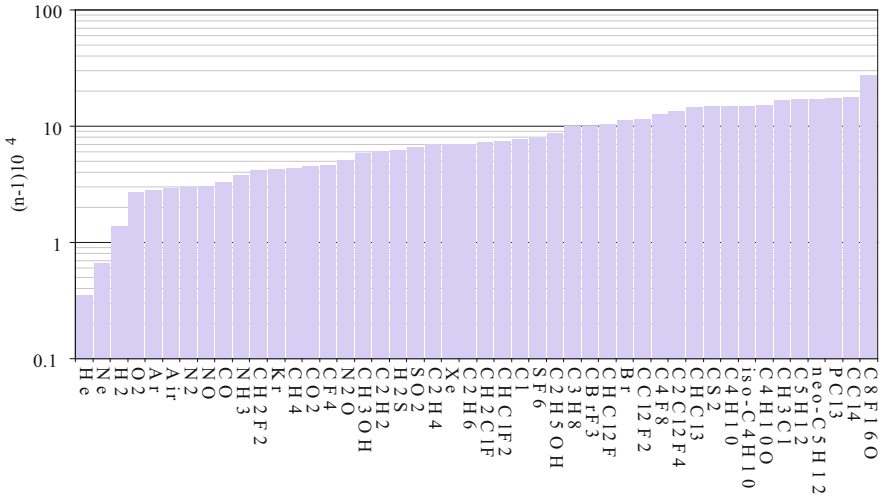


Fig. 7.14 Refractive index for some common fluids. D-line (589 nm). Data from [16, 22]

7.4.2.3 Fluids as Radiators

The relationship between the refractive index of a gas and the corresponding liquid, is given by:

$$\left[\frac{n^2-1}{n^2+2}\right]_{\text{gas}} = \left[\frac{p}{RT}\right]_{\text{gas}} \left[\frac{M}{\rho}\right]_{\text{liq}} \left[\frac{n^2-1}{n^2+2}\right]_{\text{liq}} \tag{7.27}$$

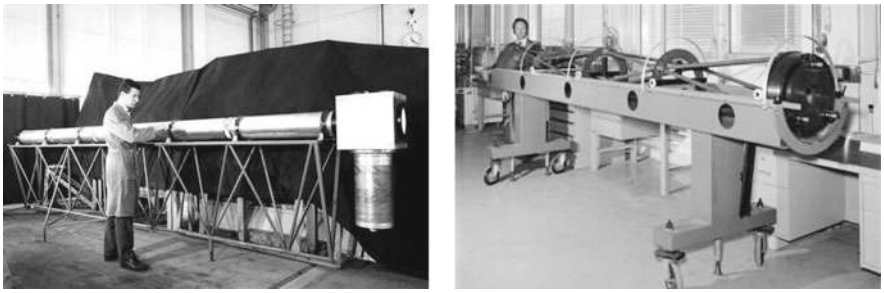
where p and T is the pressure and temperature of the gas, M and ρ is the molecular weight and density of the liquid and R is the gas constant (based on pressure and volume units $R = 82.0575$ (cm^3 atm)/(K mol)).

The refractive index for a number of fluids is plotted in Fig. 7.14.

7.4.3 Threshold Cherenkov Detectors

As soon as photon detectors, Chap. 3, coupled with the associated electronics, had the sensitivity to detect the low level of photons emitted through Cherenkov radiation, the first threshold Cherenkov detectors, see Figs. 7.15 and 7.16, were used in high energy experiments. The best known of these early experiments, is probably the discovery of the antiproton at the Radiation Laboratory of the University of California at Berkeley in 1955 [42].

The design of these threshold detectors is simple as is shown in Fig. 7.16a. In this sketch, the radiator is a gas. There is no problem to change it by inserting



(a) (b)

Fig. 7.15 (**a**) A threshold gas Cherenkov counter as used to tag particles in the secondary beams. CERN IT 6304088. (**b**) CEDAR counter (internal part). Here on the mounting bench. The counter is a differential Cherenkov, corrected for chromaticity, able to differentiate pions from kaons up to 350 GeV. Counters of this type were used in all SPS hadron beams. CERN PHOTO 7603033

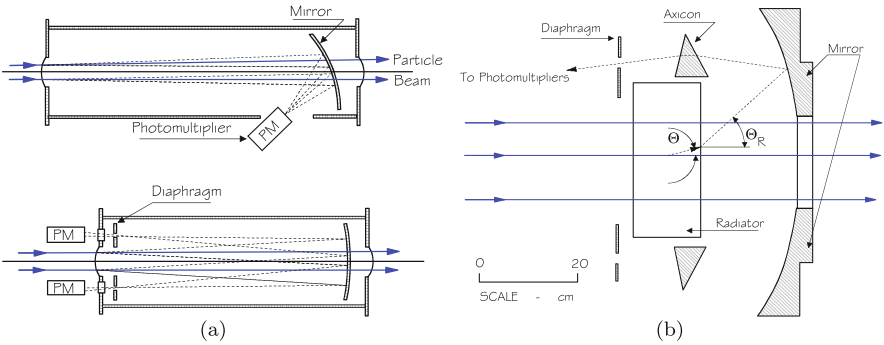


Fig. 7.16 (**a**) Top and bottom shows the working principles of respectively a threshold and a differential Cherenkov detector. (**b**) Is an achromatic liquid differential Cherenkov detectors, DISC: Differential Isochronous Self-Collimating; adapted from [43]

a solid or a liquid radiator, nor to change the pressure of the gas. It only affects the radiation length seen by the traversing particles. The solid angle covered by the detector is only limited by the design of the optics. A threshold Cherenkov detector can therefore be used both in the incoming beam to define the flavour of the primary particles as well as for identifying the secondaries. It should be noted that by introducing two, or more, detectors in series, positive particle identification can be achieved over a large momentum range.

A differential Cherenkov detector is shown in Fig. 7.16a. It is designed for a given value of the Cherenkov angle, such that:

$$\Theta = r/F \tag{7.28}$$

where r is the mean radius of the aperture of the diaphragm and F is the focal length of the mirror. The use of these detectors is mainly limited to parallel beams.

Assuming high energy particles and gas radiator, the resolution power can be written as:

$$\left[\frac{\Delta\beta}{\beta}\right]_{\text{limit}} = \tan\Theta \cdot \Delta\Theta \tag{7.29}$$

The coma[12] is the main error, given by:

$$\Delta\Theta_{\text{coma}} = \Theta^3 + \frac{\Theta^2}{4}\left[3\frac{b}{L} - \Theta\right] = \frac{3}{4}\Theta^3 \quad \text{if} \quad b \ll L \tag{7.30}$$

where b is the diameter of the incoming particle beam and L is the length of the gas radiator. The chromatic angular dispersion is given by:

$$\Delta\Theta_{\text{chrom}} = \frac{\Theta}{2\nu}\left[1 + \frac{1}{\gamma^2\Theta^2}\right] \quad \text{where} \quad \nu = \frac{n(\lambda_2) - 1}{n(\lambda_1) - n(\lambda_3)}, \tag{7.31}$$

representing the optical dispersion in the gas. λ_1 and λ_3 are the wavelengths appropriate for the limits of the spectral range. λ_2 is the mean wavelength. The total angular dispersion is then:

$$\Delta\Theta \approx \Theta^3 + \frac{\Theta}{2\nu}\left[1 + \frac{1}{\gamma_i^2\Theta^2}\right], \tag{7.32}$$

$i = 0, 1$ depending on the particle. We then get the limit for the maximum Cherenkov angle:

$$\Theta^4 + \frac{\Theta^2}{2\nu} \leq \frac{1}{2p^2}\left[m_1^2 - m_0^2 - \frac{m_i^2}{\nu}\right]. \tag{7.33}$$

For most applications, the Cherenkov angle will be smaller than this limit. The design will therefore be governed by the chromatic error.

To further diminish the errors, and thereby minimize $\Delta\beta/\beta$, a Differential Isochronous Self-Collimating, DISC, Cherenkov detector can be used. See Fig. 7.16b. With an optimized optics design a nearly achromatic condition can be achieved. That is,

$$\frac{\Delta\Theta(\lambda)}{\Delta\lambda} = 0 \quad \rightarrow \quad \frac{\Delta\beta}{\Delta\lambda} = 0 \tag{7.34}$$

[12] The aberration known as *coma* affects rays from points not on the axis of a lens. It is similar to spherical aberration in that both arise from the failure of the lens to image central rays and rays through outer zones of the lens at the same point. Coma differs from spherical aberration in that a point object is imaged not as a circle but as a comet-shaped figure (whence the term *coma*).

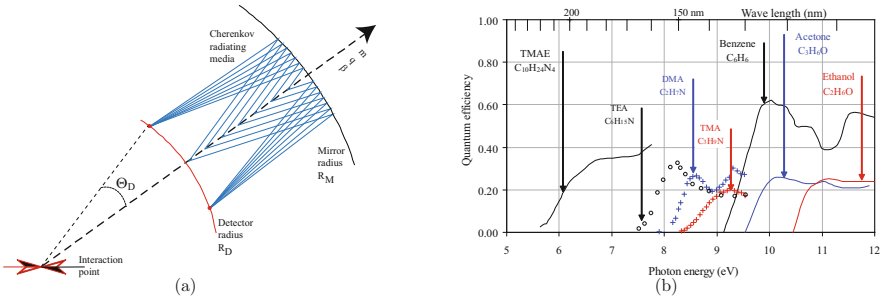


Fig. 7.17 (**a**) Ring imaging optics for particles emerging from a target or interaction region with zero impact parameter. The detected and emitted Cherenkov angles (Θ_D, Θ) are equal if the detector radius is correctly chosen.[45]. (**b**) The quantum efficiency for some photo sensitive vapours as function of photon energy. Adapted from [17]

Velocity resolution $\Delta\beta/\beta \sim 10^{-6} - 10^{-7}$ has been achieved [44]. These are very beautiful detectors, but with a somewhat limited usage as they require a near parallel beam, offer a limited solid angle and the material budget is not negligible.

7.4.4 Ring Imaging Cherenkov Detectors

The quest to make a ring imaging detector and thereby utilize all the inherent properties of Cherenkov radiation as described in Sect. 7.4, was long thwarted by the inability to get a high spatial resolution photon detector which was sensitive to single photons and compatible with photon absorptions, Fig. 7.8, in the media and photon transmission through windows, Fig. 7.9b. The breakthrough came in 1977 with the work of J. Séguinot and T. Ypsilantis [45, 46]. See Fig. 7.17a. Their work during the initial phase was mainly concentrated around MWPC, Chap. 4, and a photoionizing vapour additive to the chamber gas.

7.4.4.1 Photo Sensitive Vapours

Figure 7.17b shows the quantum efficiency for some photo sensitive vapours. The work with TEA,[13] Triethylamin $C_6H_{15}N$, and especially TMAE,[14] Tetrakis-(dimethylamino)-ethylene $C_{10}H_{24}N_4$ [47, 48], made it possible to work in the wavelength range from about 200 to 160 nm and thereby use fused silica as windows.

[13] http://webbook.nist.gov/cgi/cbook.cgi?ID=C121448&Units=SI.
[14] http://webbook.nist.gov/cgi/cbook.cgi?ID=C996703&Units=SI.

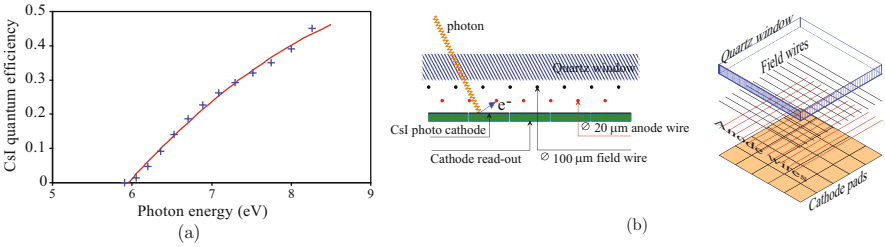


Fig. 7.18 (**a**) CsI quantum efficiency. (**b**) Sketch of a MWPC with CsI photo cathode

TMAE was the chosen photoionizing vapour, together with drift chambers, Chap. 3, for the first generation RICH detectors [49–51]. However, these fluids are difficult to handle and their usage is therefore now very limited. TEA and TMAE are chemically not reactive with respect to normal MWPC gases. They will, however, require an O_2 and water content of the carrier gas $\leq$ 10 ppm for stable operation. A drawback by using these molecules is the photon feedback. The photons created in the gas amplification process have a probability to convert. The main source of this background is from the ionization due to the charged particle going through the detector. The chambers were normally run at an amplification around $1 - 5 \cdot 10^5$ in order to be sensitive to single photons. The total probability for re-conversion thereby became larger than 1 and the chamber would break down. The number of feed-back photons can be written as $N_{\text{fp}} = \iota \cdot G$ where G is the total chamber gain.[15] $\iota \sim 7 - 8 \cdot 10^{-6}$ in CH_4 due to photon absorption for wavelengths below 143 nm. See Eq. (7.21).

A number of ingenious chamber designs were made to minimize the photon feed-back. The designs are a compromise between detection efficiency, ease of operation and fabrication and drift of electrons in a $\mathbf{B}\times\mathbf{E}$ configuration. Even at stable operating conditions, some photons will escape and give rise to an event correlated background. This background is difficult to disentangle from the real signal in high occupancy events and particularly with TMAE due to its long photon conversion length.

7.4.4.2 CsI Photo Cathode

The next step in high spatial granularity, or pixilated, photon detectors for RICH came with the CsI photon detector [52]. CsI is an alkali halide crystal which has a good quantum efficiency, Fig. 7.18a, below 200 nm and is stable in normal dry and O_2 free chamber gases [53]. The development was triggered by the need for a faster detector at the arrival of LHC and similar accelerators. In a MWPC structure,

[15] The measured chamber gain might be smaller due to charge sharing and electronics time constants.

the CsI photo cathode can either be deposited as a reflective, Fig. 7.18b, or as a semi-transparent layer [54]. The latter would, in the case of Fig. 7.18b lay-out, be a layer on the quartz window. The maximum quantum efficiency for semi-transparent CsI is for a thickness of about 11 nm in the wavelength range from 210 to 170 nm. The thickness does not matter for a reflective photo cathode and is normally in the range of 150–200 nm. A semi-transparent CsI photo cathode will have a quantum efficiency of about 0.7 compared to a reflective one. It should be noted that the photon conversion efficiency is strongly depending on the bulk structure and morphology of the CsI layer; that is, the roughness of the substrate and the connectivity of the layer. Particularly thin layers can become a collection of unconnected islands. Post-production heat treatment has proven advantageous.

As for the photosensitive vapours, a CsI photo cathode will be sensitive to the photon feed-back from the gas amplification process, see Sect. 7.4.4.1. A stable operation of the chamber is therefore a compromise between single photon efficiency, electronics sensitivity, signal shaping and gas amplification.

As few, if any other photon detector, can beat a gas based detector in cost efficiency and geometrical acceptance, a number of similar, but not identical, detector set-ups are proposed and investigated. The main emphasis is on limitation of photon feed-back, on better and more stable photo cathodes and on time resolution. This work is also partially driven by very large Cherenkov detectors for astrophysics. A very promising research and development is in gaseous micro pattern detectors with Bialkali photo cathodes. We will not discuss these here, but refer the reader to [55]. An overview of the current status and perspectives of gaseous photon detectors can be found in [56].

7.4.4.3 Vacuum Based Photon Detectors

The working principles of vacuum based photon detectors like photo multiplier tubes are discussed in Chap. 3. Although small diameter PM tubes, diameter 10 mm upwards, have been used for a long time in Cherenkov detectors, cost, balanced with space resolution and material budget, made them less attractive. The introduction of multi anode and pixilated silicon anode detectors, together with fast and sensitive electronics changed this. The first generation of multi anode photo tubes required a lens system [57] in order to give good geometrical acceptance. See Fig. 7.19a.

The schematic of a Hybrid Photon Detector [60, 61], HPD, is shown in Fig. 7.19b. In these detectors the encapsulated pixilated silicon detector is bump-bonded onto the read-out electronics. The capacitance is thereby small and the associated noise low. It also requires only a few vacuum feed-throughs. The photo cathode is normally a S20.[16] Under the influence of the electric field, the photo-electron is accelerated onto the silicon detector. In the example given in Fig. 7.19b, the 20 kV potential between anode and cathode gives a cross-focusing field with a

[16] S20 is a tri-alkaline (Sb-Na-K-Cs) semi-transparent photo cathode.

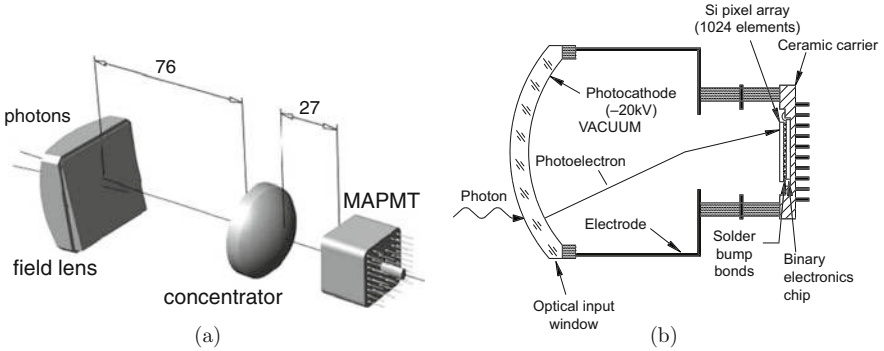


Fig. 7.19 (**a**) Optical arrangement of the COMPASS MAPMT and the fused silica lens telescope. With permission [58]. (**b**) Schematic arrangement of the LHCb Hybrid Photon Detector. With permission [59]

demagnification of ∼5. Other field configurations can be used [61]. The granularity of the silicon detector can be tailored as function of the required geometrical resolution.

These new photon detectors with a maximum quantum efficiency of about 30–35% around 300 nm, have made the choice of Cherenkov radiators and photon windows much more flexible. It has for instance allowed the use of aerogel in Ring Imaging Cherenkov detectors. See Sect. 7.4.2.2 and Fig. 7.12c.

Current research and development is mainly concentrated on faster and cheaper detectors with large geometrical acceptance. These are detectors like silicon avalanche photo diodes, micro channel plates and large area flat panel multi-anode PMTs. The reader is referred to Chap. 3.

An overview of the current status and perspectives of vacuum-based photon detectors can be found in [62].

7.4.5 Optics

We can broadly divide the light collection system of Ring Imaging Cherenkov detectors into two distinctive classes.

- Proximity focusing, or direct light collection as in Fig. 7.20.
- Concave mirrors as in Fig. 7.17a in Sect. 7.4.4.

7.4.5.1 Proximity Focusing

In the first case with proximity focusing optics, the resolution relies on the thinness, l, of the radiator in comparison to the expansion length, L. That is, $l \ll L$. The Cherenkov light will then describe a thin cone around the charged particle and

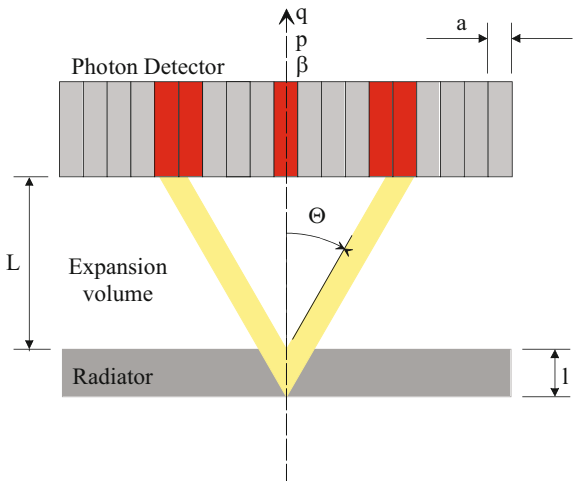


Fig. 7.20 Proximity focusing arrangement

will give rise to a finite width, conic section image where the detector plane intercepts this cone. If the particle is not perpendicular to the radiator and the detector planes, this circular image becomes distorted to an elliptic or a hyperbolic image. Depending on the refractive index of the radiator, photon window material and the expansion gap, light might be trapped due to total internal reflections. See Sect. 7.4.2.1. As the photon detector has to be placed in the path of the charged particle, the material budget may become prohibitive. However, this detector configuration is well adapted to 4π detectors with high refractive index radiators [51, 63].

7.4.5.2 Focusing Mirrors

Detectors which cover large solid angles require large focusing mirrors as in Fig. 7.21. There are two.[17] options, parabolic [65] and spherical [66, 67] mirror. The choice of mirror substrate is a balance between cost, ease of fabrication and performance. Whereas the material budget is normally not an issue in astrophysics, see Sect. 6.1 and Fig. 7.21a and b, it is one of the main concerns in accelerator based experiments as the mirrors must be inside the acceptance. If spherical aberration becomes a dominant contribution to the total error in the Cherenkov angle calculation, parabolic mirrors should be used.

[17]We will not discuss here ellipsoid nor hyperbolic mirrors. For correctors like Schmidt and Maksutov, the reader is referred to [40].

(a) (b) (c)

Fig. 7.21 (**a**) and (**b**) The VERITAS Telescope 1 as installed at the Whipple Observatory base camp. The collector dish has a diameter of 12 m and a focal length of 12 m and comprises 350 mirror facets. A 499-PMT camera is installed in the box at the focal point. Courtesy of the VERITAS Collaboration [64]. (**c**) COMPASS [67] mirror wall of RICH 1. CERN EX 0106007 01

Table 7.3 Basic material properties for some mirror substrates together with substrate rigidity, K, and the rigidity divided by material thickness in units of radiation length

Material	X_0 [cm]	E [10^4 MPa]	α [10^{-6}/°C]	Relative rigidity K	K/X_0 relative
Beryllium	35.3	28.9	11.3	1	1
Plexiglas	34.4	0.33	70	0.012	0.011
Pyrex glass	12.7	6.17	3.2	0.213	0.076
Aluminium	8.9	6.9	23.9	0.238	0.060

X_0 is the radiation length, E the Young's module and α is the coefficient of thermal expansion

The material option for the mirror substrate is a balance between radiation length, size of the substrate and stability. Some options[18,19] are given in Table 7.3.

The rigidity, K, of a thin mirror substrate is roughly given by:

$$K \propto \frac{Et^3}{D^2} \tag{7.35}$$

where E is Young's modulus, t is the substrate thickness and D is the diameter. The superiority of substrate materials like beryllium is clear in Table 7.3. In this table

[18]Plexiglas is Poly(methyl methacrylate) (PMMA) by Evoniks Business Unit Performance Polymers.

[19]Pyrex, Corning Incorporated, is made of 4% boron, 54% oxygen, 3% sodium, 1% aluminium, 38% silicon, and less than 1% potassium.

substrates of diameter 500 and 5 mm thickness are compared. However, beryllium is not a good reflector nor a good support for a reflecting surface. A thin glass face is therefore required on the beryllium as support for the reflector [68]. This glass surface can also be used to adjust the focal length of the mirror. The main challenge is to use a glass which has the same thermal expansion coefficient as beryllium.

Thin and robust mirror substrates can be made as a sandwich assembly. The kernel is normally a honeycomb or foam and the inner and outer skin are preformed to about the right radius of curvature. The final adjustment is done at the assembly stage or by reshaping, by polishing, the reflecting skin later. The skin can be high strength carbon fibre sheets [69], easily formed Plexiglas [70] or simple metal structures [71]. Glass with glass-foam kernel has also been built [72].

Glass is still the most used substrate for mirrors. It is easily shaped and machined and the ageing behaviour is well known. Stresses in the material can be simply relieved. It is also inert in most Cherenkov radiators. It is normally slumped to the required shape and then polished to the final focal length. Its principal drawback is the radiation length.

7.4.6 *The Reflective Surface*

The reflectivity of a surface is a function of the incident angle and energy of the light and the dielectric structure of the surface. The principle is discussed in [40] and more specifically in [73]. See Fig. 7.22a. A high reflectivity layer is over-coated by one or more transparent films of high and low refractive indices. Aluminium and silver are good reflectors with peak reflectivity of respectively ∼92% and ∼96%. Aluminium, the most widely used metal for reflecting films, offers consistently high reflectance throughout the visible, near-infrared, and near-ultraviolet regions of the spectrum.

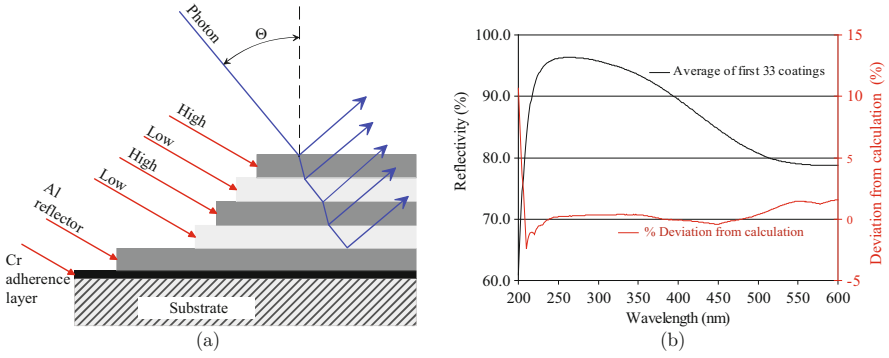


Fig. 7.22 (**a**) Schematic representation of a metal multi-dielectric mirror [73]. (**b**) Measured and calculated reflectivity of a multi-dielectric mirror coating. The stack is Cr-Al-SiO_2-HfO_2. Adapted from [73]

Table 7.4 Typical process parameters for a multi-dielectric mirror coating [73]

Material	Purity [%]	Chamber pressure [Pa]	Deposition rate [nm/s]	Thickness (geom.) [nm]
Cr	99.98	$2 \cdot 10^{-5}$	1	20
Al	99.999	$2 \cdot 10^{-5}$	5	85
SiO_2	99.99	10^{-3} O_2	0.2	28
HfO_2	99.9	10^{-3} O_2	0.2	38

While silver exhibits slightly higher reflectance than aluminium through most of the visible spectrum, the advantage is temporary because of oxidation tarnishing. Aluminium also oxidizes, though more slowly, and its oxide is tough and corrosion resistant. Oxidation significantly reduces aluminium reflectance in the ultraviolet and causes slight scattering throughout the spectrum. Generally, all reflective layers need a protective film.

Material like SiO_2 and MgF_2 have low refractive index in comparison to HfO_2 and TiO_2. Properties like residual stress, adherence, resistivity to abrasion and humidity and coating yield are essential in the selection process for these layers. The optical thickness of the layers, $d_{\text{opt}} \propto \cos\Theta$, is normally chosen to be $\lambda/4$. A dielectric coating will lead to a wavelength and angle dependent modulation of the reflectivity. The larger the ratio between the refractive indices in a Low/High pair, the higher is the peak reflectivity and width of the enhanced region. Adding more pairs for the same wavelength range, will enhance the peak reflectivity, but narrow the wavelength range. The layer stack will normally be terminated with a high refractive index layer. In this way the mirror reflectivity can be optimized for the wavelength range of the photon detector.

Mathematically approximation codes[20] will predict the behaviour of the multilayer film. The accuracy only depends on the knowledge of the refractive index and the absorption in the deposited layers. These optical properties are however dependent of the deposition method and processing parameters.

An example is shown in Fig. 7.22b. Layers of Cr, Al, SiO_2 and HfO_2 are used on a glass substrate. See Table 7.4 for process parameters. This coating is optimized for a wavelength of 275 nm in order to match a S20, footnote 16, photo cathode and compared with calculations. See footnote 20 for the calculation.

7.4.6.1 Mirror Imaging Quality

The error introduced by the imaging quality of a RICH mirror should be small compared to all other errors in the detector. If the mirror is a perfect spherical

[20] FilmStar Design, FTG Software Associates, Princeton, NJ,
SCI Film Wizard, Scientific Computing International, Carlsbad, CA
or similar.

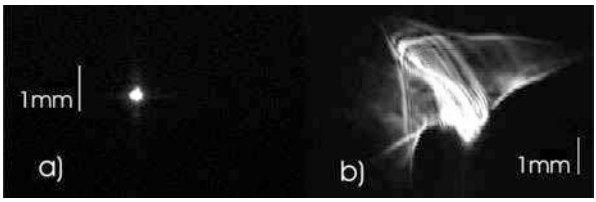


Fig. 7.23 (**a**) Spot image for a high precision glass mirror. (**b**) Spot image for a thin glass mirror [66]

surface, the spot on the focal plane would have the size given by the diffraction limit. For a circular mirror of diameter D and a radius of curvature R, the diffraction limited spot diameter, d, at the third maximum, corresponding to 95.3% of the focused light, is given by:

$$d = 2R\tan\alpha \quad \text{for} \quad \sin\alpha = \frac{\lambda x}{\pi D} \quad \text{and} \quad x = 3.7\pi \tag{7.36}$$

For a wavelength $\lambda = 641\,\text{nm}$,[21] $D = 0.50\,\text{m}$ and $R = 8\,\text{m}$, $d = 76\,\mu\text{m}$.

Real mirrors have real imperfections. Fig. 7.23 shows the difference between a high precision and a thin glass mirror. The mirror in Fig. 7.23a is a 50 mm thick glass mirror of diameter 400 mm and a radius of curvature of 7.8 m. The Fig. 7.23b mirror is 7.5 mm thick with a diameter of 400 mm and a radius of curvature of 7.8 m. 95% of the focused light for the first mirror is inside circle of diameter 0.23 mm. The corresponding diameter for the second mirror is 3.4 mm. This mirror also features irregularities at the edges of the surface. The average quality of a mirror is well described by the spot size at the focal plane and is normally sufficient as a qualification parameter. Let D_0 be the diameter of this spot which encompasses 95% of the light. $\sigma_s = D_0/4$ is the RMS of the light distribution if this distribution was Gaussian. The error induced by the mirror is then given by:

$$\sigma_\Theta = \frac{\sqrt{\sigma_s^2 + \sigma_p^2}}{2R} \approx \frac{\sigma_s}{2R} = \frac{D_0}{8R} \tag{7.37}$$

where σ_p is the resolution of the point source.

The determination of the spot shape can be an invaluable tool in the development and fabrication process. The quantification of the variation in the radius of curvature across a substrate can be used to improve the resolution of the system. It can be particularly important for large mirrors.

Shack-Hartmann sensors, Ronchi test method, Foucault method and similar measurement methods are described in detail in [74]. We will only show the power of these methods with one example.

A sketch of a Ronchi test set-up is shown in Fig. 7.24a. A beam of coherent, quasi-monochromatic light is brought to focus by an optical system that is under-

[21] Red laser diode.

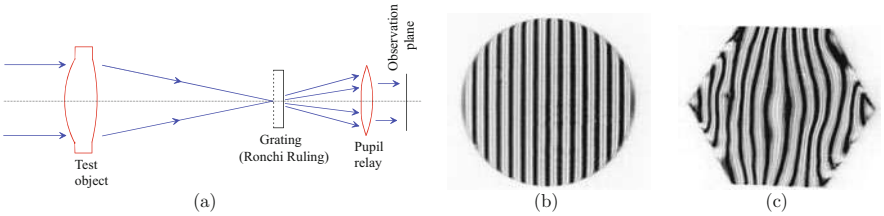
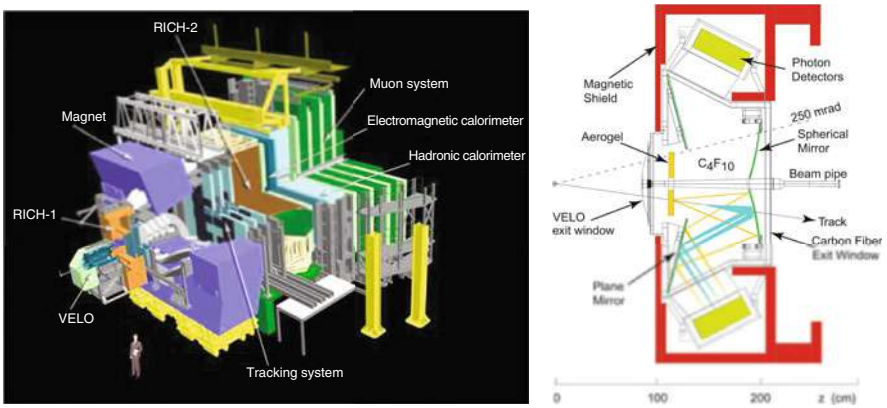


Fig. 7.24 (**a**) General set-up for a Ronchi test. (**b**) Ronchigram of a high precision spherical glass mirror. Thickness 50 mm. (**c**) Ronchigram of a thin spherical glass mirror. Thickness 4.5 mm. Ronchi ruling 1 mm

going tests to determine its aberrations. A lens, or more generally any optical system consisting of an arrangement of lenses and mirrors, is placed in the position *Test Object*. A diffraction grating, placed perpendicular to the optical axis in the vicinity of the focus, breaks up the incident beam into several diffraction orders. The diffracted orders propagate, independently of each other, and are collected by a pupil relay lens, which forms an image of the exit pupil of the object under test at the observation plane. For a concave mirror, deviation from a spherical surface will result in deformation of the fringes. The measurement is only sensitive to changes in radius of curvature perpendicular to the grating direction. Results are shown in Fig. 7.24. Figure 7.24b is a Ronchigram for a high precision spherical mirror, whereas Fig. 7.24c is for a thin large mirror. For the first mirror, the interference lines are straight which shows that the deviation from the ideal shape is smaller than the resolution of the Ronchi ruling. For the second mirror, the interference lines are distorted. In the centre, the lines bow outward and indicate parabolic deformation. On the edges, the lines bow inward to indicate an oblate spheroid surface.

7.4.7 Ring Finding and Particle Identification

As explained in Sect. 7.4, Cherenkov light is produced in a cone at polar angle Θ_C relative to the particle trajectory, as given by Eq. (7.8) for a particle travelling at velocity β. In a RICH detector the light is focussed onto a detector plane as a ring image. For the classical RICH geometry illustrated in Fig. 7.17a and [46], the detected photons corresponding to a track passing through the detector would form a circular ring image centred on the track impact point on the detector. The issues discussed in this section are the finding of the ring, i.e. the pattern recognition to associate the detected photons to a given track, and the particle identification, i.e. the determination of the particle type, given the photons that are associated to its track. Examples are taken from LHCb, Fig. 7.25, the dedicated B physics experiment at the LHC, which has two RICH detectors [75]. A review of other approaches can be found in [76].



(a) (b)

Fig. 7.25 (**a**) View of the LHCb detector. (**b**) Side view schematic layout of the RICH 1 detector. Reference [75]

For the simple detector geometry of Fig. 7.17a, and for a single track passing through the detector, the circular image implies that the photons from the track all lie at a constant radius on the detector plane, when measured from the track impact point. The radius r is related to the Cherenkov angle, Θ_C, by:

$$r = R\Theta_C/2 \tag{7.38}$$

where R is the radius of curvature of the spherical focussing mirror. For a given track the pattern recognition could therefore simply be performed by plotting the radius of all photons in this way, and searching for a peak in the distribution. Due to the finite resolution, this signal peak will have a roughly Gaussian shape, with width corresponding to the resolution. Sources of finite resolution include the pixel size of the photon detector, and the fact that the refractive index has some dependence on the photon wavelength, leading to a chromatic term in the resolution. Background hits that are distributed randomly across the detector plane, for example from noise in the photon detector, will appear as a contribution in the plot of detected photon radius that increases roughly linearly with radius (due to the increasing area swept out on the detector plane as the radius increases). This situation is illustrated in Fig. 7.26a.

Given the reconstructed radius r, the Cherenkov angle can be calculated from Eq. (7.38), and thus the velocity β of the particle determined from Eq. (7.8). To make the final step of identifying the particle, the momentum p must also be known, usually from the tracking system of the experiment that measures the curvature of the track in a magnetic field. Then the mass m of the particle can be determined using relativistic kinematics:

$$m^2 = p^2(\beta^{-2} - 1)/c^2 \tag{7.39}$$

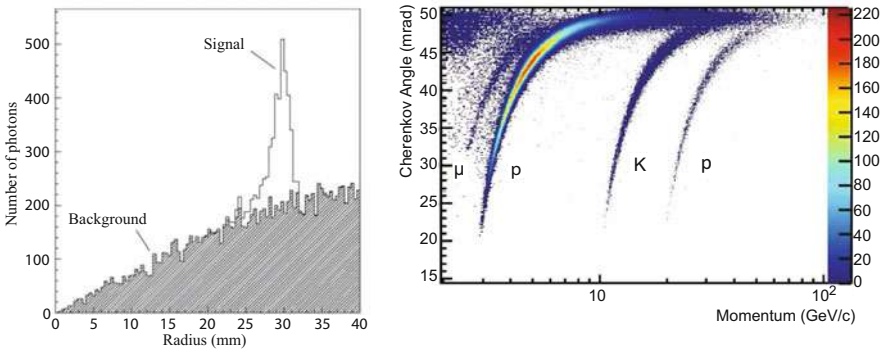


Fig. 7.26 (**a**) Distribution of photons in radius around the track, for a set of tracks in one of the LHCb RICH detectors; the peak from the photons associated to the track is visible, along with background from other sources. (**b**) Reconstructed Cherenkov angle for isolated tracks, as a function of track momentum in the C_4F_{10} radiator. The Cherenkov bands for muons, pions, kaons and protons are clearly visible. Reference [77]

and this identifies the particle type. An example is shown in Fig. 7.26b where the reconstructed Cherenkov angle has been plotted versus momentum for all the particles in a set of events, and the loci of points corresponding to particles with different masses are clearly seen.

In practical implementations of the RICH technique, the optical system usually differs from the simple classical layout, so as to avoid the material of the photon detectors being placed within the acceptance of the spectrometer. For example, the RICH detectors of the LHCb experiment involve a spherical focussing mirror that is tilted with respect to the track direction, and an additional planar mirror to bring the Cherenkov light to photon detectors sited outside the acceptance, while limiting the overall size of the detector system. This complicates the reconstruction somewhat, as the ring images are no longer circular but become distorted into roughly elliptical shapes, and the track no longer passes through the detector plane, but its image on that plane has to be calculated from knowledge of the optics. There is also an additional contribution to the resolution, due to the spherical aberration resulting from imaging the photons from off-axis tracks, but this can usually be arranged to be smaller than the limiting chromatic effect. The distortion of the ring image can be exactly corrected for by reconstructing the Cherenkov angle for each photon-track pair. For a spherical focussing mirror the analytical solution of this calculation involves the solution of a quartic equation. See [78]. For reasons of speed, a numerical approach can be used instead, ray-tracing photon candidates through the optical system and calibrating the distortion of the ring image in this fashion. The peak search is then performed in Cherenkov angle space, rather than radius on the detector plane.

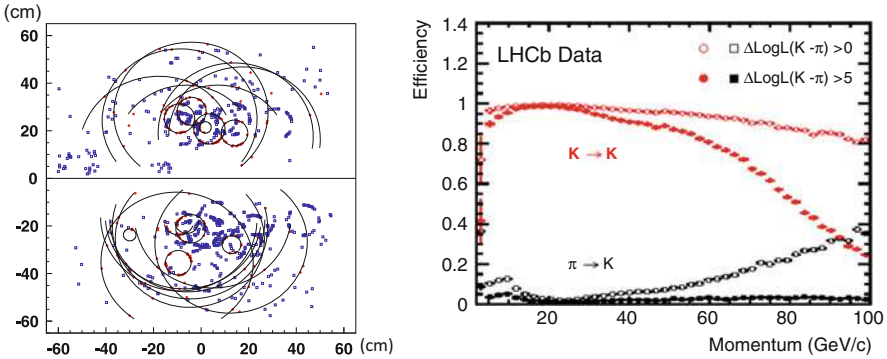


Fig. 7.27 (**a**) Ring images from tracks passing through the RICH 1 detector of LHCb, from a single proton-proton collision event at the LHC. (**b**) Kaon identification efficiency and pion misidentification rate as measured using data. Two different $\Delta\log\mathcal{L}$ (K-π) requirements have been imposed on the samples, resulting in the open and filled marker distributions, respectively. Reference [77]

This approach of peak searching works well in situations of low track multiplicity, where the ring images from tracks are well separated. However, at the LHC the track density is high, as illustrated for a typical event in Fig. 7.27a. In this case the main background to the reconstruction of the ring image of a given track comes from the overlapping rings from other tracks. It is therefore advantageous to consider the optimization of photon assignment to all of the tracks in the event simultaneously, in a so-called global approach. Since a momentum measurement is required to convert a measured ring image into particle identification, as discussed above, it makes sense to use the reconstructed tracks in the event as the starting point for pattern recognition. Trackless ring searches have been developed, but are mostly relevant for background suppression, rather than particle identification [76]. Furthermore, the number of stable charged particle types that are required to be identified is rather limited, typically five: e, μ, π, K, p. The pattern recognition can be made faster by just searching for these particle types, i.e. hypothesis testing. For applications where speed is crucial, such as use in the trigger of the experiment, the number of hypotheses compared can sometimes be further reduced, depending on the physics process that is being selected, e.g. simply comparing π and K hypotheses [79]. On the other hand, if one is interested in an unbiased search for charged particles (such as exotic states) then alternative approaches exist that do not rely on preselected hypotheses [80].

The pattern recognition then proceeds by taking the most likely hypothesis for each of the tracks in the event, typically the π hypothesis as they are the most abundantly produced particle (at the LHC). The likelihood is then calculated that the observed pattern of photons was produced by the particles, under these first choices of mass hypotheses. Conceptually this corresponds to taking the product of terms for each photon according to how close it is to the nearest ring image, assuming

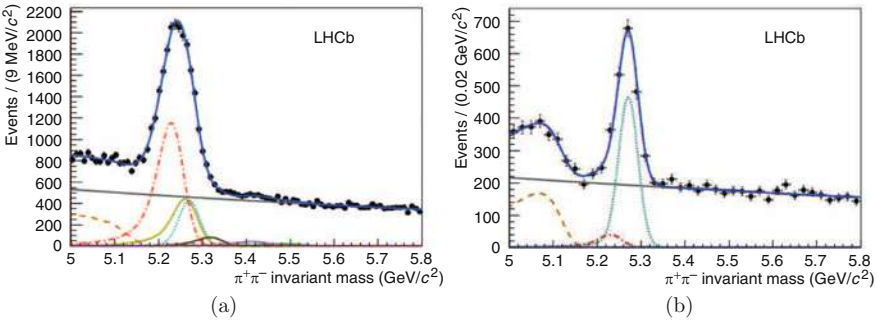


Fig. 7.28 (**a**) Invariant mass distribution for B→ h^+h^- decays in the LHCb data before the use of the RICH information, and (**b**) after applying RICH particle identification. The signal under study is the decay $B^0 \to \pi^+\pi^-$, represented by the turquoise dotted line. The contributions from different b-hadron decay modes ($B^0 \to K\pi$ red dashed-dotted line, $B^0 \to$ 3-body orange dashed-dashed line, $B_s \to$KK yellow line, $B_s \to K\pi$ brown line, $\Lambda_b \to$pK purple line, $\Lambda_b \to p\pi$ green line), are eliminated by positive identification of pions, kaons and protons and only the signal and two background contributions remain visible in the plot on the right. The gray solid line is the combinatorial background. Reference [81]

a Gaussian probability distribution around each ring. A term is also added to the likelihood from the comparison of the total number of photons assigned to a track, compared to the expected number given the mass hypothesis and momentum. The tracks in the event are then all checked to see which would give the greatest increase in the total likelihood of the event, if its hypothesis were to be changed, and the mass hypothesis of the one giving the greatest increase is then changed. This procedure is iterated until no further improvement in the likelihood can be achieved, at which point the maximum-likelihood solution to the pattern recognition has been found. By the use of various computational tricks [78] this algorithm can be reasonably fast, typically taking a similar CPU time to the track finding algorithm. The performance of this approach to particle identification when applied to LHCb events (of the type shown in Fig. 7.27a) is illustrated in Fig. 7.27b. The efficiency for identifying kaons and the misidentification rate of pions are both shown as a function of momentum.

An example from the LHCb experiment of the resulting powerful particle identification in B→ h^+h^- decays is shown in Fig. 7.28. The LHCb experiment moves to a fully software trigger where the RICH information is embedded.

7.5 Transition Radiation Detectors

A charged particle in uniform motion in free space will not radiate. It can radiate if it traverses a medium where the phase velocity of light is smaller than the velocity of the charged particle. This is Cherenkov radiation as discussed in Sect. 7.4 and was first correctly described by P.A. Cherenkov and S.I. Vavilov in 1934 and formulated

by I.M Frank and I.E. Tamm in 1937 [15]. This radiation was worked into the Bethe-Bloch formalism in 1940 by E. Fermi, see Chap. 2 and [82].

There is another type of radiation when the charged particle traverses a medium where the dielectric constant, ε, varies. This is transition radiation. It is analogous to bremsstrahlung. In both cases the radiation is related to the phase velocity of the electromagnetic waves in the medium and the velocity of the particle. In the case of transition radiation, the phase velocity changes whereas the particle velocity changes for bremsstrahlung. Transition radiation is, like bremsstrahlung, strongly forward peaked.

V.L. Ginzburg and I.M. Frank predicted in 1944 [83] the existence of transition radiation. Although recognised as a milestone in the understanding of quantum mechanics, transition radiation was more of theoretical interest before it became an integral part of particle detection and particle identification [84].

The exact calculation of transition radiation is complex and we will not repeat the mathematics here. The reader is referred to [18, 85, 86]. Specific discussions can be found in [87, 88]. We will here just recall some of the central features.

Transition radiation is emitted when a charged particle traverses a medium with discontinuous dielectric constant. Let $[\mathbf{E}_1, \mathbf{H}_1]$ be the Lorentz transformed Coulomb field of the charged particle in medium 1 and $[\mathbf{E}_2, \mathbf{H}_2]$ the corresponding one in medium 2. See Fig. 7.29a. $[\mathbf{E}_1, \mathbf{H}_1]$ and $[\mathbf{E}_2, \mathbf{H}_2]$ do not match at the boundary. In order to satisfy the continuity equation, a solution of the homogeneous Maxwell equation must be added in each medium. This is the transition radiation. The angular distribution of transition radiation by a perfectly reflecting metallic surface is of the form:

$$J(\Theta) = \omega \frac{dN}{d\omega d\Omega} = \frac{\alpha}{\pi^2} \left(\frac{\Theta}{\gamma^{-2} + \Theta^2} \right)^2 \tag{7.40}$$

where $\gamma = E/m \gg 1$ in natural units, $\hbar = c = 1$, $\alpha \simeq 1/137$ is the fine structure constant and $\Theta \ll 1$ is the angle of the photon with respect to the velocity vector $\mathbf{v}$ of the charged particle. Θ is along $\mathbf{v}$ for forward transition radiation or its mirror

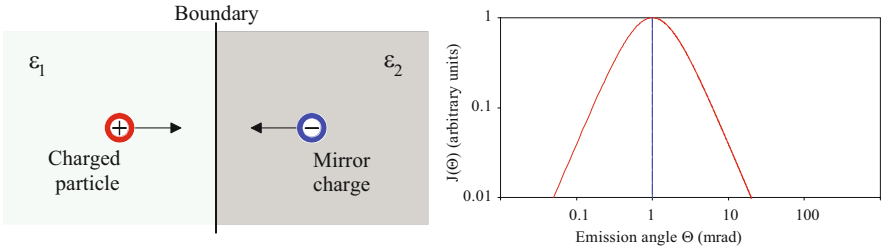


Fig. 7.29 (**a**) Schematic representation of the production of transition radiation at a boundary. (**b**) Transition radiation as function of the emission angle for $\gamma = 10^3$. Eq. (7.40)

direction for backward transition radiation. N is the total number of emitted photons. Equation (7.40) is plotted in Fig. 7.29b.

The energy radiated from a single surface, assuming $\varepsilon_0 \to \varepsilon$, is given by:

$$W = \frac{1}{3}\alpha Z^2 \omega_p \gamma \tag{7.41}$$

where ω_p is the plasma frequency.

7.5.1 Plasma Frequency

The influence of the plasma frequency was shown in the saturation of the relativistic rise expressed by the Bethe-Bloch formula, Chap. 2 and [82], due to the polarisation of the medium:

$$\frac{\delta}{2} = \ln\frac{\omega_p}{I} + \ln\beta\gamma - \frac{1}{2} \tag{7.42}$$

where I and ω_p are respectively the mean excitation energy and the plasma frequency of the medium and δ is the density correction.

The plasma frequency, ω_p, is the natural frequency of density oscillations of free electrons and its value depends only weakly on the wavelength. Longitudinal plasma waves are resonant at ω_p. Transverse electromagnetic waves are absorbed below ω_p. If $\omega < \omega_p$, the index of refraction has an imaginary part and the electromagnetic waves are attenuated or reflected. If $\omega \gg \omega_p$, the index is real and a metal becomes transparent. For large ω one can write

$$n^2 = 1 - \left(\frac{\omega_p}{\omega}\right)^2 \tag{7.43}$$

The plasma frequency is given as:

$$\omega_p^2 = \frac{NZe^2}{\varepsilon_0 m} \tag{7.44}$$

and depends only on the total number, NZ, of free electrons per unit volume. The plasma frequency can be approximated with:

$$\omega_p(\text{eV}) \simeq 28.8\sqrt{\frac{\rho(\text{g/cm}^3)\cdot z}{A}} \tag{7.45}$$

where z is the effective number of free electrons per unit volume. Table 7.5 gives the corresponding calculated and measured wavelength, λ_p, for alkali metals. $z = 1$ for alkali, group 1a, metals. The calculated plasma energies in Si, Ge and InSb are

Table 7.5 Ultraviolet transmission limits of alkali metals in nm [89]

Material	A	Z	λ_p [nm]	
			Calculated	Measured
Li	6.939	3	155	155
Na	22.99	11	209	210
K	39.10	19	287	315
Rb	85.47	37	322	340
Cs	132.95	55	362	–

Table 7.6 Radiator material properties [90]

Material		ρ [g/cm^3]	ω_p [eV]	X_0 [cm]
Lithium		0.534	13.8	148
Beryllium		1.84	26.1	34.7
Aluminium		2.70	32.8	8.91
Polyethylene	CH_2=CH_2	0.925	20.9	49
Mylar	$C_5H_4O_2$	1.38	24.4	28.7
Air		$2.2 \cdot 10^{-3}$	0.7	$30.9 \cdot 10^3$

based on four valence electrons per atom. In a dielectric the plasma oscillation is physically the same as in a metal: the entire valence electron sea oscillates back and fourth with respect to the ion core. Table 7.6 tabulates properties of some commonly used radiator material.

7.5.2 *Formation Zone*

A minimum thickness is required in order to efficiently produce the transition radiation as the evanescent field has a certain extension. This is the formation zone and is illustrated in Fig. 7.30 for a stack of aluminium, $\omega_p(\mathrm{Al}) \sim 32.8\,\mathrm{eV}$, and air, $\omega_p(\mathrm{air}) \sim 0.7\,\mathrm{eV}$. The length of the formation zone, d, can be written as:

$$d = \frac{2c}{\omega}\left[\gamma^{-2} + \Theta^2 + \left(\frac{\omega_p}{\omega}\right)^2\right]^{-1} \tag{7.46}$$

which has a maximum, $d_{\max}$, at $\omega = \gamma\omega_p/\sqrt{2}$ for $\Theta = \gamma^{-1}$, which is equivalent to the maximum intensity as can be seen from Eq. (7.40) and Fig. 7.29b.

$$d_{\max}(\mu\mathrm{m}) \sim 140 \cdot 10^{-3}\frac{\gamma}{\omega_p(\mathrm{eV})} \tag{7.47}$$

Inserting Eq. (7.45) in Eq. (7.47), we see that for media with a density in the order of 1, $\omega_p \simeq 20\,\mathrm{eV}$ and $\mathrm{d_{max}} \simeq 7\,\mu\mathrm{m}$ for $\gamma = 1000$. For a gas, ω_p is about 30 times smaller due to the reduced density and $\mathrm{d_{max}}$ thereby 30 times longer for same γ.

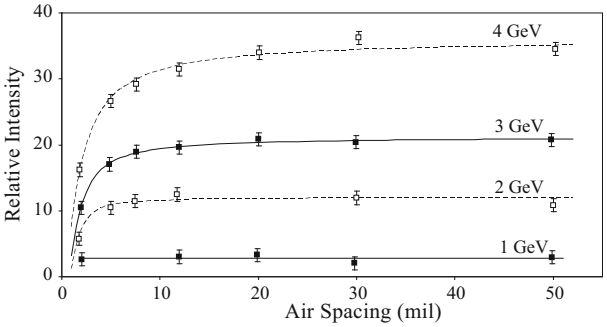


Fig. 7.30 Relative intensity of transition radiation for different air spacing. Each radiator is made of 231 aluminium foils 1 mil thick. (1 mil = 25.4 μm). Particles used are positrons of 1–4 GeV energy (γ = 2000–8000). Adapted from [91]

Using numbers for the experimental set-up in Fig. 7.30, we get $d_{max} \sim 1.5$ mm for $\gamma = 8000$.

7.5.3 Transition Radiation Detectors

From the discussion above, transition radiation can be characterized by the following:

- Transition radiation is a prompt signal.
- Transition radiation is not a threshold phenomenon.
- The total radiated power from a single interface is proportional to γ.
- The mean emission angle is inversely proportional to γ.

In general terms, there are two different types of transition radiation detectors:

1. The detectors working in the low energy, optical, range.
2. The detectors working in the X-ray range.

We will briefly introduce the first one and use a little more space on the second class of X-ray transition radiation detectors.

7.5.3.1 Optical Transition Radiation Detectors

J.E. Lilienfeld [92] was probably the first,[22] in 1919 to observe that in addition to X-rays, radiation ranging from visible light through the ultraviolet is emitted

[22] This statement has been contested over the years and could be due to a confusion between transition, Cherenkov radiation and bremsstrahlung. See [93].

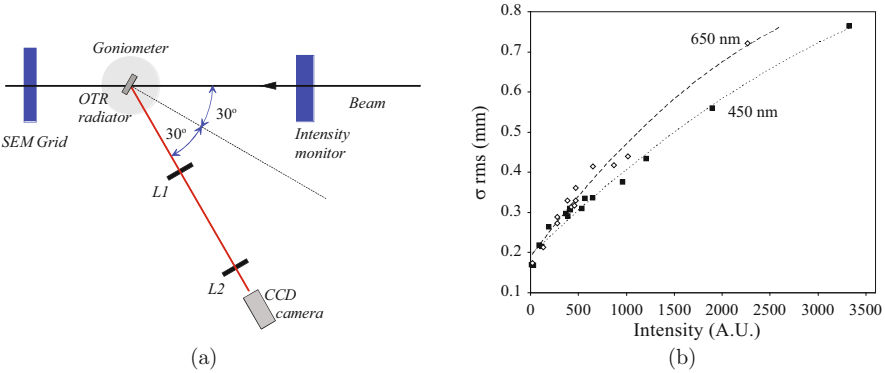


Fig. 7.31 (**a**) Sketch of an experimental set-up for measurement of optical transition radiation with secondary emission, SEM, grid and beam intensity monitor. The transition radiation foil is tilted by 30° with respect to the beam line. The optical system is defined by two lenses and a CCD camera. (**b**) Measured rms beam size values as a function of the total intensity for $\lambda = 450$ and 650 nm at 2 GeV. Adapted from [94]

when electrons approach a metallic surface. This radiation has a characteristic polarisation, spectrum and intensity. A variation to this radiation occurs when the charged particle moves roughly parallel to a conducting undulated surface. An oscillating dipole will be set up with a frequency related to the particle velocity and the undulation. The radiated power is small, but due to the microscopic source area, the brightness can be large. This has, amongst a range of other usages, found an application in accurate beam diagnostics equipments.

As an example, we will use an experiment to investigate the geometrical resolution of optical transition radiation as shown in Fig. 7.31a [94]. Integrating Eq. (7.40) over the solid angle gives:

$$\frac{dN}{d\omega} \simeq \frac{2\alpha}{\pi\omega} \ln(\gamma\Theta_{\max}) \tag{7.48}$$

where $\Theta_{\max}$ is the angle of maximum emission, measured by the optical spectrometer. The number of photons emitted is small. This must be compensated by a large number of particles in the beam.

The mathematics for such a set-up is given in [95]. The diffraction, or the Heisenberg uncertainty principle in the transverse phase-space of the photon, sets the lower limit for the size of the emitting surface:

$$\Delta b_i \geq \frac{\lambda}{2\pi} \frac{1}{2\Delta\Theta_i} \quad (i = x, y) \tag{7.49}$$

where $\lambda \sim 600$ nm is the observed wavelength. b_i and Θ_i are the components of the impact parameter **b** and the photon direction. $\Delta\Theta_i$ and Δb_i refer to rms values. Setting $\Delta\Theta = \gamma^{-1}$, or full acceptance for the photons, the resolution becomes

Table 7.7 Parameters for the fit to the data [94] and plotted in Fig. 7.31b

Parameter	$\lambda = 450\,\text{nm}$	$\lambda = 650\,\text{nm}$
ρ	$176 \pm 12\,\mu\text{m}$	$163 \pm 25\,\mu\text{m}$
a	$(9 \pm 5) \cdot 10^{-5}$	$(6 \pm 3) \cdot 10^{-5}$
b	1.12 ± 0.09	1.12 ± 0.06

proportional to γ. $\gamma = 10^5$ would give $\Delta b \geq 5\,\text{mm}$. This effect can be limited by the introduction of an iris in the optical path as in [96].

The results from [94] are shown in Fig. 7.31b. As expected, the resolution is weakly dependent on the intensity of the beam, but the total uncertainty is small. The measurement points are fitted to $\sigma_{\text{rms}} = \sqrt{\rho^2 + aI^b}$, where a and b are fit parameters, ρ is the real beam dimension and I is the beam intensity. These are given in Table 7.7.

Another promising application for optical transition radiation is in aerogel[23] Cherenkov detectors [97].

7.5.3.2 X-ray Transition Radiation Detectors

Following [98], the total radiated energy from a single surface per unit of frequency, can be approximated by:

$$\left[\frac{dW}{d\omega}\right]_{\text{s.s.}} = \frac{\alpha}{\pi}\left[\frac{1 + r + 2X_1^2}{1 - r}\ln\frac{X_1^2 + 1}{X_1^2 + r} - 2\right] \tag{7.50}$$

where

$$X_1 = \frac{\omega}{\gamma\omega_{p1}} \quad \text{and} \quad r = \frac{\omega_{p2}^2}{\omega_{p1}^2} \sim \frac{\rho_2}{\rho_1} \tag{7.51}$$

The suffix $_1$ and $_2$ denote medium 1 and 2. ω_{pi} is the plasma frequency for medium i. r will be assumed to be small and in the range of 10^{-3}, which corresponds to a $\rho = 1$ to gas interface.

By analysing Eq. (7.50), which is plotted in Fig. 7.32a, three distinct regimes can be examined:

1. If $\gamma \ll \omega/\omega_{p1}$ then $X_1 \gg 1$ and $dW/d\omega \sim \alpha/6\pi X_1^4$, which is a small number. This results in a frequency cut-off and thereby $\omega \leq \gamma\omega_{p1}$.
2. If $\omega/\omega_{p1} \ll \gamma \ll \omega/\omega_{p2}$ then $dW/d\omega \propto \ln X_1^{-1}$. That is, the total radiated power increases logarithmically with γ.
3. If $\gamma \gg \omega/\omega_{p2}$ then $X_1 \ll \sqrt{r}$. Then the total radiated power is approximately constant.

[23] See Sect. 7.4.2.2.

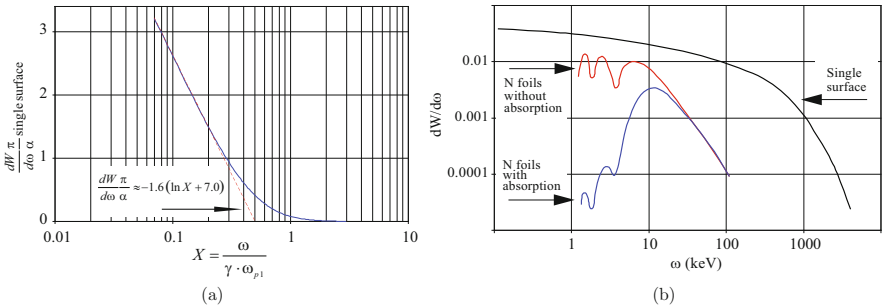


Fig. 7.32 (**a**) Total radiated energy from a single surface per unit of frequency as function of the dimensionless variable $X = \omega/\gamma\omega_{p1}$. (**b**) Intensity of the forward radiation divided by the number of interfaces for 20 μm polypropylene ($\omega_p = 21$ eV) and 180 μm helium ($\omega_p = 0.27$ eV). Adapted from [99]

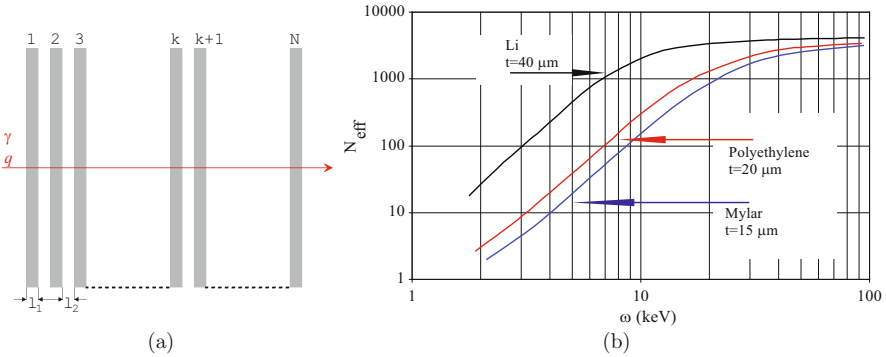


Fig. 7.33 (**a**) Sketch of a periodic transition radiation radiator. (**b**) The effective number of foils in a radiator as function of photon energy. Adapted from [90]

It can be shown that the mean radiated energy in this single surface configuration can be written as:

$$W \simeq 2\alpha\gamma\omega_{p1}/3 \tag{7.52}$$

and that the number of high energy photons produced are of the order of α when taking into account the frequency cut-off discussed above:

$$N_{\text{photons}}(\omega > 0.15\gamma\omega_{p1}) \simeq \alpha/2 \tag{7.53}$$

A large number of interfaces are therefore required to have an effective detector with a sufficient signal-to-noise ratio. A periodic transition radiation radiator is sketched in Fig. 7.33a. It should be noted that the radiators do not need to be rigorously periodic, but it is helpful for the calculation of the yield.

The basic mathematics can be found in [85, 90, 98]. Computational models can be found in [100]. The effective final number of transition radiation high energy photons at the end of the radiator stack is a function of constructive and destructive effects. See Fig. 7.32b. We will list the main effects here:

- The total radiated energy of a single surface is proportional to the plasma frequency and thereby proportional to $\sqrt{Z}$ of the material. Equation (7.44). The absorption of these photons is governed by photo-electric effects and the absorption coefficients in the stack. This goes approximately like Z^5. The radiator material should therefore be of low Z.
- The thickness of the radiator material, $\mathtt{l}_1$ in Fig. 7.33a, must be large enough to contain the formation zone for the required γ, but short enough not to introduce multiple scattering effects and bremsstrahlung. The gas density will always introduce a negative effect and should be kept as low as possible.

For a practical transition radiation radiator and following [90], the expression of the total flux, is then represented by an integration over the emission angle Θ and a function which represents the incoherent addition of the single foil intensities and includes the photon absorption in the radiator. The effective number of foils in the radiator can then be expressed as:

$$N_{\text{eff}} \simeq \frac{1 - \exp(-N\sigma)}{1 - \exp(-\sigma)} \tag{7.54}$$

where $\sigma = (\kappa\rho t)_{\text{foil}} + (\kappa\rho t)_{\text{gas}}$ and κ, ρ and t are respectively the absorption coefficient, density and thickness of the material. The self-absorption of the photons from transition radiation limits the yield and $N_{\text{eff}} \to 1/\left[1 - \exp(-\sigma)\right]$ for $N \to \infty$. A typical mean energy for the photons in a practical radiator is in the range of 10 keV. See Fig. 7.32b. The spectrum will be softer for foils with lower plasma frequencies. Since N_{eff} in Eq. (7.54) is depending on the absorption coefficient through the frequency of the photon, N_{eff} will saturate for high frequencies as shown in Fig. 7.33b.

7.5.3.3 X-ray Detectors

Any detector which has a sufficiently high efficiency for X-rays of the order of 10 keV can be used. In the design of the detector it should be noted that the number of transition radiation photons is small and produced very close to the path of the charged particle which will normally also traverse the detector. The traditional detector is a MWPC-like detector, Chap. 3, which directly follows the radiator. In order to enhance the signal-to-noise ratio and efficiently use the space as the effective number of interfaces in the radiator will saturate, a transition radiation detector is therefore normally many radiator/detector assemblies.

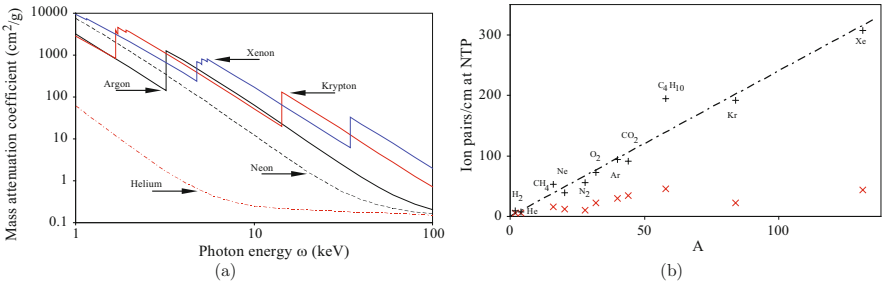


Fig. 7.34 (**a**) X-ray mass attenuation coefficient, μ/ρ, as function of the photon energy. $\mu/\rho = \sigma_{tot}/uA$, where $u = 1.660 \times 10^{-24}$ g is the atomic mass unit, A is the relative atomic mass of the target element and σ_{tot} is the total cross section for an interaction by the photon. Data from http://physics.nist.gov/PhysRefData/. (**b**) The (×) primary and (+) total number of ion pairs created for a minimum ionizing particle per cm gas at normal temperature and pressure as function of molecular mass A [101]

The ionization loss, dE/dx, from the charged particle will create charge clusters. Some of them rather far from the track due to δ-electrons. The absorption of transition radiation photons will produce a few local strong charge clusters. The choice of gas is therefore a compromise between photon absorption length, Fig. 7.34a, and the background from dE/dx, Fig. 7.34b. The optimal gas thickness is about one absorption length for 10 keV. Xenon is the preferred gas with a chamber thickness of 10–15 mm. See discussion in [90]. CO_2, or similar, is added as quencher.

A minimum ionizing particle, MIP, will produce a total of ∼310 ion pairs per cm xenon gas. Figure 7.34b. The relativistic rise is about 75% in xenon at 1 atm, or about 550 ion pairs/cm will be produced by a high γ charged particle. The average energy required to create an ion pair in a gas, is typically 25–35 eV. For xenon it is measured to 22.1 ± 0.1 eV [102], or about double the ionization energy for the least tightly bound shell electron. A 10 keV transition radiation photon will then produce about 450 ion pairs. The signal-to-noise ratio will be further reduced due to Landau-fluctuations and gain variations in the detector and electronics. Additional background might arise from curling in a magnetic field, bremsstrahlung and particle conversions. The challenge is then to correctly identify the photon cluster from a dE/dx signal of about the same strength. We will illustrate this by looking more closely at the choices made by the ALICE [103] and ATLAS [104] experiments.

7.5.3.4 ATLAS Transition Radiation Tracker

In the ATLAS experiment, the transition radiation tracker (TRT) in the barrel comprises many layers of gaseous straw tube elements interleaved with transition radiation material. Figure 7.35. With an average of 36 hits per track, it provides continuous tracking to enhance the pattern recognition and improve the momentum

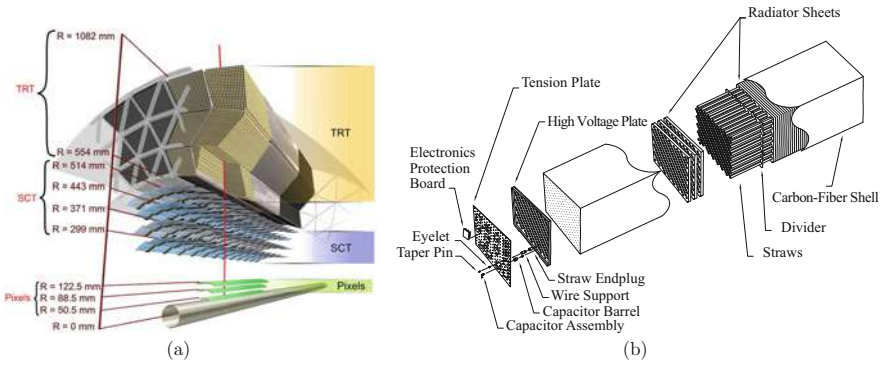


Fig. 7.35 (**a**) ATLAS Detector. Drawing showing the sensors and structural elements traversed by a charged track of 10 GeV p_t in the barrel inner detector (pseudo rapidity $\eta = 0.3$). The track traverses approximately 36 axial straws of 4 mm diameter contained in the barrel transition-radiation tracker modules. [104]. (**b**) Layout of an ATLAS Barrel TRT module. The ATLAS TRT collaboration et al. [105] with permission

resolution over $|\ \eta\ | < 2.0$[24] and electron identification complementary to that of the calorimeter over a wide range of energies. A similar detector is placed in the forward direction.

The transition radiator material which completely surrounds the straws inside each module, Fig. 7.35b, consists of polypropylene-polyethylene fibre mat about 3 mm thick. The fibres are typically 19 μm in diameter and are formed from polyethylene clad polypropylene material. The fibres are formed into fabric plies with 3 mm thickness and a density of about 0.06 g/cm^3. The absorption length for the lowest energy photons of interest (5 keV) is about 17 mm in the radiator material.

The ATLAS TRT uses two thresholds to discriminate between digitisations from tracks and those from transition radiation:

1. Low threshold, LT, for tracking which is set to ∼300 eV with 8 digitisations over 25 ns.
2. High threshold, HT, set in the range 5–7 keV with 1 digitisation over 25 ns and read out in 75 ns segments.

As the $\beta\gamma$ of the traversing particles will vary greatly, and thereby the ionization in the straw tubes, a Time-over-Threshold parameter can be defined from the LT digitisations in order to enhance the signal-to-noise estimate for the transition radiation signal.

Particle identification properties of the TRT Barrel using transition radiation were studied at several different beam energies. The good agreement between 2 GeV low

[24] Pseudo rapidity, η, is describing the angle of a particle relative to the beam axis. $\eta = -\ln\left[\tan\left(\frac{\Theta}{2}\right)\right] = \frac{1}{2}\ln\left[\frac{|\mathbf{p}|+p_L}{|\mathbf{p}|-p_L}\right]$. Θ is the angle between the particle momentum and the beam axis, $\mathbf{p}$ is the momentum vector and p_L is the longitudinal momentum component.

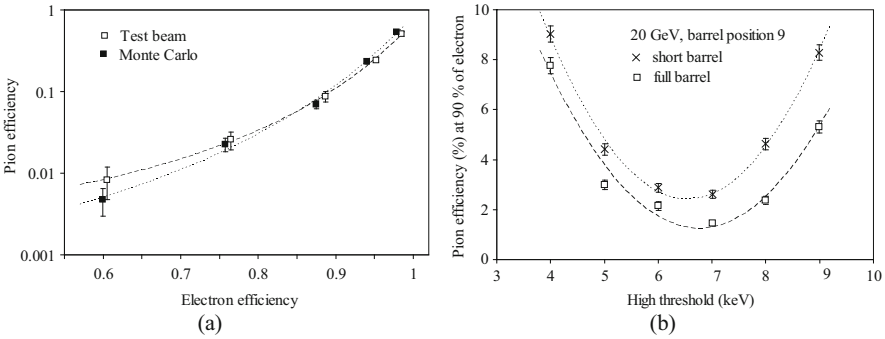


Fig. 7.36 (**a**) ATLAS TRT test beam. Pion rejection curve for a 2 GeV e/π beam. Cornelissen and Liebig [106] with permission. (**b**) ATLAS TRT test beam. e/π rejection power as a function of the high level threshold. Full barrel: all barrel straw layers are active. Short barrel: particle crosses the barrel in the central area where the first 9 layers do not have active anode wires. The ATLAS TRT collaboration et al. [105] with permission

energy data and simulation is shown in Fig. 7.36a. The results for 20 GeV beam energy are shown in Fig. 7.36b. On this figure the pion rejection power is shown as a function of the high level threshold at two beam positions along the straw. The upper points are when beam particles crossed the Barrel module 40 cm from its edge. At this position the first 9 straw layers are not active. The lower points are when the beam is positioned 20 cm from the edge of the Barrel where all 73 straw layers are active. As seen in this figure the best particle identification properties for the TRT Barrel are at a threshold of about 7 keV. Pion mis-identification in that case is 1.5–3% at 90% of the electron efficiency.

7.5.3.5 ALICE Transition Radiation Detector

The main purpose of the ALICE Transition Radiation Detector (TRD) [103, 107] is to provide electron identification in the central barrel for momenta above 1 GeV/c. Below this momentum electrons can be identified via specific energy loss measurement in the TPC. Above 1 GeV/c transition radiation from electrons passing a radiator can be exploited together with the specific energy loss in a suitable gas mixture to obtain the necessary pion rejection capability. The chamber geometry and the read-out electronics were chosen to reconstruct track segments. Since the angle of the track segment with respect to the origin is a measure of the transverse momentum of the electron, this information is used in the first level trigger within 5 μs of the collision.

The pion rejection is governed by the signal-to-background ratio in the measurement of J/Ψ production and its p_t dependence. This led to the design goal for the pion rejection capability of a factor 100 for momenta above 1 GeV/c in central Pb-Pb collisions.

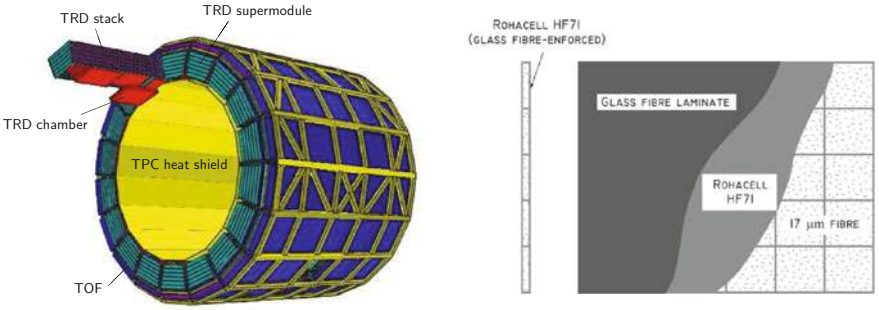


Fig. 7.37 (**a**) Schematic drawing of the TRD layout in the ALICE space frame. Shown are 18 super modules each containing 30 readout chambers (red) arranged in five stacks of six layers. One chamber has been displaced for clarity. On the outside the TRD is surrounded by the Time-Of-Flight (TOF) system (dark blue). On the inside the heat shield (yellow) towards the TPC is shown. The ALICE Collaboration et al. [103] with permission. (**b**) The principle design of the TRD sandwich radiator. The ALICE Collaboration et al. [107] with permission

The TRD consists of 540 individual readout detector modules. Figure 7.37a. Each detector element consists of a carbon fibre laminated Rohacell[25]/polypropylene fibre sandwich radiator, Fig. 7.37b, of 48 mm thickness, a drift section of 30 mm thickness, or about 2 μs, and a multi-wire proportional chamber section (7 mm) with pad readout.

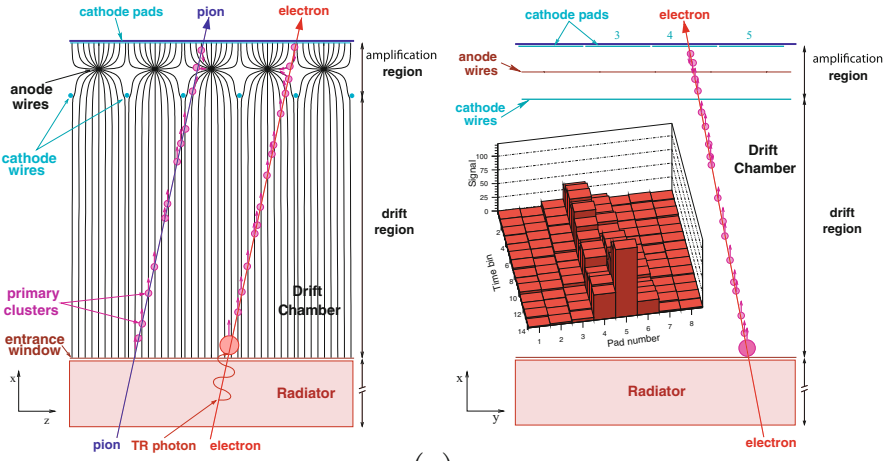
Following [108], employing the drift time information in a bidimensional likelihood [109], the pion rejection capability can be improved by about 60% [110] compared to the standard likelihood method on total deposited charge. This method is the simplest way of extending the standard method. However, it does not exploit all recorded information, namely the amplitude of the signal in each time bin. Along a single particle track this information is highly correlated, Fig. 7.38a, due to

- the intrinsic detector signal, in particular since a Xe-based mixture is used
- the response of the front-end electronics used to amplify the signals.

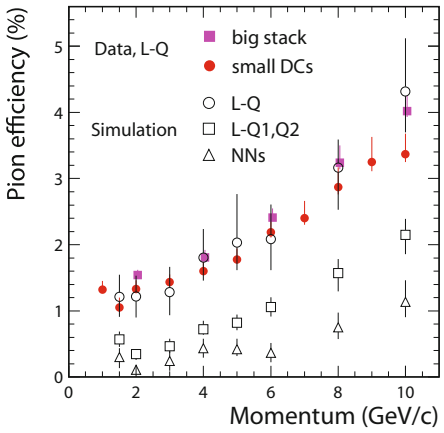
Under these circumstances, the usage of a neural network (NN) algorithm is a natural choice for the analysis of the data. The result of the data analysis from a 2–6 GeV/c mixed e/π test beam is shown in Fig. 7.38b [108]. Neural Network algorithm might improve the pion rejection significantly by a factor larger than 3 for a momentum of 2 GeV/c compared to other methods.

The detector was completed in the LS 1 before RUN 2 at LHC. Since then it provides coverage of the full azimuthal acceptance of the central barrel. Figure 7.39 shows the p_T spectra of electron candidates with 6 layers identified using the TPC and the TOF in the minimum-bias and triggered data sample. The expected onset

[25] ROHACELL is a close cell polymethacrylimide- (PMI-) rigid foam by Evonik Industries AG, Germany.



(a)



(b)

Fig. 7.38 (**a**) Schematic cross-sectional view of an ALICE detector module in rz and $r\phi$-direction. The inset shows the charge deposit from an inclined track which is used for momentum reconstruction. The ALICE Collaboration et al. [103] with permission. (**b**) Measured pion efficiency as a function of beam momentum applying likelihood on total deposited charge (L-Q) (full symbols) measured with a stack of six chambers and smaller test chambers. Results are compared to simulations (open symbols) for 90% electron efficiency and six layers. These simulations were extended to two-dimensional likelihood on deposited charge and position (LQ1, Q2) and neural networks (NN). The ALICE Collaboration et al. [103] with permission

at the trigger threshold of 3 GeV/c is observed for the triggered events and shows in comparison to the corresponding spectrum from minimum-bias collisions an enhancement of about 700. At 90% electron efficiency, a pion rejection factor of

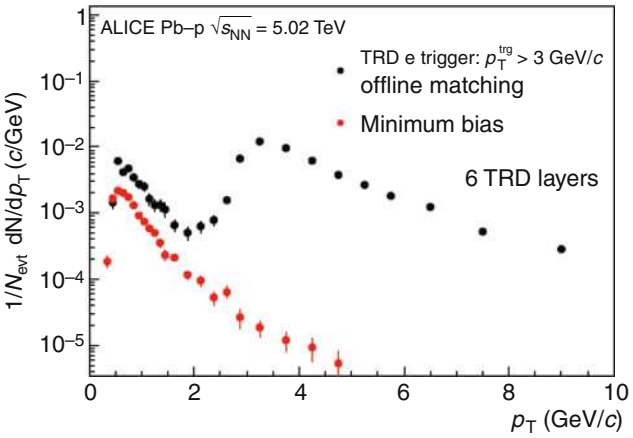


Fig. 7.39 p_T spectra of identified electrons for the minimum-bias and TRD-triggered data sample of Pb-p collisions at $\sqrt{s_{NN}} = 5.02$ TeV. For the result of the TRD-triggered sample, electrons from photon conversions in the detector material were rejected by matching the online track with a track in the TPC. Reference [111]

about 70 is achieved at a momentum of 1 GeV/c for simple identification algorithms. When using the temporal evolution of the signal, a pion rejection factor of up to 410 is obtained.

References

1. B. Dolgoshein, Complementary particle ID: transition radiation and dE/dx relativistic rise, Nucl. Instrum. Methods Phys. Res., A: 433 (1999) 533
2. C. Ward et al., A Large Aperture Time-Of-Flight Counter System, Nucl. Instr. and Meth. 30(1964)61–68
3. G. Charpak, L. Dick and L. Feuvrais, Location of the position of a particle trajectory in a scintillator, Nucl. Instr. and Meth. 15(1962)323–326
4. W. Klempt, Review of particle identification by time of flight techniques, Nucl. Instrum. Methods Phys. Res., A: 433 (1999) 542
5. A. Rabl and R. Winston, Ideal concentrators for finite sources and restricted exit angles, Applied Optics, Vol. 15 Issue 11, pp.2880–2883 (1976)
   R.H. Hildebrand and R. Winston, Throughput of diffraction-limited field optics systems for infrared and millimetric telescopes: erratum, Applied Optics, Vol. 21 Issue 17, pp.3065–3065 (1982)
6. W. Blaschke and G. Thomsen, Vorlesungen Uber Differential-Geometrie und Geometrische Grundlagen von Einsteins Relativitatstheorie. I: Elementaire Differentialgeometrie, Dritte Erweiterte Auflage. Bearbeitet under herausgegeben, DOVER PUBLICATIONS (1945)
7. T. Kobayashi and T. Sugitate, Test of Prototypes for a Highly Segmented TOF Hodoscope, Nucl. Instrum. Methods Phys. Res., A: 287 (1990) 389–396
8. C. Field et al., Development of photon detectors for a fast focusing DIRC, Nucl. Instrum. Methods Phys. Res., A 553 (2005) 96–106

9. J. Blazej et al., Gallium-based avalanche photodiode optical crosstalk, Nucl. Instrum. Methods Phys. Res., A: 567 (2006) 239–241, and references therein.
10. J. Va'vra et al., A 30 ps timing resolution for single photons with multi-pixel Burle MCP-PMT, Nucl. Instrum. Methods Phys. Res., A: 572 (2007) 459–462
11. A.N. Akindinov et al., Latest results on the performance of the multigap resistive plate chamber used in the ALICE TOF, Nucl. Instrum. Methods Phys. Res., A: 533 (2004) 74–78
12. J. Va'vra, PID techniques: Alternatives to RICH methodes, Nuclear Instruments & Methods in Physics Research A (2017), http://dx.doi.org/10.1016/j.nima.2017.02.075
13. A. Bravar, Recent Results from NA61 (Flux Related Systematic Uncertainties) and Recent Results from T2K (Overall Systematic Uncertainties), http://www.t2k.org/docs/talk/225, 2015-11-06
N. Abgrall et al., Measurements of cross sections and charged pion spectra in proton-carbon interactions at 31 GeV/c, PHYSICAL REVIEW C 84, 034604 (2011)
14. S. Afanasiev et al., The NA49 Large Acceptance Hadron Detector, CERN-EP/99-001, 6. January 1999
15. P.A. Cherenkov, Visible glow under exposure of gamma irradiation, Dokl. Acad. Nauk. USSR 2(1934)451
S.I. Vavilov, About possible reasons of blue gamma-glow of liquids, Dokl. Acad. Nauk. USSR 2(1934)457
I.M. Frank and I.E. Tamm, Coherent Radiation of Fast Electrons in a Medium, Dokl. Acad. Nauk. USSR 14(1937)107
16. J.V. Jelley, Čerenkov Radiation and its Application, London: Pergamon Press, 1958
V.P. Zrelov, Čerenkov Radiation in High Energy Physics, Jerusalem, Israel Program for Scientific Translations, 1970
17. J. Séguinot, Les Compteurs Cherenkov: Applications et Limites pour l'Identification des Particules. Développements et Perspectives, CERN-EP/89–92, LPC/89-25, 31 juillet 1989
18. J.D. Jackson, Classical Electrodynamics, John Wiley & Sons, Inc.,1998,
ISBN 0-471-30932-X
19. W.W.M. Allison and P.R.S. Wright, The Physics of Charged Particle Identification, in:
T. Ferbel (ed.), Experimental Techniques in High-Energy Nuclear and Particle Physics, ISBN 981-02-0867-7, World Scientific Publishing Company, Inc. 1999.
20. Landolt-Börnstein, Eigenschaften der Materie in ihren Aggregatzuständen, 8. Teil Opische Konstanten.
21. C.C. Chen and W.A. Steel, Molecular calculation of the dielectric constant of N_2 and CO_2, J. Chem. Phys., 75(1)383–387
22. DuPont, Freon Technical Bulletin B-32, 32A
23. P.W. Langhoff and M. Karplus, Pade Summation of the Cauchy Equation, Journal of the Optical Society of America, 59(1969)863
24. J. Koch, Nova acta Soc. Upsal. (4) Nr. 5, 1909
C. u. M. Cuthbertson, Proc. Roy. Soc. London (A) 83, 151, 1910
25. E. Albrecht et al., VUV absorbing vapours in n-perfluorocarbons, Nucl. Instr. and Meth. in Phys. Res. A 510(2003)262–272
26. D.E. Gray (ed.), American Institute of Physics handbook, McGraw-Hill Book Company. Inc., 1957
27. E. Efimov and S. Stone, A novel LiF radiator for RICH detectors, Nucl. Instr. and Meth. in Phys. Res. A 371(1996)79–81
28. B. Dey et al., Design and performance of the focusing DIRC detector, Nuclear Instruments and Methods in Physics Research A 775(2015)112–131
J. Schwiening, The DIRC detector at the SLAC B-factory PEP-II: operational experience and performance for physics application, Nucl. Instr. and Meth. in Phys. Res. A 502(2003)67–75
P. Coyle et al., The DIRC counter: a new type of particle identification device for B factories, Nucl. Instr. and Meth. in Phys. Res. A 343(1994)292–299

29. Blair N. Ratcliff, Imaging rings in Ring Imaging Cherenkov counters, Nucl. Instr. and Meth. in Phys. Res. A 502(2003)211–221
30. J. Fast, The Belle II imaging Time-of-Propagation (iTOP) detector, Nuclear Instruments & Methods in Physics Research A (2017), http://dx.doi.org/10.1016/j.nima.2017.02.045
Y. Enari et al., Progress report on Time-Of-Propagation counter—a new type of ring imaging Cherenkov detector, Nucl. Instr. and Meth. in Phys. Res. A 494(2002)430–435
31. G. S. Varner, Waveform-sampling electronics for Cherenkov detectors, https://indico.cern.ch/event/393078/contributions/2241768/attachments/1334816/2007370/1-WaveformReadout_Varner_RICH2016.pdf
32. M.J. Charles and R. Forty, TORCH: Time of flight identification with Cherenkov radiation, Nuclear Instruments and Methods in Physics Research A 639(2011)173?176
T. Gys et al., The TORCH detector R & D: Status and perspectives, Nuclear Instruments and Methods in Physics Research A (2017), http://dx.doi.org/10.1016/j.nima.2017.02.060
33. C. Schwarz et al., The PANDA DIRC Detectors at FAIR, arXiv:1707.09269 [physics.ins-det]
34. M. Cantin et al., Silica aerogels used as Cherenkov radiators, Nucl. Instr. and Meth. 118(1974)177–182
35. H. Gordon et al., The Axial Field Spectrometer at the CERN ISR, CERN-EP/81–34
36. L. A. Paquette (ed), D. Crich (ed), Ph. L. Fuchs (ed) and G. Molander (ed), Encyclopedia of Reagents for Organic Synthesis, Wiley; 2 edition (May 18, 2009)
L. Rösch, P. John and R. Reitmeier, Silicon Compounds, Organic, Ullmann's Encyclopedia of Industrial Chemistry, Wiley-VCH; Sixth edition (February 21, 2003)
37. T. Bellunato et al., Refractive index of aerogel: uniformity and dispersion law, Nucl. Instr. and Meth. in Phys. Res. A 595(2008)183–186
T. Bellunato et al., Refractive index dispersion law of silica aerogel, Eur. Phys. J. C52(2007)759–764
38. D. Perego. Private communication, March 2008
39. A.Yu. Barnyakov et al., Focusing Aerogel RICH optimization, Nucl. Instr. and Meth. in Phys. Res. A 595(2008)100–103
40. M.Born and E. Wolf, Principles of Optics, University Press, Oxford, ISBN 0-521-64222-1
41. R. Pestotnik et al., The aerogel Ring Imaging Cherenkov system at the Belle II spectrometer, Nuclear Instruments & Methods in Physics Research A (2017), http://dx.doi.org/10.1016/j.nima.2017.04.043
42. O. Chamberlain, E. Segre, C. Wiegand and Th. Ypsilantis, Observation of antiprotons, Phys. Rev. 100, No. 3, 947 (1955).
43. J. Litt and R. Meunier, Čerenkov counter technique in high-energy physics, Annu. Rev. Nucl. Part. Sci.: 23(1973)1–43
R. Meunier, Čerenkov Detectors, 1973 International Conference on Instrumentation for High Energy Physics, Frascati, Italy, May 8–12, 1973, Published by Laboratori Nazionali del CNEN
44. M. Bourquin et al., Particle and antiparticle production by 200 GeV/c protons in the charged hyperon beam at the CERN SPS, Nucl. Phys. B 153(1979)13–38
45. J. Séguinot and T. Ypsilantis, Photoionization and Cherenkov Ring Imaging, Nucl. Instr. and Meth., 142(1977)377; for an extensive list of RICH references, see http://alice-hmpid.web.cern.ch/alice-hmpid/basic_references.htm
46. J. Séguinot and T. Ypsilantis, A historical survey of ring imaging Cherenkov counters, Nucl. Instr. and Meth. in Phys. Res. A, 343(1994)1–29
J. Séguinot and T. Ypsilantis, Theory of ring imaging Cherenkov counters, Nucl. Instr. and Meth. in Phys. Res. A, 343(1994)30–51
47. Carl A. Heller and Aaron N. Fletcher, Oxidation and Chemiluminescence of Tetrakis (dimethylamino) ethylene. I. Reversible Reactions of Oxygen with Tetrakis (dimethylamino) ethylene and n-Decane, J. Phys. Chem.; 1965; 69(10); 3313–3317
J.P. Paris, Chemiluminescence of tetrakis-(dimethylamino)-ethylene, Photochemistry and Photobiology, 4(6)(1965)1059–1065
Sidney Toby, Paul A. Astheimer and Frina S. Toby, Chemiluminescence in the gas phase

reaction between tetrakis-(dimethylamino)-ethylene and oxygen, Journal of Photochemistry and Photobiology A: Chemistry 67(1992)1–12
W.R. Carpenter and E. M. Bens, Influence of oxidation products on the chemiluminescent oxidation of tetrakis-(dimethyl amino)-ethylene. Naval ordnance test station China Lake CA, Report 0526608, FEB 1967
W.P. Norris, Reactions of Tetrakis-(Dimethylamino)-Ethylene with Weak Acids. Naval ordnance test station China Lake CA, Report 0111409, OCT 1972
48. D. F. Anderson, A xenon gas scintillation proportional counter coupled to a photoionization detector, Nucl. Instr. and Meth., 178(1980)125–130
Wiliam H. -M. Ku, Charles J. Hailey and Michael H. Vartanian, Properties of an imaging gas scintillation proportional counter, Nucl. Instr. and Meth., 196(1982)63–67
D. F. Anderson, Measurement of TMAE and TEA vapor pressures, Nucl. Instr. and Meth. in Phys. Res. A, 270(1988)416–418
Richard A. Holroyd et al., Measurement of the absorption length and absolute quantum efficiency of TMAE and TEA from threshold to 120 nm, Nucl. Instr. and Meth. in Phys. Res. A, 261(1987)440–444
49. H.W. Siebert et al., The Omega RICH, Nucl. Instr. and Meth. in Phys. Res. A, 343(1994)60–67
50. D. Aston et al., Development and construction of the SLD Cherenkov ring-imaging detector, Nucl. Instr. and Meth. in Phys. Res. A, 283(1989)582–589
51. E. Albrecht et al., Operation, optimisation, and performance of the DELPHI RICH detectors, Nucl. Instr. and Meth. in Phys. Res. A, 433(1999)47–58
52. M. Starič, A. Stanovnik and P. Križan, Tests of a solid CsI photocathode in a thin multiwire proportional chamber, Nucl. Instr. and Meth. in Phys. Res. A, 307(1991)145–148
J. Almeida et al., Development of large area fast-RICH prototypes with pad readout and solid photocathodes, Nucl. Instr. and Meth. in Phys. Res. A. 348(1994)216–222
53. J.R. Hardy and A.M. Karo, The Lattice Dynamics and statics of alkali halide crystals, Plenum Press 1979, ISBN 0-306-4022-1
I.P. Csorba, Imaging Tubes, Howard W, Sams & Co., Inc. 1985, ISBN 0-672-22023-7
54. C. Lu and K. T. McDonald, Properties of reflective and semitransparent CsI photocathodes, Nucl. Instr. and Meth. in Phys. Res. A. 343(1994)135–151
A. Braem et al., Technology of photocathode production, Nucl. Instr. and Meth. in Phys. Res. A, 502(2003)205–210
55. T. Meinschad, L. Ropelewski and F. Sauli, GEM-based photon detector for RICH applications, Nucl. Instr. and Meth. in Phys. Res. A. 535(2004)324–329
A. Breskin et al., Recent advances in gaseous imaging photomultipliers, Nucl. Instr. and Meth. in Phys. Res. A. 513(2003)250–255
J.M. Maia et al., Single-UV-photon 2-D imaging with multi-GEM detectors, Nucl. Instr. and Meth. in Phys. Res. A. 580(2007)373–376
L. Periale et al., Detection of the primary scintillation light from dense Ar, Kr and Xe with novel photosensitive gaseous detectors, Nucl. Instr. and Meth. in Phys. Res. A. 478(2002)377–383, and references therein.
56. F. Tessarotto, Status and perspectives of gaseous photon detectors, Nuclear Instruments & Methods in Physics Research A (2017), http://dx.doi.org/10.1016/j.nima.2017.023.011
57. I. Ariño et al., The HERA-B RICH, Nucl. Instr. and Meth. A 453 (2000) 289–295
58. P. Abbon et al., Read-out electronics for fast photon detection with COMPASS RICH-1, Nucl. Instr. and Meth. A 587 (2008) 371–387
59. M. Moritz et al., Performance Study of New Pixel Hybrid Photon Detector Prototypes for the LHCb RICH Counters, IEEE TRANSACTIONS ON NUCLEAR SCIENCE, 51(2004)1060–1066
60. LHCb Collaboration, LHCb RICH, Technical Design Report, CERN/LHCC/2000-0037, 7 September 2000
61. A. Braem et al., Development, fabrication and test of a highly segmented hybrid photodiode, Nucl. Instr. and Meth. in Phys. Res. A. 478(2002)400–403

62. T. Gys, Micro-channel plates and vacuum detectors, Nuclear Instruments & Methods in Physics Research A 787(2015)254–260
T. Iijima, Status and perspectives of vacuum-based photon detectors, Nuclear Instruments & Methods in Physics Research A 639(2011)137–143
T. Gys, Status and perspectives of vacuum-based photon detectors for single photon detection, Nuclear Instruments & Methods in Physics Research A 595(2008)136–141
63. ALICE Collaboration, ALICE Technical Design Report, Detector for High Momentum PID, CERN/LHCC 98–19, ALICE TDR 1, 14 August 1998.
64. J. Holder et al., The first VERITAS telescope, Astroparticle Physics 25 (2006) 391–401
65. P. Baillon et al., An improved method for manufacturing accurate and cheap glass parabolic mirrors, Nucl. Instrum. Methods Phys. Res., A 276(1989)492–495
P. Majumdar et al., Angular Resolution of the Pachmarhi Array of Cerenkov, Telescopes Astropart. Phys. 18(2003)333–349
66. M. Laub, Development of opto-mechanical tools and procedures for the new generation of RICH-detectors at CERN, CERN-THESIS-2006-028; LHCb-2001-130; CERN-LHCb-2001-130, Prague: Prague TU, 2001
67. E. Albrecht et al., The mirror system of COMPASS RICH-1, Nucl. Instrum. Methods Phys. Res., A 502 (2003) 236–240
68. G.J. Barber et al., Glass-coated beryllium mirrors for the LHCb RICH1 detector, Nucl. Instrum. Methods Phys. Res., A 570(2007)565–57
69. J. Friese, R. Gernhäuser, P. Maier-Komor and S. Winkler, A new carbon based VUV mirror of high radiation length for the HADES RICH, Nucl. Instrum. Methods Phys. Res., A 502(2003)241–245
F.C.D. Metlica and On behalf of the LHCb Collaboration, Development of light-weight spherical mirrors for RICH detectors, Nucl. Instrum. Methods Phys. Res., A 595(2008)197–199
70. T. Bellunato et al., Light composite mirrors for RICH detectors: production, characterisation and stability tests, Nucl. Instrum. Methods Phys. Res., A 538(2005)458–464
71. C. Bigongiari et al., The MAGIC telescope reflecting surface, Nucl. Instrum. Methods Phys. Res., A 518(2004)193–194
72. U. Müller et al., The Omega RICH in the CERN hyperon beam experiment, Nucl. Instrum. Methods Phys. Res., A 433(1999)71–76
73. A. Braem, C. David and C. Joram, Metal multi-dielectric mirror coatings for Cherenkov detectors, Nucl. Instrum. Methods Phys. Res., A 553(2005)182–186
74. D. Malacara (Ed.), Optical Shop Testing (Wiley Series in Pure and Applied Optics), Wiley-Interscience, 2007, ISBN: 978-1574446821
D. Malacara, M. Servín and Z. Malacara, Interferogram Analysis For Optical Testing, Second Edition (Optical Engineering), CRC; 2005, ISBN: 978-0471484042
75. The LHCb Collaboration, A. Augusto Alves Jr. et al, The LHCb Detector at the LHC, JINST 3 (2008) S08005.
76. G. Wilkinson, In search of the rings: Approaches to Cherenkov ring finding and reconstruction in high-energy physics, Nucl. Instr. and Meth. A 595 (2008) 228.
77. M. Adinolfi et al., Performance of the LHCb RICH detector at the LHC, Eur. Phys. J. C73 (2013) 2431, arXiv:1211.6759.
78. R. Forty and O. Schneider, RICH Pattern Recognition, LHCb note LHCb-98-040.
79. C. Buszello, LHCb RICH pattern recognition and particle identification performance, Nucl. Instr. and Meth. A 595 (2008) 245.
80. M. Starič, Track based maximum likelihood ring search, Nucl. Instr. and Meth. A 595 (2008) 237
81. LHCb collaboration: R. Aaij et al., Measurement of b-hadron branching fractions for two-body decays into charmless charged hadrons, JHEP10(2012)037
82. E. Fermi, The Ionization Loss of Energy in Gases and in Condensed materials, Phys. Rev. 57(1940)485–493

83. V.L. Ginzburg and I.M. Frank, Radiation of a uniformly moving electron due to its transition from one medium into another, Zh. Eksp. Teor. Fiz. 16 (1946)15.
a shortened (English) version J. Phys. USSR 9 (1945), 353.
84. C.W. Fabjan and W. Struczinski, Coherent Emission of Transition Radiation in Periodic Radiators, Physics Letters 57B(1975)483–486
85. C. Leroy and P-G. Rancoita, Principles of radiation interaction in matter and detection, World Scientific, 2004, ISBN: 978-9812389091
86. M.L.Ter-Mikaelian, High-Energy Electromagnetic Processes in Condensed Media, John Wiley & Sons, Inc, 1972, ISBN 0-471-85190-6
87. G.N. Afanasiev, V.G. Kartavenko and Yu.P. Stepanovsky, Vavilov-Cherenkov and transition radiations on the dielectric and metallic spheres, Journal of Mathematical Physics, 44(2003)4026–4056
88. V.L. Ginzburg and V.N. Tsytovich, On the Derivation of the Transition Radiation Intensity, Physics Letters 79A(1980)16–18
V.L. Ginzburg, On the radiation from uniformly moving sources, Nucl. Instrum. Methods Phys. Res., A: 248 (1986) 13–16
89. Charles Kittel, Introduction to solid state physics. - 7th ed. / New York, NY: Wiley, 1996
90. B. Dolgoshein, Transition radiation detectors, Nucl. Instrum. Methods Phys. Res., A: 326 (1993) 434–469
91. L.C.L. Yuan, C.L. Wang and H. Uto, Formation-Zone effects in transition radiation due to ultrarelativistic particles, Phys. Rev. Lett. 25 (1970) 1513–1515
92. J.E. Lilienfeld, Die sichtbare Strahlung des Brennecks von Röntgenröhren, Physik Zeitschrift, XX(12) 280, 1919
Mario Rabinowitz, Lilienfeld Transition Radiation Brought to Light, physics/0307047 (July 2003)
93. H. Boersch, C. Radeloff and G. Sauerbrey, Experimental Detection of Transition Radiation, Phys. Rev. Lett. 7(1961)52–54
94. X. Artru et al., Experimental investigations on geometrical resolution of optical transition radiation (OTR), Nucl. Instrum. Methods Phys. Res., A: 410 (1998) 148–158
95. J. Ruŝiŝka and J. Mehes, Properties of optical transition radiation for charged particle inclined flight through a finite thick plate IX, Nucl. Instrum. Methods Phys. Res., A: 250 (1986) 491–502, and references therein.
96. S.D. Borovkov et al., On studying the possibility to use optical transition radiation for proton beam diagnostics, Nucl. Instrum. Methods Phys. Res., A: 294 (1989) 101–107
97. J. Ruzicka et al., On optical transition radiation of charged particles in SiO_2-aerogels, Nucl. Instrum. Methods Phys. Res., A: 384 (1997) 387–402
98. X. Artru, G.B. Yodh and G. Mennessier, Practical theory of the multilayered transition radiation detector, Phys. Rev. D 16(1975)1289–1306
99. L. Fayard, Transition radiation, in: Instrumentation en Physique Nucléaire et en Pysique des Particules, les éditions de physiques, 1988, 327–340
100. J. Apostolakis et al., Parameterization models for X-ray transition radiation in the GEANT4 package, Comput. Phys. Commun. 132, 3 (2000) 241–50
V.M. Grichine and S.S. Sadilov, GEANT4 X-ray transition radiation package, Nucl. Instrum. Methods Phys. Res., A: 253 (2006) 299–302
101. F. Sauli, Principles of Operation of Multiwire Proportional and Drift Chambers, CERN 77–09
102. P. Cwetanski, Straw performance studies and quality assurance for the ATLAS transition radiation tracker, Ph.D. Thesis, Helsinki 2006, ISBN 952-10-2122-5
103. The ALICE Collaboration, K. Aamodt et al., The ALICE experiment at the CERN LHC, 2008 JINST 3 S08002
104. The ATLAS Collaboration, G. Aad et al., The ATLAS Experiment at the CERN Large Hadron Collider, 2008 JINST 3 S08003
105. The ATLAS TRT collaboration, E. Abat et al., The ATLAS TRT Barrel Detector, 2008 JINST 3 P02014

106. T. Cornelissen and W. Liebig, ATLAS Inner Detector Results from the 2004 Combined Test Beam Data, Nuclear Physics B (Proc. Suppl.) 172 (2007) 292–295
Tuan Vu Anh, Private communication, February 2009
107. ALICE collaboration, ALICE transition-radiation detector: Technical Design Report, CERN-LHCC-2001-021
108. C. Adler et al., Electron/pion identification with ALICE TRD prototypes using a neural network algorithm, Nucl. Instr. and Meth. Phys. Res. A 552 (2005) 364–371
109. M. Holder and H. Suhr, Separation of electrons and pions by a transition radiation detector using flash ADCs, Nucl. Instr. and Meth. Phys. Res. A 263 (1988) 319
110. A. Andronic, Electron identification performance with ALICE TRD prototypes, Nucl. Instr. and Meth. Phys. Res. A 522 (2004) 40–44
111. ALICE Collaboration, S. Acharya et al., The ALICE Transition Radiation Detector: construction, operation, and performance, arXiv:1709.02743v1 [physics.ins-det] 8 Sep 2017, CERN-EP-2017-222 29 August 2017