4.1. BASIC REQUIREMENTS FOR SENSING IONIZING RADIATION

4.1.1. GENERAL INTRODUCTION

To understand the performance of a sensor or make a choice between different sensors for a detection task it is important to have an understanding of how the recorded signals are generated. In the case of sensing ionizing radiation, signals are generated after the incoming radiation has undergone interactions within the sensor material. The types of interactions and the subsequent resultant particles depend upon the type of incident radiation, its energy, and the material that makes up the sensing medium. This introductory section deals with these interactions and discusses the processes that lead to detectable signals.

There are many textbooks that deal with these aspects in depth. For more details than are given in this introductory section, the interested reader should consult Evans (1955), Heitler (1954), ICRU (1984, 1992), Kember (1994), Knoll (2000), and Smith (2000).

4.1.2. BASIC REQUIREMENTS

Table 4.1.1 lists the possible mechanisms for the detection of different radiations that are relevant to biomedical applications.
Table 4.1.1. Radiation types and the mechanisms for their detection

<table>
<thead>
<tr>
<th>CLASSIFICATION</th>
<th>TYPE OF RADIATION</th>
<th>DETECTION REQUIREMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Electromagnetic</td>
<td>X-rays, γ-rays</td>
<td>Must interact to produce a directly ionizing particle.</td>
</tr>
<tr>
<td>2 Directly ionizing</td>
<td>α-particle, β-particle, protons</td>
<td>Energy is transferred to the detection medium through interactions with the electrons or nucleons of the medium.</td>
</tr>
<tr>
<td>3 Indirectly ionizing</td>
<td>Neutron</td>
<td>Must interact to produce a directly ionizing particle.</td>
</tr>
</tbody>
</table>

Note that for all radiations, the only mechanism for detectable effects in a material is the transfer of energy to the medium.

From Table 4.1.1 it can be seen that unless incident energy is transferred first to ionizing particles and then to the detection medium, an event will not be sensed. It is therefore clear that if electromagnetic or indirectly ionizing radiations are to be detected, it is important that the detection material have a significant cross section for interaction with those radiations so that secondary, directly ionizing particles are produced. Directly ionizing particles will always interact with the detection medium. To maximize the detection probability, it is important that the resultant transfer of particle energy to some method of detection be as efficient as possible. There are two principal methods: chemical changes and collection or movement of charges. Chemical changes have been used in many imaging techniques, although modern methods are now based on the collection of charges. Most quantitative techniques are based on the collection or movement of charges.

If a charge is to be collected or its movement controlled and monitored, then a detector bias will be required. This collection or movement of charge represents the signal that will be used to register the detection of ionizing radiation. The relationship between how much charge is created, that is, the magnitude and development of the signal, and some property of the incident radiation (energy, particle type, location of the interaction, etc) is an important relationship. To understand this relationship it is essential to appreciate the details of how radiation interacts with the sensor material.

4.1.2.1. Sensor Materials

The material of which a sensor is made plays a significant role. Table 4.1.2 presents a range of materials and outlines the reasons why they make good choices as active components for sensors.

Table 4.1.2. Materials used in sensors for ionizing radiation

<table>
<thead>
<tr>
<th>MATERIAL</th>
<th>APPLICATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argon</td>
<td>Fill gas for high-pressure gas detectors</td>
</tr>
<tr>
<td>BaF</td>
<td>A phosphor used in intensifying screens</td>
</tr>
<tr>
<td>CaW</td>
<td>A phosphor used in intensifying screens</td>
</tr>
<tr>
<td>CsI</td>
<td>A scintillator usually doped with Tl or Na</td>
</tr>
<tr>
<td>CdTe</td>
<td>A compound semiconductor used to make compact sensors</td>
</tr>
<tr>
<td>CdZnTe</td>
<td>An improved version of CdTe by the introduction of Zn to help with charge trapping</td>
</tr>
<tr>
<td>GaAs</td>
<td>Used in high-quality spectroscopy systems for photons</td>
</tr>
<tr>
<td>Ge</td>
<td></td>
</tr>
</tbody>
</table>
4.1.3. INTERACTIONS OF CHARGED PARTICLES

For the majority of interactions that are of interest in biomedical sensors for ionizing radiation, the interaction will be a Coulomb interaction between the incident charged particle and the electrons of the medium making up the active volume of the sensor. With this in mind, it is convenient to place these interactions into two categories:

1. Heavy charged particle interactions—protons, ions. In this case, the mass of the incident particle is far greater than the mass of the particle in the material. The form of energy transfer is restricted to multiple small depositions of energy and little change in the path of the incident particle (see Fig. 4.1.1).

2. Light charged particles—electrons, positrons. In this case, the mass of the incident particle matches that of the particle in the material. Energy transfers can be significant and will certainly cause large changes in the incident particle trajectory. Again, the total loss of particle energy will be made up of a very large number of individual interactions (see Fig. 4.1.1).

Other mechanisms of interaction are possible, but rarely have any relevance to the detection of ionizing radiation except in the case of neutron detection. One of the important characteristics of the mechanism of energy transfer is the distance travelled during this process. To discuss this property it is important to distinguish the distance travelled along the track (pathlength) and the distance that the particle penetrates into the material (range). There are a number of different definitions for range.

4.1.3.1. Particle Range

If the number of particles transmitted through a material, \( N \), is counted in an experiment, as shown in Figure 4.1.2, then the plot of \( N \) versus the thickness of the material, \( x \), can be characterized by several parameters. The most important quantity is the mean range, \( r_0 \), the thickness to reduce the number of particles by half. The continuous curve plot in Figure 4.1.2 applies to heavy charged particles and exhibits a nearly constant fraction of transmission for a considerable thickness of material. For light charged particles (shown as the broken curve plot in Fig. 4.1.2), the tortuous path leads to a different

---

**Figure 4.1.1.** Particle tracks within the medium of a sensor. The heavy charged particle track (a) is almost straight, whereas the light charged particle (b) follows a very tortuous track.
shape for the transmission curve. However, the same definition of mean range can be applied. Typical values for $r_0$ are given in Table 4.1.3.

4.1.3.2. Relevance of Particle Range in Selecting a Sensor

When selecting a sensor it is important that the thickness or volume of the sensor is large enough to ensure that most, if not all of the energy associated with an ionizing photon or particle is deposited within the active part of the sensor. Clearly the range of the incident particle or the secondary electrons created by an incident photon plays an important role in selecting the thickness. In general, if photons are being detected, the major consideration for the detector volume is to provide enough material to ensure that the probability of interaction is acceptably high. Any volume chosen on this basis will be more than sufficient to cover the range of the secondary charged particles, particularly in a solid-state detector material. However, particle beams can use much smaller volumes for detection, as can be seen from Table 4.1.3.

4.1.3.3. Collisional and Radiative Energy Transfer

Energy loss mechanisms within a medium can be split into two broad categories: radiative, where energy is transferred to electromagnetic radiation, and collisional, where the transferred energy creates excitations and ionizations of the medium. Radiative loss is only significant for energetic electrons

Table 4.1.3. Range values for different sensor materials for electrons of different energies

<table>
<thead>
<tr>
<th>ENERGY (keV)</th>
<th>CARBON</th>
<th>SILICON</th>
<th>GERMANIUM</th>
<th>CESIUM IODIDE (CsI)</th>
<th>SODIUM IODIDE (NaI)</th>
<th>PHOTOEMULSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1.66</td>
<td>1.49</td>
<td>0.92</td>
<td>1.39</td>
<td>1.53</td>
<td>1.22</td>
</tr>
<tr>
<td>100</td>
<td>94.2</td>
<td>78.2</td>
<td>44.7</td>
<td>60.2</td>
<td>69.2</td>
<td>58.8</td>
</tr>
<tr>
<td>1000</td>
<td>2920</td>
<td>2312</td>
<td>1230</td>
<td>1605</td>
<td>1879</td>
<td>1650</td>
</tr>
</tbody>
</table>

All values are in micrometers.
passing through high atomic number materials. Since most biomedical sensing applications involve relatively low-energy secondary charged particle creation, collisional losses are the major mechanism for transferring energy to the material of the sensor. The rate at which energy is transferred as the particle moves through the sensor is another important property of the particle and the medium that makes up the sensor. This quantity is called the stopping power.

### 4.1.3.4. Stopping Power and Its Relevance to Selecting a Sensor

Stopping power \( \frac{dE}{ds} \) is defined as the rate of energy loss per unit path length, \( s \). Developing a theoretical model to describe this process has been successfully undertaken for both classical as well as quantum mechanical approaches. The quantum mechanical result is often referred to as the Bethe equation and is given as

\[
\frac{dE}{ds} = \frac{4\pi r^2 mc^2}{\beta^2 A} \frac{\rho Z z^2}{u} \left[ \ln \left( \frac{2mc^2\beta^2}{(1 - \beta^2) I} \right) - \beta^2 \right]
\]

for heavy charged particles and

\[
\frac{dE}{ds} = \frac{2\pi r^2 mc^2}{\beta^2 A} \frac{\rho Z}{u} \left[ \ln \left( \frac{T}{I} \right)^2 + \ln \left( 1 + \frac{\tau}{2} \right) + F(\tau) - \delta \right]
\]

for electrons. Here, \( r \) is the classical radius of the electron and \( \beta \) is the ratio of the particle velocity to the speed of light. Also in these equations, \( u \) is the atomic mass unit, \( Z \) is the atomic number of the medium, \( A \) is the atomic weight, \( I \) is the mean excitation energy (describing the average energy lost to the medium at each Coulomb interaction), \( T \) is the kinetic energy of the incident particle, and \( \tau = T/mc^2 \). \( \delta \) is a correction for the so-called density effect, and \( F \) is a factor that allows for the difference in behavior between electrons and positrons. All quantities in the square brackets are slowly changing functions of energy, and hence the major behavior is governed by the nonconstant values in front of the square brackets. It can be seen that for both heavy charged particles and electrons, the rate of loss of energy per unit path length is proportional to \( \rho \left( \frac{Z}{u} \right) \), where \( z \) is the charge on the particle (1 for electrons), \( v \) is the velocity of the particle, and \( \rho \) is the density.

### 4.1.3.5. Bragg Curve

One of the effects of the dependence of stopping power on particle velocity is that energy loss is not uniformly distributed along the path length, but is much less when the particle first moves through the material than when it is near the end of its path. This observation, shown graphically in Figure 4.1.3, is called the Bragg curve. Note that if all the energy associated with an event is to be deposited within a given sensor, it is most important that the end of the Bragg curve be included, as this may constitute a significant proportion of the total energy.

### 4.1.3.6. Minor Effects of Charged Particle Interactions

There are several interesting features of charged particle interactions that occur at low energies. As particles (primary events or secondary charged particles) lose energy along their path, a point will come
when their energy is either insufficient to ionize atoms effectively or their velocity becomes comparable
to the orbital velocity of the atomic electrons in the medium of the sensor. Both of these will affect
the total charge created by an event. The results of these effects are discussed in the appropriate sections
on dosimetry.

4.1.3.7. Summary of Charged Particle Interactions

The following are important points:

- Ionizing radiation can only be detected or sensed by the effects of charged particle interactions.
- If the radiation to be detected is electromagnetic (X- or $\gamma$-rays), then the energy associated with
  the quanta must first be converted to charged particles—generally electrons or positrons in the
  case of biomedical applications.
- Important parameters such as range and stopping power (and to a lesser extent the Bragg curve)
  contribute to the choice of a suitable material for a sensor.
- For biomedical applications, charged particles will travel only short distances within a sensor material.

4.1.4. INTERACTIONS OF X- AND GAMMA RAYS

X- and gamma rays are part of the electromagnetic spectrum and are generally referred to as photons,
so it is assumed that most of their behavior is best described by particle-like characteristics. This is true
for the important interactions insofar as biomedical sensors are concerned. In biomedical applications,
the energy range of interest is usually from 1 keV to 25 MeV, and within this range there are four
processes by which the photon can interact. Some, but not all, of these processes result in the creation
of secondary charged particles that can lead to detection of the photon.

When a beam of many photons enters a material, some of the photons will be transmitted and
some will interact. This process is described by considering the absorption and scattering of the beam
leading to the concept of attenuation. If the material being considered is a sensor, then the attenuation
provided by the sensor will lead to the concept of efficiency, usually referred to as quantum efficiency. Consider the sensor in Figure 4.1.4. The incident beam provides an intensity (energy passing through unit area in unit time) of $I_0$. Some of the photons in the beam do not interact in the sensor and pass through as intensity $I$. Some interact and deposit all their energy in the sensor. Others interact and may or may not deposit some energy in the sensor, but will have their direction of travel changed and thus become scattered photons. An ideal sensor has properties that produce no transmitted beam, with all the scattered photons being reabsorbed within the sensor. This would lead to 100% quantum efficiency if all the deposited energy could be measured.

Transmitted and incident intensity are related by $I = I_0 e^{-\mu x}$, where $\mu$ is the total linear attenuation coefficient of the sensor material and $x$ is the thickness. Note that this relation does not describe the level of scattered radiation. However, it is clear that thick, large-volume sensors decrease the level of both transmitted and scattered radiation and increase the quantum efficiency. To fully understand the different contributions to the energy deposited in the sensor, and hence its output, it is necessary to consider the different types of interactions. For each interaction type there are two aspects to consider. First, the energetics of the interaction; that is, if an interaction takes place, what actually happens? Second, a discussion about the factors that govern the probability of the interaction taking place.

4.1.4.1. Photoelectric Interaction

4.1.4.1.1. Energetics of Photoelectric Interactions

When an incident photon interacts with a bound electron in the sensor material, the energy associated with that photon is transferred to the electron. This energy is used to overcome the binding energy so as to eject the electron. The ejected electron, known as a photoelectron, has energy equal to the incident photon energy minus the binding energy it had when in the atomic structure. This energetic secondary electron loses energy, as described in section 4.1.3. The atom from which the electron has been removed is unstable, having an inner shell electron missing, and so rearranges its electron structure to maintain stability. This is achieved by an electron from another shell or subshell moving
into the vacancy: this move is governed by quantum mechanical selection rules. The result of this rearrangement means that further energy becomes available when the energy of the electron states is taken into account. This energy can either give rise to a characteristic X-ray (usually the case when the nucleus of the atom concerned has a high atomic number) or an Auger electron. An Auger electron is an intershell vacancy transfer that is produced when the energy made available is given directly to one of the outer shell electrons, causing it to be ejected by the atom. Since most sensors are made of high Z materials (see Table 4.1.2), characteristic X-ray generation is the most likely outcome of a photoelectric interaction in the sensor. In order for all the energy associated with the photon to be captured within the sensor, it means that either these characteristic X-rays must be stopped in the sensor or the result of the primary interaction must be to produce an Auger electron. If neither of these conditions is met, then only part of the energy will be deposited.

The effects of this are most significant when considering spectroscopy (see section 4.4). If all the energy is left in the sensor, the signal generated will be directly related to the incident photon energy. It is worth noting that as characteristic photons are created in high Z materials, they are not likely to travel great distances in such a material. Furthermore, if the interaction is in a low Z material, the “excess” energy appears as an Auger electron. Thus, whether the interaction is in a high or low Z material, the secondary radiations are quickly stopped. This often leads to the assertion that, provided the volume of the material in which a photoelectric interaction takes place is on the order of a few cubic millimeters, the photoelectric interaction is one of total absorption. However, this statement should not be interpreted to say that a sensor needs to be only a few cubic millimeters in volume to be an effective detector.

4.1.4.1.2. Probability of Interaction

Theoretical descriptions of the photoelectric interaction can be used to develop the dependence of the interaction upon the atomic number of the sensor medium Z, the incident photon energy $E$, and the physical density of the sensor medium $\rho$. For the energy range of interest in biomedical sensors, the relationship is often simplified as $\tau \propto \frac{Z^3}{E^3}$, where $\tau$ is the photoelectric linear attenuation coefficient (the probability that a photoelectric interaction will take place per unit thickness of the sensor material). In practice, noninteger powers are required that also have a dependence on $E$ and $Z$ (White, 1977). Clearly, if the probability increases, then so does the quantum efficiency of the sensor. It can be seen that higher Z materials and materials with high density will produce a high probability of interaction, and hence detection. It is also clear that lower energy photons are easier to detect.

The relationship given in the previous paragraph for $\tau$ is true for the interaction of an incident photon and a given shell or subshell electron that takes part in the interaction. When interactions from a beam of photons of differing energies are considered, the effect of absorption edges must be taken into account. Figure 4.1.5 plots the value of $\tau/\rho$ for different sensor materials over the energy range of 10 keV to 100 keV. These data have been taken from XCOM (see http://www.nist.gov). The effects of absorption edges are clear in this figure. Silicon (Si) and germanium (Ge) have similar values of $\tau/\rho$ at 10 keV, but because of the absorption edge for Ge at 11 keV, Ge has a value of $\tau/\rho$ an order of magnitude greater above 11 keV. Similar effects can be noticed for CsI above 35 keV. The performance of a sensor can be significantly altered by the presence or otherwise of an absorption edge; hence, for increased sensitivity it is always advisable to choose a sensor material with an absorption edge below the photon energy to be detected.
4.1.4.2. Rayleigh Scattering

Rayleigh scattering is an interaction process whereby no energy is deposited in the material and hence, in itself, does not allow a photon event to be detected. However, it can play an important role if absolute measurements of event rates are to be determined because it will need to be taken into account in any such measurement.

4.1.4.3. Compton Effect

4.1.4.3.1. Energetics of Compton Scattering

Compton scattering is an inelastic scattering process whereby a photon loses part of its energy to the medium of the sensor and the reduced energy photon is scattered in a new direction. For this event to take place, the incident photon energy must significantly exceed the binding energy of the electron in the medium. This condition is satisfied in many sensor materials for incident photons greater than 50 keV, and above this energy, this will become significant. The process is usually described as a “billiard ball” type of collision, where the incident photon, energy $E$, scatters off the atomic electron at an angle $\phi$. The scattering event imparts energy $T$ to the electron, which leaves in some direction $\theta$. There are various relationships that can be derived from these energetics, but for understanding the performance of a sensor, the most important is that $E = T + E'$, where $E'$ is the scattered photon energy. This simple result indicates that not all the incident photon energy is deposited in the medium. If $E'$ is significant, the scattered photon may have sufficient energy to leave the volume of the sensor. Under these conditions only part of the incident photon energy is recorded. This can have two consequences. It may mean that insufficient energy is deposited for the event to be recognized above the noise of the system, or, if spectroscopy is to be undertaken, the event will be registered with the wrong energy.

4.1.4.3.2. Probability of Compton Scattering

Theoretical understanding of the Compton interaction was developed by Klein and Nishina (1929). Their work developed differential collision and scattering cross sections that must be integrated to

![Figure 4.1.5. Photoelectric attenuation data demonstrating the increase in probability at particular energies. These are called absorption edges. Data taken from XCOM (http://www.nist.gov).](image-url)
obtain Compton attenuation coefficients. The result of these integrations is that \( \sigma \propto f(E) \rho \), where \( \sigma \) is the linear Compton attenuation coefficient, \( \rho \) is the physical density of the sensor medium, and \( f(E) \) is a function that describes the variation of \( \sigma \) with \( E \). \( f(E) \) is a slowly varying function of \( E \), falling faster as \( E \) increases. Up to values of 100 keV it is often assumed to take on a constant value of approximately 0.2 \( \rho \) cm\(^{-1}\) when \( \rho \) is measured in grams per cubic centimeter (g cm\(^{-3}\)). This simplification can be useful to estimate the sensor thickness required to ensure adequate quantum efficiency.

4.1.4.4. Pair Production

4.1.4.4.1. Energetics of Pair Production

Pair production is an event that only occurs at incident photon energies in excess of 1.02 MeV. The process involves the incident photon energy being converted into matter within the favorable conditions offered by the electrostatic force of the nucleons. This force is sensed close to the nucleus and its magnitude can be modified by inner shell electron effects in large, complicated nuclei. The conversion into matter requires at least sufficient rest mass energy to create the resultant particles. At energies in the biomedical range of interest, this means the creation of an electrically neutral pair of low-mass particles, that is, an electron and a positron. Each has a rest mass energy of 0.511 MeV, hence the threshold for this effect to be observed is 1.02 MeV. Energy in excess of this threshold is divided approximately equally between the particles. Both particles move off through the medium, undergoing Coulomb interactions and losing energy, as described in section 4.1.3. At very low, almost zero energy there is a marked difference in the behavior of the electron and positron. The electron takes up a position in the outer shell of an atom. The positron, on the other hand, does not exist naturally and is annihilated with an electron in the medium, turning matter back into energy. If the positron had effectively zero kinetic energy, and likewise the electron with which it is annihilated, the result of the annihilation is two 0.511 MeV photons traveling in opposite directions. These are relatively energetic photons and therefore have a finite probability of escaping the confines of the detector. If this is the case, their energy will not be deposited within the detector, affecting the magnitude of any detected signal.

4.1.4.4.2. Probability of Interaction

Theoretical models developed to study the pair production interaction draw parallels with those describing the photoelectric interaction. In the case of the photoelectric interaction, an electron (the photoelectron) is ejected from a positive energy state. In the case of pair production, the electron is ejected from a negative energy state. This gives rise to the electron, and the vacancy left in the negative energy state appears as a positron in the laboratory frame of reference. Hence the pair is produced. This allows the interaction to be handled in a similar way to the photoelectric interaction. The result is that \( \kappa \propto Z \rho \ln E \), where \( \kappa \) is the probability of pair production per unit thickness of material, \( \rho \) is the physical density of the material, and \( E \) is the incident photon energy. It is important to note that as the energy increases, so does the probability of interaction. Clearly high \( Z \) and high \( \rho \) will increase the quantum efficiency of a detector in this energy range.

4.1.4.5. Summary of Photon Interactions

The following points are most important to remember:

- Detection of photons requires conversion of the photon into charged particles. Thus only those interactions that create charged particles are useful for event detection.
• Only if all the photon energy is converted to charged particles can the energy of the photon be
determined.
• Secondary photons that escape the detector reduce the detected signal.
• Photoelectric interactions provide the best opportunity for full energy deposition.
• Too many Compton interactions generally lead to poor detector performance.
• Combining all the interactions that can take place gives the overall energy response that can be
expected from a given sensor material.

4.1.5. NEUTRON INTERACTIONS
4.1.5.1. Energetics of Neutron Interactions
There are three possible mechanisms for neutron interactions:
• The neutron is scattered by the nuclear potential of the medium through which it is passing
(direct scattering).
• The neutron enters the nucleus to form a compound nucleus and the subsequent deexcitation
can be either elastic (a neutron of the same energy is emitted), inelastic (a neutron of lower energy
and one or more gamma rays or particles are emitted), or radiative (only gamma rays are emit-
ted, often leaving an unstable nucleus that will decay).
• Breakup of the nucleus can occur at very high energies—this is not relevant to biomedical sensors.

Of these, the most useful for neutron detection are reactions such as \( (n,\alpha) \) and \( (n,p) \), where part
of the neutron energy is converted to heavy charged particles. Reactions leaving gamma rays are not
so useful because the gamma rays can be difficult to detect due to their energies. These reactions only
occur at “slow” neutron energies (i.e., less than 0.5 eV), so it is often necessary to slow down (moder-
ate) higher energy neutrons using hydrogenous materials.

4.1.5.2. Probability of Interaction
Entering a nucleus to form a compound nucleus is a resonance phenomenon that occurs with greater
probability when the incident neutron energy matches the rest-mass energy plus the excitation energy
of one of the resonance states. Outside this region it can be shown that \( \sigma_{\gamma} \propto \frac{1}{v} \), where \( \sigma_{\gamma} \) is the
probability that radiative capture will take place and \( v \) is the velocity of the incident neutron.

4.1.6. EFFECTS OF INTERACTIONS ON SENSOR OPERATION
The output from a sensor is governed by the mechanisms that have deposited energy within the sen-
sor volume. Whether the primary radiation consists of charged particles or electromagnetic radiation,
interactions within the sensor create further charged particles that are then responsible for forming
the signal. This process of forming the detector response can be chemical or a mechanism to either
directly or indirectly sense the charge. If quantitative estimates of the energy deposited are required,
then important parameters need to be considered. The amount of energy it takes to create ionization
describes the ionization potential of the sensor material. The average energy required to form an ion
pair or charge pair is usually the most relevant parameter when considering how the detected signal
might be used in dosimetry. This quantity is usually called the $W$-value in gases or the $w$-value in semiconductor materials.

4.1.6.1. Ion Pair or Electron-Hole Creation

An important parameter when selecting a sensor is how much energy, on average, is required to produce an ion or charge pair. Knowledge of this quantity is essential in dosimetry, but it also dictates the statistical precision with which the event can be characterized. This translates into the energy resolution in spectroscopy and the low contrast performance of imaging detectors.

A further effect to be considered is when light charged particles are responsible for creating the sensor response (as in most cases). The value of $w$ (or $W$) changes as the charged particle energy falls to low values (because there is a greater probability of creating excitation rather than ionization) and hence needs to be considered.

4.1.6.2. Charge Collection

Sensors whose output depends upon charge collection can usually be operated in one of two basic configurations: integrating or pulse mode. In an integrating detector, the total charge created by all events that occur in the detector in a given time is summed to form the output. Thus instantaneous evaluation of this current gives an output expressed as a rate (e.g., dose rate or exposure rate). In pulse mode, the charge associated with each event is recorded.

Charge is collected by forming an electric field gradient within the sensor and allowing the charge to drift under its influence. Alternatively, and sometimes in conjunction with drift, a potential well is created within the sensor to hold charge within a region. If operating with drift, at the point when the ion or charge pair is formed, the positive and negative components induce signals on the anode and cathode of the sensor. If operated in pulse mode, the rates at which the charges drift to their respective electrodes create the temporal response of the sensor. If operated in integrating mode, a near steady state of charge flow is set up, and only at the start and finish of the total exposure are any effects due to drift differences noticeable.

The shape of the electric field distribution within the sensor is critical to the detector output. In some gas detectors (ion chambers, for example) a uniform electric field strength is desirable. On the other hand, the rapidly changing electric field strength inside a Geiger-Mueller tube or proportional counter enables the moving charge to gain energy, create further charge through collisions, and hence increase the output of the sensor. This increase can either be controlled and predictable (proportional gas counter) or in the form of an avalanche (Geiger-Mueller). Similar considerations apply to solid-state detectors, although only recently have multiplying effects within the sensor been used. Generally, uniform fields are used to maintain good charge collection efficiency, leading to excellent spectroscopic performance.

4.1.6.3. Photon Counting, Spectroscopy, and Integration

If the sensor signal is derived from the collection of charge, then the signal processing electronics as well as the electrical characteristics of the sensor govern the final signal used to derive the output (see Fig. 4.1.6). Processing electronics with a fast time constant ($R \times C$) compared to the charge collection time within the sensor can effectively follow the temporal development of the charge in the sensor.

Alternatively, a long time constant allows the full charge packet associated with the event to be presented to the electronics before any part of the signal transits through any processing stages. The former case is called an integrating system and the latter is called a pulse counting or photon counting
system. The temporal responses are shown in Figure 4.1.7. The total deposited energy is converted to charge, as described in sections 4.1.3 to 4.1.5. This charge forms an instantaneous current $i(t)$ that is presented to the processing electronics. Two cases are shown for the output from the processing electronics, where the value of the time constant is short or long compared to the time to collect charge, $t_c$. When $RC \gg t_c$ it is possible to relate the value of $V_{max}$ to the total energy deposited in the sensor. If the sensor is carefully chosen, this value should represent the total energy of the event.

**4.1.7. DEFINITION OF QUANTITIES RELATED TO RADIATION DOSIMETRY**

Since the beginning of research in both radiotherapy and X-ray imaging, an attempt has been made to identify quantities that could be related to the risk for the patient and to develop methods for

![Figure 4.1.6. Schematic diagram of a sensor and processing electronics. The value of C includes any capacitance associated with the sensor. The output of the sensor appears as a voltage across the load resistance R.](image)

![Figure 4.1.7. The top figure represents the current flowing in the sensor during the charge collection period. The other graphs show the output from the processing electronics with different values of RC.](image)
measuring them. Delivered dose—the energy imparted per unit mass—is currently believed to be the most effective quantity for estimating patient risk. The need for high-precision dosimetry is particularly relevant in radiotherapy, the goal of which is to maximize the damage to the neoplastic tissue while minimizing the damage to healthy tissue.

It is well known that tumor control probability (TCP) and normal tissue complication probability (NTCP) are both increasing functions of the delivered dose, but they have different increase rates, the slope being steeper for normal tissue (AAPM, 2004), as shown in Figure 4.1.8. A correct estimation of the delivered dose is therefore crucial for effective tumor control without severely damaging healthy tissue. An underestimation of delivered dose can result in a complete control of the tumor, but also in damage to normal tissue, while an overestimation of the delivered dose can result in inadequate tumor control.

As for imaging dosimetry, there are less stringent guidelines stating that the delivered dose to the patient must be kept “as low as reasonably achievable” (ALARA), but correct dosimetry can still be of crucial importance because of the risk of radiation-induced carcinogenesis. This is particularly the case in screening procedures, such as mammography, aimed at identifying lesions in asymptomatic and thus potentially healthy subjects. In such cases the risk of the procedure has to be carefully balanced against the expected benefits.

This section focuses mainly on photon and electron radiotherapy dosimetry and on diagnostic dosimetry, although some applications to other fields, such as environmental dosimetry, will occasionally be mentioned. The main quantities of dosimetric interest will be introduced and the main requirements for dosimetric measurements will be discussed.

4.1.7.1. Quantities of Dosimetric Interest
4.1.7.1.1. Photon Fluence and Energy Fluence
Photon fluence is a measure of the number of photons impinging on a unit area. The definition of photon fluence (ICRU, 1980) is

$$\Phi = \frac{dN}{da},$$  \hspace{1cm} (4.1.1)

where $dN$ is the number of photons incident on a sphere of cross-sectional area $da$.

![Figure 4.1.8. Tumor control probability (TCP) and normal tissue complication probability (NTCP) as a function of dose.](image)
A quantity related to photon fluence is energy fluence, which is defined in a similar way as

\[ \Psi = \frac{dR}{da}, \tag{4.1.2} \]

where \( dR \) is the radiant energy entering a sphere of cross-sectional area \( da \) (ICRU, 1980).

4.1.7.1.2. Exposure

Since the earliest interest in dosimetric issues, it appeared a good idea to measure a beam by the charge it created in a medium. In particular, the first dosimetric quantity—exposure—was related to the charge released in air. This was convenient because air, with an atomic number of 7.6, is a tissue equivalent (tissue has an atomic number of 7.4). Another obvious advantage was the availability of air.

Exposure is defined as

\[ X = \frac{dq}{dm}, \tag{4.1.3} \]

“where the value of \( dq \) is the absolute value of the total charge of the ions of one sign produced in air when all the electrons (negatrons and positrons) liberated by photons in air of mass \( dm \) are completely stopped in air” (ICRU, 1980).

Consider this definition. As discussed in section 4.1.4, photons interact with air, producing electrons by photoelectric and Compton effects and, at energies above 1.022 MeV, by electron/positron pair production. These charged particles travel across the medium, in turn producing further ionization. The “total charge” referred to in the definition is the total charge produced by them along their path before they are stopped. They may also produce bremsstrahlung photons, which are subsequently reabsorbed in the medium. The ionization produced by charged particles deriving from bremsstrahlung photons must not be taken into account; in other words, the interaction coefficient related to exposure is the energy absorption coefficient, not the energy transfer coefficient.

The SI unit for exposure is coulomb per kilogram (C kg\(^{-1}\)), but the historical unit roentgen (R) is still accepted: 1 R = 2.58 × 10\(^{-4}\) C kg\(^{-1}\).

4.1.7.1.3. Kerma

The acronym kerma stands for “kinetic energy released per unit mass.” Kerma for a nondirectly ionizing beam (i.e., photons and neutrons) in a given material is given by

\[ K = \frac{dE_{\text{e}}}{dm}, \tag{4.1.4} \]

where \( dE_{\text{e}} \) is the sum of the kinetic energies of all charged particles initially released in a given mass and \( dm \) is the mass itself.

---

1 Bremsstrahlung, or “braking radiation,” is the process by which charged particles moving in a medium lose energy, as electromagnetic radiation, through interaction with the nuclei of the medium. The power emitted is proportional to the effective atomic number of the medium and to the particle energy; hence, for soft tissue, characterized by low \( Z \), it is negligible up to a few megaelectron volts. On the other hand, it is the main process involved in the production of diagnostic X-rays, where electrons of 10 keV to 150 keV are accelerated onto a high-\( Z \) material such as tungsten or molybdenum.
The main conceptual difference between kerma and exposure lies in the fact that kerma is related to the initial kinetic energy of the charged particles released by noncharged particles: that is, it takes into account the energy that could subsequently be lost in the bremsstrahlung process. For a beam of energy fluence $\Psi$, $K = \Psi \mu_{tr}/\rho$, where $\mu_{tr}/\rho$ is defined as the mass energy transfer coefficient of the material of interest. Kerma is not completely defined unless a material is specified. Finally, kerma is not only defined for a photon beam, but also for a neutron beam. The SI unit for kerma is the gray (Gy): 1 Gy = 1 J kg$^{-1}$.

4.1.7.1.4. Absorbed Dose

Absorbed dose was introduced for measuring the effects of radiation in neutron beams because the initially defined quantity—exposure—referred only to photons. It is also used for directly ionizing radiation.

Absorbed dose at a point is defined as

$$D = \frac{d < \varepsilon >}{dm},$$

where $d < \varepsilon >$ is the mean energy imparted by ionizing radiation to a material of mass $dm$. The presence of a mean value is due to the peculiar mechanism of energy deposition, which for photons and neutrons is a discrete process. The SI unit for absorbed dose is the gray (Gy). The gray has replaced the rad: 1 rad = 100 erg g$^{-1}$ = 10$^{-2}$ Gy.

Consider the differences between kerma and absorbed dose. Kerma describes the energy initially released when the radiation beam interacts with matter, regardless of the processes that take place afterward, whereas absorbed dose describes the energy absorbed from the secondary particles created by the radiation beam in the medium in question.

To better understand this mechanism, consider a photon beam entering a medium and assume first that the beam is not attenuated by the medium and that the fraction of energy lost in bremsstrahlung is negligible (Fig. 4.1.9a). In this case, kerma is constant across the thickness of the material because the number of electrons produced per unit length is constant. Before producing ionization, and hence contributing to absorbed dose, secondary electrons travel a distance related to their range in the medium. For this reason, dose is small close to the entrance surface of the beam and increases with the depth in the medium. This mechanism is called dose buildup. At a certain depth the ionization produced within a layer of material by electrons generated outside it is equal to the ionization produced outside the layer by electrons generated within the layer, and the absorbed dose is equal to kerma. This is the condition of charged particle equilibrium.

When the beam is attenuated by the medium (Fig. 4.1.9b), kerma decreases exponentially. Dose buildup occurs until the secondary electron fluence is high enough to provide an energy deposition larger than the energy transfer described by kerma. At greater depth, dose buildup is compensated for by attenuation of the photon beam, and absorbed dose starts decreasing at the same rate as kerma. This regime is called transient charged particle equilibrium.

The distinction between kerma and dose is more relevant for therapy beams than for diagnostic beams, as dose buildup is negligible for diagnostic beams.

If the fraction of energy lost in the bremsstrahlung process is not negligible, the relationship between kerma and dose at charged particle equilibrium is

$$D = K(1 - g) = \frac{\mu_{en}}{\rho} \Psi,$$

(4.1.6)
where \( g \) is the fraction of energy lost in the bremsstrahlung process, and the mass energy absorption coefficient, \( \mu_{\text{en}}/\rho \), is correlated to the mass energy transfer coefficient by \( \mu_{\text{en}}/\rho = (1 - g)\mu_{\text{tr}}/\rho \).

At charged particle equilibrium

\[
D_{\text{air}} = X \frac{W_{\text{air}}}{e},
\]

(4.1.7)

where \( W_{\text{air}} \) is the mean energy required to produce an electron-ion pair in air, and \( e \) is the electronic charge.

4.1.7.1.5. Dose Equivalent

Absorbed dose is ineffective for describing the actual biological effects of radiation. For this purpose, dose equivalent is defined as

\[
H = Dw_R,
\]

(4.1.8)

where \( D \) is the absorbed dose and \( w_R \) is the so-called radiation weighting factor, taking into account the beam type. Table 4.1.4 summarizes the values of \( w_R \) for different types of radiation. It can be seen that dose and dose equivalent are numerically the same only for photon and electron irradiation, and that the biological effect is greater for all other beam types.

Although \( w_R \) is a dimensionless quantity, the unit for the dose equivalent is not the same as the unit for dose (i.e., Gy), but is the sievert (Sv): 1 Sv = 1 J kg\(^{-1}\).

4.1.8. DEFINITION OF QUANTITIES RELATED TO RADIATION IMAGING

This section considers the main parameters that should be taken into account when evaluating an imaging detector. Only the basic concepts will be discussed, as a comprehensive analysis is far beyond the scope of this chapter and a vast amount of literature is available on the subject.
4.1.8.1. Overall Detector Characteristics: Area Coverage, Uniformity, Stability, and Linearity

Before addressing some specific quantities in detail, it is worth discussing some general features that apply to all detectors and provide the primary lines along which the choice of a detector should be made, according to the targeted application. Among these, area coverage, uniformity, and stability are rather obvious but nonetheless primary issues. Regarding area coverage, not only the size of the sample that has to be imaged must be taken into account, but also the expected degree of magnification, in case a nonnegligible distance between the sample and detector is foreseen. This is often the case, as scattering rejection techniques can be based on air-gap methods, and magnification itself is frequently employed to increase the system spatial resolution. Area coverage values are standardized for most radiological applications (35 × 43 cm² for chest radiography, 18 × 24 cm² for mammography, etc.), while in the biological field they are much more application driven.

Uniformity refers to the ability of each detector element (pixel) to provide the same response when exposed to the same amount of radiation. A nonuniform response results in fixed pattern noise (FPN), that is, the image obtained by exposing the system to a uniform radiation field is not perfectly flat. Different systems have different degrees of uniformity (for instance, complementary metal-oxide semiconductor [CMOS]-based sensors are historically regarded as more subject to FPN than charge-coupled devices [CCDs]), but a certain degree of nonuniformity is encountered in practically all systems. This is usually corrected by means of simple algorithms, and the higher the system stability, the more effective the result of the correction procedure. System stability is the ability of the system to provide the same response to the same input over different times, which is clearly a basic requirement in order to effectively apply the above corrections.

Correction algorithms depend on another basic property of the system, namely its linearity, which refers to the direct proportionality between the radiation input and the detector response. If the system is linear, the correction requires the acquisition of a flat field (i.e., the system response to a uniform radiation field, or simply to a uniform object if one wants to correct for the nonuniformities due to the radiation source) and a dark field image (i.e., an image without any radiation impinging on the detector), providing slope and intercept values, respectively, to the (linear) correction function, which

---

**Table 4.1.4. Radiation weighting factors for different radiation beams**

<table>
<thead>
<tr>
<th>Type of Radiation</th>
<th>( w_r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photons</td>
<td>1</td>
</tr>
<tr>
<td>Electrons</td>
<td>1</td>
</tr>
<tr>
<td>Neutrons, energy</td>
<td></td>
</tr>
<tr>
<td>&lt; 10 keV</td>
<td>10</td>
</tr>
<tr>
<td>10 keV to 100 keV</td>
<td>20</td>
</tr>
<tr>
<td>&gt; 100 keV to 2 MeV</td>
<td>10</td>
</tr>
<tr>
<td>&gt; 2 MeV to 20 MeV</td>
<td>5</td>
</tr>
<tr>
<td>&gt; 20 MeV</td>
<td>5</td>
</tr>
<tr>
<td>Protons, energy &gt; 2 MeV</td>
<td>5</td>
</tr>
<tr>
<td>( \alpha ) particles</td>
<td>20</td>
</tr>
</tbody>
</table>

ICRU, 1980.
is then applied on a pixel-by-pixel basis. If the system is not linear, more images have to be taken at intermediate exposure levels, and the appropriate (nonlinear) correction function is obtained by interpolating or fitting the acquired data.

4.1.8.2. Dynamic Range and Related Topics

The dynamic range is usually defined as the ratio between the maximum signal achievable in the individual detector pixel and the amount of signal that is stored in the individual pixel because of noise. Both the noise due to the detector and that due to the X-ray source have to be taken into account, and they are usually summed in quadrature. The phenomena of X-ray emission and interaction are subject to fluctuations described by Poisson statistics, in which the variance on a number of quanta is equal to the number of quanta. Hence this is an intrinsic limit that cannot be overcome and to which any further source of noise due to the detector system has to be added (in quadrature). Details on detector noise modeling are provided later, but it should be noted at this point that the dynamic range properties of a system benefit substantially from the design of detectors with optimum noise performance.

The maximum achievable signal depends on the specific detector design, for example, on the well capacity (i.e., how many electrons the single potential well can contain) in a CCD detector. However, with digital detectors it is possible to take two subsequent exposures and sum the resulting images, thus increasing this value; this possibility is clearly not available in film imaging.

In this framework it should be noted that in photon counting devices, in order to enable a proper counting of every single photon, any source of noise due to the detector system has to be eliminated by proper threshold settings. In such systems the only source of noise is due to the statistical nature of X-ray interaction, which means that the noise performances are kept at their theoretical maximum. Moreover, as each single photon is counted, there are no restrictions related to the pixel capacity, as occurs in integrating systems, and the maximum achievable signal depends only on the number of bits in the integrated counter (usually $2^{16}$, but higher capacities can be implemented). As a consequence, in counting systems the dynamic range can be pushed to the maximum values. Nevertheless, many applications are still based on integrating devices, as the ability of counting systems to handle high photon fluxes is still a subject of discussion. Moreover, many recently developed integrating systems feature a level of detector noise so small that it can be considered almost negligible when summed in quadrature with the X-ray Poisson values.

Finally, it should be mentioned here that, for practical purposes, the main image quality parameter is the signal-to-noise ratio (SNR), which in an image is defined as the contrast of a detail divided by the average noise surrounding it, the contrast being the difference between the average number of counts outside and within the detail. According to Rose (1973), the minimum SNR that the human eye can perceive ranges between 4 and 5. As a consequence, a more “practical” minimum for the dynamic range should be 4 or 5, while in the definition given above, the minimum possible value is 1. Moreover, if a detail has to be detected within a noisy background, the number of pixels forming the detail plays a major role: the higher this number (i.e., the greater the detail in the pixel scale), the higher the visibility. As a consequence, Yaffe proposed a modified definition of the dynamic range in which, before the ratio is evaluated, the maximum achievable signal is multiplied by the number of detail pixels and the overall noise level is multiplied by 4 or 5 according to the Rose criterion (Yaffe & Rowlands, 1997). This definition takes SNR requirements correctly into account but has less generality, as it depends on the specific imaging task.
4.1.8.3. Spatial Resolution

Before discussing spatial resolution in detail, it is worth clarifying the distinction between pixel aperture and spacing. Digital sensors are usually matrices of elements arranged in a regular array, and in many cases the entire surface of the single element is not sensitive to radiation (Fig. 4.1.10). The dimension of the active part of the element is called the aperture, while the dimension of the entire element, that is, the distance between the same point in two adjacent elements, is called spacing (for simplicity’s sake, assume here that the elements are square). The aperture defines the intrinsic spatial resolution of the system, while the spacing determines the sampling frequency. The ratio between the active portion and the overall detector surface is called the fill factor. For an efficient use of radiation, the fill factor should be as close as possible to unity. For a single detector device, it is equal to the ratio between the square of the aperture and the square of the spacing, provided that any further nonsensitive area along the outer edges (guard rings, front-end electronics, etc.) is not irradiated. If the overall detector is obtained by tiling several devices, the fill factor can be further reduced.

The active part of a detector pixel is effectively described by the point spread function (PSF), which expresses the relative sensitivity of each point within the active surface (Fig. 4.1.11). Ideally one could measure it by finely scanning the entire active surface with a pencil beam of extremely small cross section and registering the response corresponding to each position. As this solution is clearly impractical, many alternative possibilities have been proposed, all of which basically measure the line spread function (LSF). The LSF is the response of the pixel to a blade of radiation, uniform in one direction and infinitely narrow in the other, which ideally should be scanned over the pixel in one direction only, as shown in Figure 4.1.11. In other words, the LSF is the PSF integrated along one direction. As in most practical cases, square pixels are employed, an isotropic behavior is assumed with respect to the pixel center, and the LSF is considered sufficiently representative of the spatial resolution properties of the device.

A simple approach to LSF measurement consists of covering the pixel with an absorbing edge and recording the pixel response as the edge is scanned in a direction orthogonal to the edge, thus progressively exposing to radiation larger portions of the pixel. The curve obtained is the integral of the LSF along the scanning direction, called the edge response function (ERF), and the LSF is then obtained by numerically differentiating the ERF (Fig. 4.1.12). This technique has the advantage of measuring individually the LSFs of all pixels in a row, but has the drawback of being time consuming.

![Figure 4.1.10. Aperture (a) and spacing (s) in a digital detector (sensitive area in grey).](image)
A faster possibility, involving taking an image of an edge placed at a very small angle with respect to the pixel orientation, was proposed by Cunningham and Fenster (1987). An image profile taken along the orientation of the pixels, thus forming a small angle with the edge, provides the ERF (averaged over a number of pixels). The LSF is then obtained by numerical differentiation. A similar method, proposed by Fujita et al. (1992), employs a slightly tilted narrow slit instead of an edge, and makes possible a direct measurement of the LSF by combining the response of several rows of pixels taken in a direction (almost) orthogonal to the slit.

Figure 4.1.11. (A) PSF and (B) LSF of a detector pixel. In (A), the pixel is ideally subdivided into squarelets and the relative efficiency of each squarelet is given as a function of the squarelet position \((x, y)\). In (B), the pixel is ideally subdivided into stripes and the relative efficiency of each stripe is given as a function of position in the direction orthogonal to the stripes.

Figure 4.1.12. A PSF (solid line) and ERF (dashed line) for an ideal (left) and a real (right) 100 \(\mu\)m pixel. In the ideal case, every point in the pixel would have the same efficiency and the PSF would be a box function. In real cases, the relative efficiency often has its maximum in the pixel center and decreases near the edges.

A faster possibility, involving taking an image of an edge placed at a very small angle with respect to the pixel orientation, was proposed by Cunningham and Fenster (1987). An image profile taken along the orientation of the pixels, thus forming a small angle with the edge, provides the ERF (averaged over a number of pixels). The LSF is then obtained by numerical differentiation. A similar method, proposed by Fujita et al. (1992), employs a slightly tilted narrow slit instead of an edge, and makes possible a direct measurement of the LSF by combining the response of several rows of pixels taken in a direction (almost) orthogonal to the slit.
The modulus of the Fourier transform of the LSF provides the modulation transfer function (MTF), which is the most commonly employed quantity for expressing the spatial resolution of an imaging system. The MTF provides the relative amplitude with which each single spatial frequency is transferred by the system to the acquired image (Fig. 4.1.13). Actually, bar pattern test objects featuring progressively narrowing bar patterns of transmitting and absorbing materials make possible a direct determination of the MTF by measuring the decreasing amplitude of the patterns in the image as a function of the increasing frequency. However, this solution has two drawbacks: first, square rather than real sinusoidal functions are input into the system, and second, it allows the measurement of MTF values only at some predetermined spatial frequencies. Converging or diverging bar patterns may allow finer MTF sampling, but in this case a precise assessment of the spatial frequencies in which the MTF values are actually taken is important.

The MTF obtained with any of the previously mentioned techniques is actually the presampling MTF. As discussed earlier, in image acquisition, the sampling is determined by the pixel spacing. This means that the acquired image can be undersampled, as the presampling MTF may contain frequencies above the Nyquist limit \((2s)^{-1}\) if \(s\) is the spacing) imposed by the sampling. This results in the well-known phenomenon of aliasing. Without going into details (see, for instance, Bracewell, 1986), all frequencies higher than \((2s)^{-1}\) will not be reproduced in the image, and an overestimation of all Fourier coefficients may occur at frequencies lower than \((2s)^{-1}\). In order to avoid aliasing, the image should be “band limited” to frequencies below the Nyquist limit. In some cases, image blurring due to the focal spot size or light diffusion in the phosphor screen in indirect detection devices can help in fulfilling this requirement, although it cannot prevent aliasing effects from high-frequency noise. A safer solution may be to increase the sampling step rather than band-limiting the image, which can be achieved by taking several images while shifting the detector to different positions and then appropriately combining the acquired data (dithering). It was also demonstrated that under certain conditions, this technique can provide images with enhanced spatial resolution, determined by the scanning step rather than by the pixel size, provided that appropriate deconvolution algorithms are applied (Olivo et al., 2000).

Figure 4.1.13. The MTF of an ideal 100 μm pixel (solid line) and a real one (dashed line). The PSFs of the same pixels are shown in Figure 4.1.12.
4.1.8.4. Noise Performance

As mentioned previously, the intrinsic statistical nature of X-ray emission and interaction, described by the Poisson statistic, imposes a lower limit to the noise performance of any X-ray imaging system. The fact that only a photon counting device, in which all noise apart from the Poisson fluctuations is cut by a threshold, can reach this theoretical limit was also mentioned.

Although the issue of efficiency is discussed in the next subsection, it is worth stressing here that if \( \varepsilon \) is the quantum efficiency of a device, which means that \( N_1 = \varepsilon N \) photons will be detected out of the \( N \) quanta impinging on the detector, \( N_1 \) is still a Poisson-distributed variable; that is, the variance of \( N_1 \) is still equal to \( N_1 \) itself. Hence a limited efficiency does not prevent a counting system from reaching its theoretical limit in terms of noise performance.

In integrating systems, X-ray interaction is usually followed by one or more gain stages (e.g., the large number of visible photons created for each X-ray interaction in indirect detection techniques, or the electron multiplication in photomultiplier tubes). Following these gain stages, the new number of quanta will therefore be \( N_2 = GN_1 = G\varepsilon N \) (where, for simplicity's sake, all gain factors are condensed into the single factor \( G \)). This has relevant consequences. First, the variance of \( N_2 \) will be increased, as fluctuations on \( G \) have to be taken into account when the uncertainty propagation is carried out (Rabbani, Shaw, & Van Metter, 1987). Second, in general \( N_2 \) will not be Poisson distributed, even in those cases in which \( G \) is Poisson distributed.

In order to describe systems in which several stages of gain and loss of quanta are encountered, the method of quantum accounting diagrams (QADs) was devised (Cunningham, Westmore, & Fenster, 1994; however, the authors declare in their paper that the term QAD was previously used by M. J. Yaffe). The system is described as a sequence of cascaded stages in which quanta are transferred from one stage to another and gains or losses occur at each transfer. An example is shown in Figure 4.1.14. In order to describe an indirect detection system (for instance, a CCD coupled with a phosphor screen via fiber optics) with the QAD method, the following steps are needed:

![Figure 4.1.14. Example of a QAD for a CCD coupled to a phosphor via fiber optics.](image-url)
1. $N$ photons impinge on the phosphor and $N_1 = \varepsilon N$ interact (obviously $N_1 < N$: a first loss has occurred). $N_1$ is the “primary quantum sink” representing the intrinsic noise limit: the ultimate SNR will be lower (or equal, in ideal counting systems) than the square root of $N_1$.

2. Each interacting X-ray gives rise to a large number of visible photons, hence the number of quanta created in the phosphor is $N_2 = GN_1$, with $N_2 > N_1$ (in the given example, $N_2$ is also much larger than $N$, as $G$ can be on the order of $10^3$).

3. A percentage of visible photons escape and are not collected by the fiber optics; hence, a loss is encountered: $N_3 = l_1 N_2$, $N_3 < N_2$.

4. The non-100% efficient transfer of visible photons through the fiber optic results in another loss: $N_4 = l_2 N_3$, $N_4 < N_3$.

5. A further loss arises from the nonperfect matching between the spectra of the visible photons emitted by the phosphor and the sensitivity curve of the CCD: $N_5 = l_3 N_4$, $N_5 < N_4$.

If $N_2$, $N_3$, $N_4$, and $N_5$ are all larger than $N_1$, the limit in the noise performances of the system is still determined by the primary quantum sink—that is, the SNR will not be higher than $(N_1)^{1/2}$—and the smaller the fluctuations in each single step of the chain, the closer it will be to this limit.

On the other hand, if, for example, $N_5$ (which in the example is the smallest of the four) is smaller than $N_1$, this represents a secondary quantum sink in the system, and the overall SNR will be smaller or equal to the square root of $N_5$. In general, if the quanta are transferred through $K$ subsequent stages, the smallest number between $N_1$ and $N_K$ will determine the ultimate limit in the noise performances of the system.

It should be noted, however, that the QAD method outlined here, although extremely helpful, is an approximate one, and is subject to two major limitations: no additive noise sources are taken into account, and perfect localization is assumed at every creation/collection of secondary quanta (i.e., the effects of scattering and diffusion are not considered). As a consequence, its use may result in an understimation of the overall system noise.

In order to evaluate the effect that noise has on each single spatial frequency, the noise power spectrum (NPS) or Wiener spectrum has to be evaluated. If $I(x,y)$ is the data set, the NPS is defined as the Fourier transform of its autocorrelation function (Blackman & Tukey, 1958). However, using the properties of the Fourier transform, it can be demonstrated that the same result can be obtained by directly calculating the square modulus of the Fourier transform of the data itself (see, for instance, Williams, Mangiafico, & Simoni, 1999). This allows an easier approach to NPS calculation, and in fact is referred to as the “direct method” of NPS calculation, the autocorrelation-based technique being called the “indirect method.” In fact, Dainty and Shaw (1974) define the NPS (for a continuous variable) as

$$\text{NPS}(u,v) = \lim_{X,Y \to 0} \frac{1}{4XY} \left( \int_{-X}^{X} \int_{-Y}^{Y} (I(x,y) - \bar{I}) e^{-2\pi i (ux + vy)} \, dx \, dy \right)^2. \quad (4.1.9)$$

where $I(x,y)$ is the two-dimensional image intensity and $\bar{I}$ is the average background intensity. The angled brackets indicate that averaging over data ensembles is required. In order to apply the definition to digital detectors, discrete Fourier transform algorithms are employed in which the integrals are replaced with sums (see, for instance, Brigham, 1974). A comprehensive treatment can be found in the third chapter (by Dobbins) of the Handbook of Medical Imaging edited by Beutel, Kundel, and Van Metter (2000).
It should be noted, however, that many nontrivial practical problems are encountered when the NPS has to be measured experimentally. In general, the necessity of using a finite range of noise data affects the results at high frequencies and determines the frequency sampling of the NPS itself, and a large number of measurements has to be averaged in order to reduce the fluctuations. Some of these practical problems are discussed by Dobbins et al. (1995), while a somewhat “classical” theoretical description and experimental method can be found in Giger, Doi, and Metz (1984).

As already discussed in the case of the MTF, when square pixels are used, an isotropic behavior is assumed and a one-dimensional NPS is considered sufficient to describe the noise performance of the system. This is usually achieved by selecting a slice of the two-dimensional NPS close to the υ or ν axis, where υ and ν are the spatial frequency axes conjugated to the spatial coordinates x and y, respectively.

4.1.8.5. Detection Efficiency

The first step in all X-ray imaging detectors is X-ray interaction, in the sensor material for direct detection systems or in the phosphor screen for indirect systems. The probability of interaction for the single X-ray, or quantum efficiency (ε), is thus determined by the thickness (t) and by the attenuation coefficient (μ) of the material in which the interaction takes place:

\[ \varepsilon = 1 - e^{-\mu(E)t} \quad (4.1.10) \]

where the energy dependence of the attenuation coefficient on the X-ray energy (E) is explicitly expressed. Equation 4.1.10 holds only when monochromatic radiation is used, which is true only in a limited number of cases. When a polychromatic spectrum is used, equation 4.1.10 should be replaced with

\[ \varepsilon' = \frac{\int_{E_{\text{min}}}^{E_{\text{max}}} \Phi(E)(1 - e^{-\mu(E)t}) \, dE}{\int_{0}^{E_{\text{max}}} \Phi(E) \, dE} \quad (4.1.11) \]

where \( \Phi(E) \) is the X-ray spectrum impinging on the detector (i.e., beyond the imaged sample), which means that beam-hardening effects due to the sample must be taken into account. In noncounting detectors, since the detected signal actually depends on the energy absorbed by the detector rather than on the number of photons, the concept of energy absorption efficiency (\( \varepsilon'' \)) is introduced in some cases:

\[ \varepsilon'' = \frac{\int_{0}^{E_{\text{max}}} \Phi(E) \frac{\mu_{\text{en}}(E)}{\mu(E)} (1 - e^{-\mu(E)t}) \, dE}{\int_{0}^{E_{\text{max}}} \Phi(E)E \, dE} \quad (4.1.12) \]

where the ratio between the energy absorption coefficient \( \mu_{\text{en}} \) and the attenuation coefficient accounts for the amount of absorbed energy per interacting X-ray photon.

In counting devices, the concepts of quantum efficiency and MTF are sufficient to describe the system performance at all spatial frequencies, while for integrating detectors, the concept of detective
quantum efficiency (DQE) has to be introduced. Before discussing the DQE, it is worth introducing the concept of noise equivalent quanta (NEQ).

As in a counting device, the only noise source is X-ray Poisson fluctuations, if \( N \) is the number of interacting X-rays, \( \sigma_N = N^{1/2} \), and \( \text{SNR}_{\text{ideal}} = N/\sigma_N = N^{1/2} \). In an integrating device, the recorded signal \( S \) is different from the number of interacting quanta, and thus in general \( \text{SNR}_{\text{nonideal}} = S/\sigma_S < N^{1/2} \). As in the ideal (counting) behavior one has \( (\text{SNR}_{\text{ideal}})^2 = N \), in the nonideal behavior it is useful to define the quantity \( N' = (S/\sigma_S)^2 < N \), which establishes a direct comparison between the nonideal case and the ideal case. The quantity \( N' = (\text{SNR}_{\text{nonideal}})^2 \) is called the NEQ.

The DQE is defined as the ratio between the square of the SNR in the image and the square of the SNR input to the detector:

\[
DQE = \left( \frac{\text{SNR}_{\text{out}}}{\text{SNR}_{\text{in}}} \right)^2,
\]

where the numerator is, in practice, the NEQ. (In fact, using the notation in equation 4.1.13, it is possible to write \( DQE = N'/N \).) Although this definition might seem rather straightforward, it is important to notice that while \( (\text{SNR}_{\text{out}})^2 \) can be measured on the acquired image, the determination of \( (\text{SNR}_{\text{in}})^2 \) is not easy. If the beam is monochromatic, this corresponds to the number of photons impinging on the detector, but, as mentioned previously, this is the case only in a restricted number of cases. Strictly speaking, the DQE is referred to a number of photons; hence, when a polychromatic beam is used, the variances of the number of quanta should be individually evaluated at each energy (in practice, in each energy bin) and then summed to give the overall variance. According to some authors, however, since the effective signal in the detector depends on the deposited energy, emphasis should be placed on the energy variance rather than on the variance in the number of counts, and an energy-weighted balance should thus be used to estimate \( (\text{SNR}_{\text{in}})^2 \).

Although thus far the DQE has been treated as a scalar quantity (for simplicity’s sake), its frequency dependence should already be clear as the MTF, which accounts for the decrease in the achievable image contrast with increasing spatial frequencies, as already discussed. Hence the DQE is actually a \( DQE(\nu) \), where \( \nu \) is the spatial frequency. (Again, although the DQE is actually a function of two independent variables, it is often approximated with a single-variable dependent function when the system is made of square pixels and an almost isotropic behavior is expected.)

Different ways of practically evaluating the NEQ (e.g., \( (\text{SNR}_{\text{out}})^2 \)) are reported in the literature. Here, referring to the previously quoted work of Dobbins et al. (1995), in which a “normalized” noise power spectrum (NNPS) is evaluated as

\[
\text{NNPS}(\nu) = \frac{\text{NPS}(\nu)}{(\text{large area signal})^2},
\]

the NEQ or \( (\text{SNR}_{\text{out}})^2 \) as a function of spatial frequency is given by

\[
\text{NEQ}(\nu) = \left( \frac{\text{SNR}_{\text{out}}(\nu)}{\text{SNR}_{\text{in}}(\nu)} \right)^2 = \left( \frac{\text{MTF}(\nu)}{\text{NNPS}(\nu)} \right)^2.
\]
4.2. SENSORS FOR DOSIMETRY

4.2.1. SPECIAL REQUIREMENTS FOR DOSIMETRIC MEASUREMENTS

4.2.1.1. Characteristics of an “Ideal” Dosimeter

Dosimeters make possible the measurement of a delivered dose via a dose-related effect such as the temperature increase in a calorimeter, the current in an ionization chamber, or the optical density of a radiographic film. The required general characteristics for dosimetry detectors include the following:

4.2.1.1.1. Linear Response

The output of the detector (current, analog-to-digital converter counts, etc.) should increase linearly with the absorbed dose and have no zero offset. Clearly this would make calibration as simple as possible. Ideally the range of linearity should be infinite, but saturation and other nonlinear behavior will occur in practice.

To a certain extent, and excluding the saturation regime, calibration can be performed in cases of nonlinear response. Figure 4.2.1 shows the response characteristics of different dosimetric systems.

4.2.1.1.2. Tissue Equivalence

In order to reproduce correctly the mechanisms of energy release in tissues, the atomic number of an ideal dosimeter must be as close as possible to that of the tissue ($Z_{\text{tissue}} = 7.4$). This implies a flat energy response across the range of use of the dosimeter.

Tissue equivalence is a major issue when a depth dose distribution or a three-dimensional (3-D) dose distribution has to be measured (for instance, for validating a treatment planning system). A change in the spectrum, typically occurring for photon beams, causes measurements made with a nontissue equivalent dosimeter to be unreliable. Tissue equivalence is even more relevant when performing neutron dosimetry, as the cross section for neutron interactions has very sharp changes, such as resonance peaks in elastic scattering.

On the other hand, tissue equivalence is a less stringent requirement in electron dosimetry because of the slow variation of the mass stopping power as a function of the beam energy (see section 4.2.1.2.1). Moreover, it is not important when measuring dose in one position because a calibration as a function of the beam energy against a tissue equivalent dosimeter can be performed. Most commercial devices are provided with calibration data for the most typical range of use.

Figure 4.2.1. Examples of typical responses of dosimetric systems: (a) linear response; (b) sublinear response (the dotted line shows the linear behavior); (c) limited range of linearity, with a region of underexposure and a saturation region.
4.2.1.3. Independence of Dose Rate
An integrating dosimeter—that is, a device measuring the total dose deposited within a certain time interval—should not depend on the dose rate. In other words, it should measure the same dose independently of the time over which the dose has been delivered.

4.2.1.4. Independence of Ambient Parameters
The response of a dosimeter should not be affected by ambient parameters such as temperature and humidity, but few dosimeters satisfy this condition completely. In order to limit the effects of ambient parameters, storage and operating conditions must be controlled, otherwise a calibration as a function of these parameters must be carried out.

4.2.1.5. Independence of Beam Direction
An ideal dosimeter should always give the same output, independent of the direction of incidence of the beam, but this does not usually happen because of the geometry and the material of the dosimeter. This is a major issue when performing depth dose measurements because the angular distribution of radiation on the detector depends on scattered radiation, which is in turn dependent on the position in the material. Typical depth dosimeters, such as silicon (Si) diodes, are calibrated for the same geometry as that in which they are used.

4.2.1.6. Long-Term Stability of Calibration
It is clearly more convenient to use a dosimeter with a constant response over time, however, this is not always achieved. For instance, whereas the calibration of ion chambers is constant over time, other dosimeters, such as thermoluminescence (TL), devices show a change in sensitivity depending on their radiation and readout histories, and hence need periodic recalibration.

4.2.1.6. Precision and Accuracy
Precision is the spread in the distribution of repeated measurements and describes random uncertainties. To estimate the actual difference between the measured value and the “true” value, systematic uncertainties resulting from several factors, such as uncertainties in the calibration or in the estimation of the physical constants, also need to be taken into account. Accuracy describes the systematic uncertainties and is a less objective quantity, depending on the judgment of the experimenter. Both inaccuracy and imprecision need to be as small as possible.

Other desirable characteristics for a dosimeter are related to the specific application and will be discussed case by case. Some characteristics that can be desirable depending on the type of application include the following:

- Long-term information storage
- Possibility of 2-D and 3-D dosimetry
- Reusability
- Portability
- No need for high voltage.
4.2.1.2. Specific Requirements for Dosimetry of Different Beams

4.2.1.2.1. Electron Dosimetry

Electrons are directly ionizing particles. At relatively low energy, they only produce other electrons, while at higher energies they start losing energy by the bremsstrahlung process. For instance, the radiation yield (i.e., the fraction of kinetic energy of the primary electron converted into bremsstrahlung) for electron beams in soft tissue is about 0.3% at 1 MeV and exceeds 10% only above 25 MeV (Berger, Coursey, & Zucker, 2000).

Electrons lose energy very quickly when passing through a material, so charged particle equilibrium (see section 4.2.3.2) is very seldom reached. In order to limit this effect as much as possible, dosimeters for electrons are usually designed to be very thin in the direction orthogonal to the incoming beam.

The tissue dose for an electron beam can be determined from the dose deposited in another material as follows:

\[
D_{\text{tissue}} = D_m s_{\text{tissue}, m}, \tag{4.2.1}
\]

where \(D_{\text{tissue}}\) is the tissue dose, \(D_m\) is the dosimetric material dose, and \(s_{\text{tissue}, m}\) is the stopping power ratio of water to that of the dosimetric material.

Tissue equivalence is not as great an issue as it is in photon and neutron dosimetry because the ratio of the mass stopping power of two materials is a slowly varying function of the electron energy, and no sharp peaks, as for resonances in neutron interactions or for K-edges in photon interactions, are present. For example, Figure 4.2.2 shows the ratio of the collision mass stopping power of electrons in air and in Si to that in water.

4.2.1.2.2. Photon Dosimetry

Photons are indirectly ionizing particles, that is, they first interact with matter to produce charged particles (electrons and, at energies above 1022 keV, positrons). These charged particles produce ionization along their tracks.

![Diagram of mass stopping power ratio for electrons in different materials](image)

Figure 4.2.2. Collision mass stopping power ratio for electrons for two dosimetric materials compared to water (data from Berger, Coursey, & Zucker, 2000).
As previously mentioned, a good material for photon dosimetry must have an atomic number as close to that of the tissue as possible. This implies that the ratio of the interaction coefficients of the material to those of the tissue is a slowly varying function of the photon energy. In this way, the conversion from the material dose to the tissue dose can easily be performed, even when dealing with polychromatic beams (i.e., in the most common case), by means of equation 4.2.2:

$$D_{\text{tissue}} = D_m \frac{\langle \mu_{en}/\rho \rangle_{\text{tissue}}}{\langle \mu_{en}/\rho \rangle_m},$$

(4.2.2)

where $D_{\text{tissue}}$ is the tissue dose, $D_m$ is the detector material dose, and $\langle \mu_{en}/\rho \rangle_{\text{tissue}}$ and $\langle \mu_{en}/\rho \rangle_m$ are the average mass energy absorption coefficients of the beam for the tissue and dosimetric material, respectively.

Figure 4.2.3 shows a comparison of the ratio of the mass energy absorption coefficients for photons for two typical dosimetric materials to that of water.

Unlike in electron dosimetry, charged particle equilibrium can be achieved relatively easily in photon dosimetry by appropriately selecting the dimensions of the sensitive volume or the dimension and the composition of the outer walls of the instrument.

4.2.1.2.3. Neutron Dosimetry

As previously pointed out in section 4.2.1.1, tissue equivalence is even more relevant for neutron dosimeters than for photon dosimeters, as the interaction probabilities for neutrons do not have a smooth dependence on energy or atomic number of the material. Hence a dosimetric material for neutrons must have an atomic composition as close as possible to the tissue. In particular, the hydrogen content must be the same, as the contribution of hydrogen to delivered dose in soft tissue is predominant, although the weight fraction of the hydrogen content is only around 10% (Greening, 1985). However, when greater sensitivity is needed, neutron dosimeters are constructed with materials that have cross sections capable of neutron reactions, such as boron or lithium.

![Figure 4.2.3. Mass energy absorption coefficient ratio for two different dosimetric materials compared to water (data from Nowotny, 1998).](image-url)
4.2.2. THE CALIBRATION CHAIN

An absolute dosimeter is a detector that provides a signal from which the dose (or a related quantity such as kerma or exposure) can be determined based on physical constants derived independently. The only devices that can be regarded as absolute dosimeters are calorimeters, ionization chambers, and Fricke dosimeters.

A relative dosimeter needs to be calibrated against an absolute instrument or in a known radiation field. Examples of the output of a relative dosimeter are the optical density of a film dosimeter and the light yield of the readout system for a TL dosimeter.

The main steps for the correct calibration of a dosimeter are shown in Figure 4.2.4.

The key elements of the calibration chain are the primary standard dosimeters. These instruments, ion chambers, or calorimeters, depending on the national protocol, are permanently located in the national standards laboratories. Secondary standards, maintained by regional calibration laboratories, are transfer instruments used for calibrating local hospital standards. A calibrated local standard can then be used for calibrating beams or other local dosimeters. A local standard must be recalibrated periodically against a regional standard. The overall accuracy of the final calibration is thus dependent on the accuracy of the intermediate steps.

4.2.3. IONIZATION CHAMBERS

4.2.3.1. General Principles

The ionization chamber is the most typically used instrument for beam monitoring and calibration. It can be used as an absolute dose monitor, provided its geometry is known with high precision.

The working principle of the ionization chamber is based on detection of the ionization created in a gas by directly or indirectly ionizing particles. The number of electron-ion pairs indirectly created by a particle is equal to the energy deposited divided by the average energy needed to create an electron-ion pair. The average energies required for electrons to create an electron-ion pair ($W$-value) in different gases are shown in Table 4.2.1.

![Figure 4.2.4. Schematic diagram of a dosimetry calibration chain (update of Planskoy, 1983).](image-url)
If an external voltage is applied to the volume in which the ionization takes place, the electrons will drift toward the positive electrode and the positive ions will drift toward the negative electrode. Thus a current can be measured in an external circuit by means of an electrometer.

Air-filled ionization chambers are particularly suitable for measuring exposure, the definition of exposure being related to the charge created in air. The dose in air can then be calculated from exposure, as discussed in section 4.1.

The main issues related to the use of an ion chamber for dose-in-air measurements are discussed in sections 4.2.3.2 through 4.2.3.5.

4.2.3.2. Charged Particle Equilibrium

Exposure at a point is defined by the total charge of one sign created when all the electrons liberated by photons in air at that point are completely stopped, thus the tracks of all secondary electrons created at that point should be followed. This condition is equivalent to the condition of charged particle equilibrium, as explained in Figure 4.2.5. Let A be the volume of interest, and consider a volume B of air surrounding A large enough to completely stop any electron produced in A. Provided that every point in B is subject to the same exposure (i.e., that the beam attenuation in B is negligible), the ionization produced in A by electrons produced by photons interacting in B (track marked “b” in the figure) is equal to the number of secondary electrons (a) produced in A and stopped in B.

Table 4.2.1. $W$-values for electrons in different gases

<table>
<thead>
<tr>
<th>Gas</th>
<th>$W$-value (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air</td>
<td>33.8</td>
</tr>
<tr>
<td>$N_2$</td>
<td>34.8</td>
</tr>
<tr>
<td>Ar</td>
<td>26.4</td>
</tr>
<tr>
<td>$O_2$</td>
<td>30.8</td>
</tr>
<tr>
<td>$CH_4$</td>
<td>27.3</td>
</tr>
</tbody>
</table>

exactly equal to the ionization produced in B by electrons produced by photons interacting in A (track marked “a” in the figure).

In order to achieve charged particle equilibrium, the sensitive volume of an air-filled ionization chamber must be surrounded by a large enough volume of air. Charged particle equilibrium is more easily reached for low-energy beams, such as diagnostic beams, because of the shorter range of low-energy electrons. More complex geometries must be adopted for reaching charged particle equilibrium for higher energy beams, as discussed in section 4.2.3.4.

4.2.3.3. Saturation Voltage

A major issue related to ion chambers is the selection of the correct voltage depending on the beam energy and the beam flux. A qualitative example of the voltage-current characteristic is shown in Figure 4.2.6 for two different energy fluxes of the incoming beam, $\Psi_1$ and $\Psi_2 > \Psi_1$.

At low voltages the electron-ion pairs have a high probability of recombination before reaching the electrodes and thus are not completely collected. Upon increasing the voltage, the probability of recombination decreases until all electron-ion couples have sufficient acceleration to reach the electrodes without recombining. The ionization current then reaches a plateau. The minimum value of the voltage at which a plateau is reached is called the saturation voltage. It is intuitive that the higher the energy flux of the incoming beam, the higher the number of electron-ion pairs produced per unit volume, and thus the higher the probability of recombination; therefore the saturation voltage is higher for more intense or more energetic beams. Saturation voltage also depends on the chamber geometry and on the gas used. More details about recombination can be found in Boutillon (1998).

4.2.3.4. Geometries for Ion Chambers

4.2.3.4.1. The Free Air Chamber

The simplest ion chamber is called a free air ionization chamber, shown in Figure 4.2.7.

Here, the cross section of the sensitive volume is defined by means of an entrance collimator, while its length corresponds to the length of the volume delimited by the guard electrodes.

Since exposure is defined as the charge deposited per unit mass in air, the exposure rate for a photon beam, provided the geometry of the system is known, can be calculated as follows:

![Figure 4.2.6. Voltage-current characteristic of an ion chamber for two different incoming beam energy fluxes.](image)
where \( I \) is the ionization current measured by an electrometer, \( \rho \) is the density of air, and \( V \) is the sensitive volume of the chamber.

Since the air density depends on both temperature and pressure, equation 4.2.3 can be written more precisely as

\[
\dot{X} = \frac{I}{m} = \frac{I}{\rho_{air} V},
\]

(4.2.3)

where \( I \) is the ionization current measured by an electrometer, \( \rho \) is the density of air, and \( V \) is the sensitive volume of the chamber.

Since the air density depends on both temperature and pressure, equation 4.2.3 can be written more precisely as

\[
\dot{X} = \frac{I}{\rho_{air} V} \frac{\rho_0}{p_0} \frac{T}{T_0},
\]

(4.2.4)

where \( \rho \) and \( T \) are the pressure and the absolute temperature of the gas; \( T_0 = 273.15 \) K, the standard temperature; \( p_0 = 760 \) mmHg, the standard pressure; and \( \rho_{air} = 1.293 \) kg m\(^{-3}\), the density of air under the standard conditions \( p_0 \) and \( T_0 \). The dose rate can then be derived from the exposure rate according to equation 4.1.7.

However, unless the geometry of the system (thickness of the sensitive volume, collimator size, and beam divergence) is very well known, calibration with an absolute instrument is recommended rather than the use of equation 4.2.4. Corrections for temperature and pressure are always needed.

The free air chamber is used only in calibration laboratories as a calibration standard for X-rays generated at less than 300 kV; other geometries are more convenient for clinical practice.

4.2.3.4.2. The Cavity Chamber

In practice it is not always possible to achieve conditions of charged particle equilibrium in air, and in particular, when high-energy beams (\( E > 300 \) keV) have to be monitored. For example, the range for 200 keV electrons in air is 0.4 m (Berger, Coursey, & Zucker, 2000). However, the same condition can be achieved by “compressing” the outer volume B (needed to ensure charged particle equilibrium) into a thinner, denser volume by replacing it with a proper thickness of a solid material with an effective

---

**Figure 4.2.7.** Schematic of a free air ion chamber. The sensitive volume is defined by the entrance collimator, the beam divergence, and the dimensions of the collecting electrode. Absolute dosimetry is only possible when all these quantities are known with high precision (Knoll, 2000).
atomic number as close as possible to that of air but with a higher density (see Fig. 4.2.8). Carbon, for example, is a good candidate, and is used for the most common ion chambers in clinical use.

Because it is difficult to know with high precision the dimensions of the air volume and, above all, to what extent the outer wall material can be regarded as air equivalent, a cavity chamber needs to be calibrated against a free air chamber.

Theoretical issues regarding cavity chambers have been widely discussed and more details can be found in Greening (1985).

4.2.3.4.3. The Parallel Plate (Plane Parallel) Ion Chamber

The parallel plate ion chamber is, together with the Farmer chamber (see section 4.2.3.4.4), the most frequently used ion chamber in clinical practice. A typical application of this type of chamber is in the calibration of low-energy photon beams and electron beams of less than 10 MeV. According to the International Atomic Energy Agency (IAEA, 1997), they must actually be used below 5 MeV.

In the parallel plate ion chamber (sketched in Figure 4.2.9), the beam direction is orthogonal to the electrodes. One of these electrodes also acts as an entrance window and typically consists of a tissue.

**Figure 4.2.8.** Principle of a cavity chamber: A is the volume of interest, B is the air volume needed for charged particle equilibrium, and \( B' \) is a volume with an atomic number close to that of air, but with a higher density.

**Figure 4.2.9.** Parallel plate ion chamber (IAEA, 1997).
equivalent polymer (Kapton, Mylar) coated with a thin layer of conductive material. Typical materials for the coating are aluminum or graphite, although the latter is preferable due to its tissue equivalence. The back side of the chamber is either a conducting plastic material or a graphite-coated insulating material. A guard ring is also present.

Several commercial parallel plate chambers are reversible and equipped with two different entrance windows designed for high-energy and low-energy applications. Because of their thin collecting volume, some parallel plate chambers are suitable for depth dose measurements with good depth resolution.

4.2.3.4.4. The Cylindrical Ion Chamber (Thimble Chamber/Farmer-Type Chamber)

This geometry, shown in Figure 4.2.10, is typically adopted for beam calibration in radiotherapy. The chamber consists of a hollow cylinder made of low atomic number material (typically graphite) acting as an electrode and an aluminum central electrode about 1 mm in diameter. The presence of a strong electric field near the central electrode increases the charge collection efficiency, thus allowing the use of a lower voltage than in planar geometry.

Cylindrical ion chambers can be manufactured in very small sizes, with active volumes down to about 0.01 cm$^3$. Small cylindrical chambers are suitable for use in radiation fields with strong gradients in one direction or for in vivo measurements in brachytherapy, although other devices, such as Si diodes, are more frequently used for such applications.

Usually, cylindrical chambers are used in the calibration of medium energy X-ray beams generated at greater than 80 kV, high-energy photon beams, and electron beams above 10 MeV. They are also suitable for proton dosimetry (IAEA, 2000).

4.2.3.4.5. Sealed Ion Chambers

In a sealed ion chamber, the gas pressure is controlled and no correction for air density changes is required. Moreover, it can be filled with a gas other than air or at a higher pressure, thus ensuring a higher sensitivity for low dose rate applications. Typical filling gases are argon, xenon, and nitrogen. Nitrogen is used for removing any dependence on air humidity rather than for increasing the sensitivity of the chamber, because the energy required for producing an electron-ion pair is very close to that required in air. High-pressure chambers are usually used as radiation survey meters.

![Figure 4.2.10. Cylindrical ion chamber (Aird & Farmer, 1972).](image-url)
4.2.3.4.6. Brachytherapy Chamber
When a low dose rate brachytherapy source has to be calibrated, maximum sensitivity must be
achieved, thus a well-shaped geometry for this operation is necessary, as shown in Figure 4.2.11. An
appropriate active volume is about 1 L. The chamber is usually sealed, and therefore does not require
any correction for air density. The source can be placed in a holder within the cavity, thus ensuring an
approximately $4\pi$ geometry.

4.2.3.4.7. The Compensated Ion Chamber
This type of chamber is designed for neutron dosimetry and consists of two separated ion chambers,
one neutron-sensitive, being coated with boron or with another material with a large cross section for
neutron interactions, while the other is sensitive to photons. If the polarities are arranged so that the
currents in the two chambers have opposite directions, the reading obtained from the electrometer
indicates the difference between these two currents, thus allowing the contribution of photons to be
separated from that of neutrons.

4.2.3.5. Advantages and Disadvantages
Because ion chambers can have many different structures and applications, it is not easy to summarize
their general characteristics. It must be pointed out that for precise beam calibration, several correc-
tions must be made, including for radiation field perturbation caused by the chamber itself, the effects
of ion recombination, chamber polarity effects, and other factors. Several protocols for the evaluation
of such corrections have been proposed and are currently applied in clinical practice (see, for example,
Almond et al., 1999; IAEA, 2000).

Some of the main advantages of ion chambers, which make them useful as one of the standards
for beam calibration, are the following:

- They can be designed to be tissue equivalent.
- They are nearly impervious to radiation damage.

![Diagram of a brachytherapy chamber](image-url)
• They have long-term calibration stability.
• Although each type of chamber is designed for a specific application, the range of applications can be extended by adding suitable buildup layers or buildup caps.
• Provided their geometric characteristics are known with high precision, some types of ion chambers are suitable for absolute measurements of dose.
• They can be made sensitive to neutrons by adding a boron or lithium compound to the entrance window.

On the other hand, they do have some disadvantages:
• Because of the need for high voltages, the portability of ion chambers is limited.
• Apart from small volume cylindrical chambers, they are not usually suitable for in vivo measurements because of their large volume.

4.2.4. LUMINESCENCE DOSIMETRY

Luminescence dosimetry is a well-established solid-state method based on the use of materials with a small number of impurities. This principle is explained in Figure 4.2.12.

The presence of an impurity introduces traps for electrons at intermediate energies between the valence band and the conduction band. When the material is irradiated, the following occurs:

• A number of electrons is created.
• Each electron can excite a number of electrons from the valence band to the conduction band.
• Each excited electron migrates through the conduction band until it falls into a trap.

During the readout process, if the trap is deep enough to avoid simple thermal excitation at room temperature, the trapped electron can be excited, by means of light or heat, to jump into the conduction band, from where it can either fall into a trap or fall into the valence band, thus emitting light. The light yield, which can be measured using proper instrumentation, is proportional to the number of excited photons and, ideally, to the dose deposited in the material.

Figure 4.2.12. The thermoluminescence or optically stimulated luminescence process. During irradiation, an electron (1) is excited into the conduction band, (2) migrates through it, and (3) falls into a trap. During readout, the electron (1) is excited into the conduction band by means of light or heat, (2) migrates through the conduction band, and (3) falls into the valence band emitting light.
Two dosimetric techniques are based on this effect: TL dosimetry and optically stimulated luminescence (OSL) dosimetry.

4.2.4.1. Thermoluminescence Dosimetry
Thermoluminescence dosimeters, in particular lithium fluoride (LiF) dosimeters, are among the most commonly used dosimeters in clinical practice.

4.2.4.1.1. General Principles
As already mentioned, TL dosimetry is based on the excitation of trapped electrons by means of heat and on the detection of the light emitted during de-excitation. The microdosimetric mechanisms of TL dosimetry are not fully understood and produce a nonlinear response; however, the physical aspects involved are widely discussed in the literature (see, e.g., Cameron, Sunthralingham, & Kenney, 1968; Horowitz, 1981; and McKinlay, 1981).

4.2.4.1.2. Characteristics of TL Dosimeters
A TL dosimetry system consists of the following:

- A set of TL dosimeters, available in different forms (powder, pellet, chip, rod).
- A readout system consisting of a heated element where the TL dosimeters are placed, ensuring uniform thermal contact, and a photomultiplier tube that detects the light emitted during the heating process (see Fig. 4.2.13).
- A high temperature oven for heating the TL dosimeters at the end of the readout process in order to free any electrons still trapped in a metastable state. This process is called annealing. This operation is not necessary for all dosimeters, but it is commonly used in practice to ensure stability of the readout.

Although LiF dosimeters are most commonly used because of their near tissue equivalence ($Z = 8.14$) and wide range of use ($10^{-5}$ to $10^3$ Gy) (Johns & Cunningham, 1983), many TL materials and combinations of dopants have been evaluated in order to find the most suitable one for each application.

![Figure 4.2.13. Schematic diagram of a TL dosimeter reader.](image)
In general, highly sensitive materials such as LiF in combination with manganese (Mg), copper (Cu), or phosphorus (P), or aluminum oxide (Al₂O₃) are recommended for low-dose dosimetry, for example, for diagnostic beams, personal, and even environmental dosimetry (Duggan et al., 1999; Saez-Vergara et al., 1999). Moreover, TL compounds of boron or lithium (having a large cross section for neutron capture) are particularly suited to neutron dosimetry (Horowitz, 1981; Toivonen et al., 1998).

The readout process. After irradiation, TL dosimeters are placed in a reader and the temperature is increased at a steady rate. During the readout process, a nitrogen flow is used in some cases to prevent chemoluminescence emissions caused by reactions with oxygen or water vapor, which may alter the readout.

A plot of the light yield as a function of temperature (the glow curve) shows several peaks at fixed values of temperature (see Fig. 4.2.14).

These peaks are caused by the emptying of the different traps in the forbidden band and the subsequent emission of light when the electrons fall back into the valence band: the smaller the energy gap between the trap and the conduction band, the lower the temperature at which the peak will appear. The lower temperature peaks are usually rejected, as they can be more easily affected by fluctuations due to thermal excitation at room temperature.

Typical TL dosimeter readers allow the user to choose between different settings related to different readout system gains, offering either high sensitivity or wide range.

A major issue affecting TL dosimetry is fading, that is, the loss of information due to light excitation, thermal excitation at room temperature, or other imperfectly understood mechanisms. The fading phenomenon varies strongly from one material to another and depends on both the readout and the storage conditions (light, temperature, humidity). For some TL materials it is negligible, while for others materials it is extremely relevant (Horowitz, 1981).

Unless the fading of the selected TL material is negligible over long periods, when a precise comparison of measurements has to be performed, it is recommended that the readout be observed after the same time interval in all cases. Moreover, the dosimeters must be kept under the same conditions of temperature, light, and humidity before the readout.

**Calibration.** The response of TL dosimeters within the same set may vary because of small differences in the weights of the chips, small variations in their composition, or their individual radiation

![Figure 4.2.14. Schematic of a glow curve for a TL dosimeter material. It is usually good practice to reject the low temperature peaks and to integrate only the area under the higher temperature peaks, as highlighted in the graph.](image-url)
histories. Thus, ideally, individual calibrations must be performed periodically, especially when relative dose measurements (e.g., depth dose distribution) have to be carried out. Alternatively, TL dosimeters can be grouped in batches according to their sensitivities.

Advantages and disadvantages. The main advantage of TL dosimeters is their small size. This provides a high degree of portability, allowing for off-site measurements and in vivo dosimetry. It also implies a small perturbation of the radiation field, thus on-patient measurements or depth dose measurements in tissue equivalent materials become feasible.

For these reasons, and also because of the wide range of available TL materials, allowing the selection of the most suitable one for each application, TL dosimetry is one of the most used dosimetry methods. However, other issues must be taken into account in order to avoid imprecise and misleading results:

- First, and most obvious, TL dosimeters must be handled with care, avoiding scratches that may alter their weight and dust that may mask the light emitted during the readout process. Moreover, a TL dosimeter can be damaged and lose sensitivity if exposed to either a very high dose during the irradiation process or excessive temperatures during the readout or annealing process. A recalibration is recommended every time any of these conditions occurs.
- For similar reasons, TL dosimeters are not recommended for use in water phantoms unless they are properly protected.
- Nonlinear behavior occurs above a certain energy (typically a few Gy). Typical behavior consists (Horowitz, 1981) of a region of supralinearity up to about 100 Gy, followed by a decrease of the readout–dose ratio until reaching a sublinear regime at dose levels of $10^3$ to $10^6$ Gy, as shown in Figure 4.2.15. Because the width of the supralinearity region depends on both the beam energy and the particles studied, apart from the TL material used and the composition and density of dopants, a calibration over the entire energy range of interest and for all the beam types involved must be performed.
- Some TL materials are not tissue equivalent. This may give rise to the problems previously discussed in section 4.2.1. Figure 4.2.16 shows a comparison of the photon mass energy absorption coefficient for several TL materials.

**Figure 4.2.15.** Generic diagram of the response of a TL dosimeter.
Although TL dosimeters are suitable for depth dose measurements, they are in general not very suitable for use in the presence of a strong radiation field gradient because their response depends on the depth at which the dose is deposited. Thus, unless they are uniformly irradiated, the result of a measurement may be misleading. Uniform irradiation in the presence of a strong radiation field gradient is possible only with small-size dosimeters, but a very small-size dosimeter is less reliable because of the large uncertainties that can affect the measurement. Better performances are obtained in this case with TL dosimeters in powder form.

Reproducibility of the heating cycle is mandatory for good reproducibility of the results. This is normally achieved with modern readout systems.

4.2.4.2. Optically Stimulated Luminescence Dosimetry

The use of OSL dosimeters has developed only in recent years and is not yet competitive with TL dosimetry. However, its main advantage is that the processes involved are purely optical, so the readout system is simple compared with that for TL dosimetry, and the number of control parameters is reduced.

As the physical mechanisms involved are related, most of the issues discussed for TL dosimetry also apply to OSL dosimetry, including the relevance of fading and the necessity for both an average and individual calibration. Moreover, OSL dosimeters and TL dosimeters share ease of portability and convenience of use.

An OSL dosimetry readout system consists of a laser for stimulating the dosimeter and a photomultiplier tube for detecting the light emitted. A similar process to annealing can be useful in depleting the traps completely after the readout, although it is less important for a good result than is the annealing of TL dosimeters.

The dosimeter can be equipped with optical fibers for channeling both the stimulating and emitted light. In this case, “near-real-time” measurements are possible. Currently the most commonly used OSL material is Al₂O₃:C (Aznar et al., 2004; Bøtter-Jensen et al., 1997; Gaza, McKeever, & Akselrod, 2005), which has TL properties, but its OSL properties are more appealing for dosimetry. These include high sensitivity, high reproducibility, and linearity over a wide range. It has been shown (Bøtter-Jensen et al., 1997) that a standard Al₂O₃:C chip is sensitive to doses below 1 μGy and that its range of linearity with dose is about 0.05 to 50 Gy. Moreover, it is less subject to fading than typical TL dosimeters.

**Figure 4.2.16.** Mass energy absorption coefficients of some TL dosimeter materials (data from Nowotny, 1998).
Because of its higher sensitivity, an OSL dosimeter can be built in smaller sizes than can a TL dosimeter, allowing reliable dose measurements in the presence of a strong radiation field gradient. Another advantage resulting from the use of OSL is the presence of a phenomenon called radioluminescence (RL), which consists of immediate and continuous luminescence during the irradiation. Although RL increases with irradiation time (Aznar et al., 2004), it can provide information about the delivered dose during the irradiation if the dosimeter is coupled with optical fibers. Another recently proposed method for obtaining near-real-time dose measurements in radiotherapy is to stimulate the dosimeter during the irradiation by using the same optical fiber for both the stimulation and the emission (Gaza et al., 2005).

The use of OSL dosimeters has been proposed in mammography (Aznar et al., 2005) and in vivo radiotherapy (Aznar et al., 2004; Yukihara et al., 2005). Currently, although TL dosimetry is a more established technique, OSL dosimetry is becoming competitive, and commercial systems are now available.

4.2.5. SEMICONDUCTOR DOSIMETERS

4.2.5.1. Silicon Diodes

4.2.5.1.1. General Principles

A Si diode can be thought of as a solid-state equivalent of an ion chamber, consisting basically of two Si structures in contact, a p-doped one and an n-doped one. The p-doped Si has an excess of holes, while the n-doped Si has an excess of electrons.

If the two structures are in contact, the electrons migrate from the n-doped to the p-doped region, recombining with holes and leaving behind fixed positive charges in the form of ionized donor impurities. Conversely, holes will migrate to the n-doped region, leaving behind fixed acceptor sites that have acquired extra electrons. The net effect of this charge distribution is a potential gradient.

The junction region remains “depleted,” that is, the density of the remaining charge carriers is negligible. When radiation hits the depleted zone, a number of electron-hole pairs proportional to the dose deposited in the chip are created. Because of the electric field created by charge displacement, electrons are attracted toward the n side and holes are attracted toward the p side.

This effect can be enhanced if the structure is reverse-biased, that is, if the n side is connected to a positive voltage. The potential gradient is enhanced and so is the charge collection efficiency when radiation hits the chip. Moreover, the thickness of the depleted region increases, increasing the efficiency of the diode.

The most typical Si diodes are called p-type and n-type diodes. A p-type Si diode consists of p-doped Si with a thin layer of n-doped Si diffused on the surface. An n-type Si diode consists of a p-doped layer diffused on the surface of an n-doped substrate.

Another diode used in dosimetry is the so-called p-i-n diode (positive-intrinsic-negative), consisting of an intrinsic Si layer with p-doped and n-doped Si diffused on opposite sides (see Fig. 4.2.17). P-i-n diodes have a higher charge collection efficiency than p-n diodes (Knoll, 2000).

4.2.5.1.2. Characteristics of Si Diode Dosimeters

Both n-type and p-type Si diodes are available for dosimetry, p-type diodes being suitable for radiotherapy because they are more radiation resistant and have a smaller dark current (Grusell & Rikner, 1993). Typical Si diode dosimeters can be manufactured in very small sizes (a few hundred microns in size). They are usually sealed to allow measurements in vivo or in water and are provided with suitable buildup caps for different applications.
In clinical practice, Si diodes are frequently unbiased in order to limit as much as possible the leakage current, and hence the noise, allowing the measurement of low doses despite a loss in sensitivity because of the smaller size (a few tens of microns) of the depleted region. This also makes “in vivo” use very convenient.

Although they are more suitable for electron dosimetry than for photon and neutron dosimetry—not being tissue equivalent—Si diodes have become a standard in radiotherapy because of their ease of use.

The performance of Si diode dosimeters is well documented, and typical response characteristics include the following:

- **Linearity**—until radiation damage occurs (see below) Si diodes are linear with dose.
- **Dependence on dose rate**—the sensitivity of most Si diode dosimeters appears to be independent of dose rate for continuous irradiation, while it is an increasing function of the dose per pulse when irradiated with pulsed beams (Grusell & Rikner, 1985; Jornet, Ribas, & Eudaldo, 2000; Marre & Marinello, 2004; Rikner & Grusell, 1983). This effect is more relevant in n-type diodes.
- **Radiation damage**—the sensitivity of both p-type and n-type Si diodes has been shown to decrease sharply initially as a function of dose and decreases at a slower rate (Rikner & Grusell, 1983). Thus frequent calibration is recommended, especially in the initial stages. As previously pointed out, this effect is less important for p-type diodes than for n-type diodes. In order to limit the initial sharp decrease of sensitivity, many commercially available Si diodes are preirradiated with a dose of about 5 kGy to 10 kGy. Another effect related to radiation damage is the loss of linearity with respect to dose rate when Si diodes are used in pulsed radiation beams. This effect is more relevant to n-type Si diodes (Grusell & Rikner, 1985).
- **Angular dependence**—a certain angular dependence related to beam quality and diode geometrical characteristics has been shown, with typical maximum variations as functions of the beam angle ranging from about 2% to 3% to more than 20% (Eveling, Morgan, & Pitchford, 1999; Jornet et al., 2000; Marre & Marinello, 2004).
• Temperature dependence—operating temperature affects the response of Si diodes. Typical changes in sensitivity due to increases in temperature are about 0.2% to 0.3% per degree Celsius (Grusell & Rikner, 1986; Jornet et al., 2000; Saini & Zhu, 2002; Welsh & Reinstein, 2001). This may slightly affect in vivo dose measurements.

4.2.5.1.3. Typical Applications

Silicon diodes are widespread in clinical dosimetry. They are typically used for in vivo measurements of entrance and exit doses, but their small size also allows intracavity use as well as precise measurements of small beams and large gradients in the radiation field. Online measurements are also possible. 2-D arrays can also be built (Jursinic & Nelms, 2003), allowing 2-D dose mapping.

Large area diodes (a few millimeters on a side) are used for measuring low doses, but are not usually used in diagnostic dosimetry because diagnostic X-ray beams are easily monitored by means of tissue-equivalent detectors such as ion chambers or some types of TL dosimeters.

4.2.5.2. MOSFETs

4.2.5.2.1. General Principles

A metal oxide semiconductor field-effect transistor (MOSFET) is a very well-known electronic device, the use of which has been well established in radiation dosimetry since the early 1990s, although it was originally proposed in 1974 (Holmes-Siedl, 1974).

A p-MOSFET, the most frequently used type (see Fig. 4.2.18), consists of an n-doped Si substrate, p-doped Si source and drain, and a metallic gate separated from the substrate by a thin oxide layer. When a gate bias is applied (negative in the case of p-MOSFETs) with an absolute value larger than a value \( V_t \), the area below the oxide is converted into a channel that allows the flow of a source-drain current.

When the chip is irradiated, electron-hole pairs are created in the oxide. A fraction of them recombine, but some of the holes are trapped in the oxide-substrate interface, creating an electric field that increases the absolute value of \( V_t \). The variation in \( V_t \) is directly proportional to the dose delivered in the detector.

Two modes of operation are possible:

Figure 4.2.18. A MOSFET dosimeter: (a) before irradiation and (b) after irradiation.
• Biased mode—a gate-source bias is applied, limiting the recombination of electrons and holes. The variation in $V_t$ is thus increased and the resulting device is more sensitive.  
• Unbiased mode—during the irradiation the chip is kept unbiased. It is read “offline” by measuring the shift in $V_t$ with respect to the nonirradiated system.

The most obvious advantage of the unbiased operation mode with respect to the biased mode is greater portability and the possibility of in vivo use. However, fading occurs in the unbiased mode because of the recombination of electrons and holes. Conversely, a major issue in the biased mode is the fluctuation in $V_t$ caused by changes in external temperature. A possible solution to this problem was proposed by Soubra et al. (1994) and consists of the coupling of two MOSFETs, each with a different initial value of $V_t$. The difference between the values of $V_t$ after the irradiation in the two chips is related to the absorbed dose; the response of the device has been shown to be constant for doses up to 50 Gy to 60 Gy.

4.2.5.2.2. Characteristics of MOSFET Dosimeters
Metal oxide semiconductor field-effect transistor dosimeters share many characteristics with Si diode dosimeters. In particular, because they are not tissue equivalent, they are more suitable for electron dosimetry.

If they are operated in biased mode, MOSFETs offer the possibility of online dose readouts, and commercial systems are available for this task. However, they cannot be used in intensity modulated radiotherapy (IMRT) because of the so-called creep-up effect, which consists of an increase in $V_t$ in consecutive readings depending on the time interval between readings. This effect decreases after a few minutes and occurs for doses above 20 Gy (Ramani, Russell, & O’Brien, 1997). The effect of too short an interval between consecutive readouts is an overestimate of the dose.

The response characteristics of MOSFET dosimeters are the following:

• Linearity: The main advantage of MOSFET dosimeters is the large range of linearity (from a few mGy to tens of Gy).
• Radiation damage: The sensitivity of MOSFETs decreases steadily with accumulated dose, thus frequent calibration is required. Their lifetimes, depending on the beam quality and the operating parameters, have been estimated to be as low as 50 Gy for low-energy X-rays and about 200 Gy for high-energy photon and electron beams (Ehringfeld et al., 2005).
• Angular dependence: Angular dependence values vary significantly from author to author. The maximum reported variation as a function of the beam incidence angle is between 2.5% and 26%, depending on the dosimeter type and beam quality (Chuang, Verhey, & Xia, 2002; Ehringfeld et al., 2005; Ramani, Russell, & O’Brien, 1997; Scalchi & Francescon, 1998).
• Temperature dependence: This is particularly relevant when performing in vivo measurements because the response will depend on the patient skin or internal temperature (Welsh & Rein-stein, 2001).

Because of individual variability, MOSFETs need individual calibration.

4.2.5.2.3. Typical Applications
Metal oxide semiconductor field-effect transistors are suitable for conventional radiotherapy, and because of their small size they are particularly appealing when a high spatial resolution is required.
The use of implanted unbiased MOSFETs has been recently proposed (Beddar et al., 2005), allowing long-term use in radiotherapy treatment. Also, because of their linearity over a wide range, MOSFETs are suitable for use in quality assurance programs.

4.2.6. FILM DOSIMETRY

4.2.6.1. Radiographic Films

4.2.6.1.1. General Principles

The use of radiographic films as dosimeters is well established. A radiographic film comprises a thin layer of polyethylene or other plastic material coated with a layer of emulsion. The emulsion typically consists of silver halide grains (in most cases silver bromide [AgBr]) dispersed in gelatin. Exposure to radiation causes ionization of the halide, forming a so-called latent image. The development process makes the image visible: the darker the area on the film, the higher the dose received.

The image can then be evaluated by means of a densitometer, an instrument capable of measuring the transmission of a small visible-light beam through the film. The output of a densitometer is the optical density (OD), defined as $OD = \log_{10}(I/I_0)$, where $I_0$ is the incident light intensity and $I$ is the transmitted light intensity. The response plot of a radiographic film is called a sensitometric curve (or HD curve, from Hurter and Driffield, who originally investigated it), and it shows the OD as a function of the exposure or the dose.

When a sensitometric curve is plotted, a logarithmic x-axis results in the sigmoid shape shown in Figure 4.2.19, and features five typical zones:

- Fog—the optical density corresponding to no exposure, which is usually subtracted from measurements.
- Toe—a sublinear variation of OD for low exposures.
- Linear—the useful portion of the exposure curve.

![Figure 4.2.19. A typical response for a radiographic film. This characteristic (“HD”) curve plots the OD against the logarithm to the base 10 of the relative exposure.](image-url)
• Shoulder—the end of linear region.
• Saturation—the region of constant OD for high exposure.

As is clear from the figure, the expected exposure range must be known in advance in order to make sure that it falls within the linearity range.

The parameters characterizing a sensitometric curve are the following:

• Latitude—the ratio between the two extremes of the linearity region. For instance, if the maximum exposure within the linearity region is 10 Gy and the minimum is 1 Gy, the latitude is 10:1.
• Gamma—the slope of the linear portion of the curve.
• Speed or sensitivity—the exposure required to produce a density of 1.0 over the fog level.

4.2.6.1.2. Characteristics of Radiographic Film Dosimeters

A radiographic film provides a 2-D distribution of dose with a spatial resolution that can be down to a few micrometers. Characteristic parameters of a radiographic film include the size of the AgBr grains, the number of grains per unit area, and the thickness of the emulsion. The combination of these parameters affects the sensitivity and the spatial resolution of the film.

The main characteristics of radiographic film dosimeters are the following:

• Tissue equivalence: Radiographic films are not tissue equivalent, thus they must be carefully calibrated as a function of the beam energy, especially when used with photons. The result of a change in energy is a change in the slope of the sensitometric curve (Cheng & Das, 1996; Muench et al., 1991).
• Linearity: Although there are significant variations from one type of film to another, the typical range of use for film dosimeters is in tenths of grays (Bos et al., 2002; Cheng & Das, 1996).
• Dose rate dependence: Unlike diagnostic radiographic films, the response of most dosimetry films is nearly dose-rate independent because the energy is deposited by photons and electrons in a single hit, not allowing fading as occurs for visible light. However, a dependence on dose rate has recently been reported for new films (Martens et al., 2002).
• Angular dependence: The angular dependence of a film is limited, unless it is nearly parallel to the radiation beam.
• Sensitivity to ambient parameters: Variations from one batch to another, which are typically related to storage conditions, have been reported (Bos et al., 2002; Childress, Dong, & Rosen, 2002). Thus it is essential that a calibration film be exposed along with each series of films. Film dosimeters are also sensitive to visible and ultraviolet (UV) light, and are commonly packaged in light-tight envelopes.
• Other factors affecting the response: The temperature of the chemicals strongly affects the response of radiographic film (Bogucki et al., 1997), but this is a minor problem because most commonly available developers keep the chemicals at a constant temperature. The type of reader used may also affect the response (Bos et al., 2002).

4.2.6.1.3. Typical Applications

Since they are not tissue equivalent, film dosimeters are particularly suited for electron dosimetry. Their use in quality assurance of radiotherapy machines (congruence of light and radiation fields, determination of collimators position, etc.) is well established.
They are also used in IMRT (Esthappan et al., 2002; Martens et al., 2002) and in interventional radiology (Morrell & Rogers, 2004; Vano et al., 1997).

A common application is in personal dosimeters and film badges; these are equipped with filters of different materials (typically aluminum and copper) and thicknesses. Comparison of readouts in all regions makes possible not only calculation of the absorbed dose, but also determination of the source that produced the exposure.

4.2.6.2. Radiochromic Film

4.2.6.2.1. General Principles
Radiochromic films (the most typical being GAFCHROMIC films) are a more recent alternative to radiographic films as dosimeters. A radiochromic film is a colorless film containing a dye that polymerizes when exposed to radiation, changing the color of the film different shades of blue. The resulting optical density can be read using a laser and photodiode system or a digitizer.

4.2.6.2.2. Characteristics of Radiochromic Film Dosimeters
The typical range of use of radiochromic films is broader and shifted toward high doses in comparison with radiographic films. Radiochromic films have shown a good response from about 10 Gy to 10⁴ Gy (McLaughlin et al., 1991). Their main advantages are the following:

- Good tissue equivalence.
- Independent of the dose rate (Butson et al., 2003; Dini et al., 2005).
- High spatial resolution (about 1200 lp/mm; McLaughlin et al., 1991) because of their grainless nature.
- Nearly flat angular response.
- Limited dependence on ambient parameters within a reasonable range of temperature and humidity (Abdel-Fattah & Miller, 1996).
- Insensitivity to visible light, which implies less stringent requirements in terms of packaging.

The sensitometric curve of radiochromic films depends strongly on the color spectrum of the densitometer in terms of both linearity and sensitivity (Butson, Cheung, & Yu, 2004; Kellermann, Ertl, & Gornik, 1998; Lee, Fung, & Kwok, 2005).

Probably the biggest disadvantage of radiochromic films is their continuous darkening after irradiation. Since this process stabilizes after a few hours (Ali et al., 2003; Meigooni et al., 1996; McLaughlin et al., 1991), it is a common practice to read radiochromic films 24 hours after exposure. This, of course, limits their clinical use.

4.2.6.2.3. Typical Applications
Because of their high spatial resolution and independence of dose rate, radiochromic films are particularly suited for measurements of high dose gradients, for instance, near brachytherapy sources and in stereotactic fields (Sharma, Bianchi, Conte, Novario, & Bhatt, 2004; Chiu-Tsao, de la Zerda, Lin, & Kim, 1994). Their use is also documented at tissue interfaces (Niromaand-Rav et al., 1996; Reinstein, Gluckman, & Meek 1998).
4.2.7. DIAMOND DOSIMETRY

4.2.7.1. General Principles

The potential of diamond detectors was investigated as early as the 1940s, although research in this field was not pursued because of the limited availability and the variability of natural diamonds. Only a very small fraction of natural diamonds—those with very low impurity levels—can be used as radiation detectors, and they need individual calibration and characterization. Research into using diamonds began again after techniques for growing synthetic diamonds, such as chemical vapor deposition (CVD), became established.

Irradiated diamonds show a linear current-voltage characteristic, the resistivity decreasing with the dose rate. This effect is due to the creation of electron-ion pairs using ionizing radiation that drift toward the electrodes when a voltage is applied.

The presence of impurities strongly influences the lifetime of charge carriers in diamond. Thus, in order to ensure full charge collection, it has been calculated that the optimal thickness of a diamond detector is about 0.1 mm to 0.3 mm (Altukhov et al., 2004).

4.2.7.2. Characteristics of Diamond Dosimeters

A typical diamond detector is a few cubic millimeters in volume and is sealed in a resin or polystyrene case. The bias voltage is applied by means of metal contacts. Due to the presence of defects in the diamond lattice, a space charge can be accumulated inside the crystal. This effect is known to stabilize after a certain dose, so diamond dosimeters need to be preirradiated with a dose that may vary from 0.1 Gy to 10 Gy (Cirrone et al., 2003; Prosvirin et al., 2004). The following are some of the characteristics of diamond dosimeters:

- **Tissue equivalence**: One of the main advantages of diamond dosimeters is their near-tissue equivalence \((Z = 6)\), deviations from which are due mainly to the geometry and the composition of the electrodes. However, the energy response is nearly flat for photons above about 200 keV, and for lower energies the energy dependence is weaker than in other detectors (Yin et al., 2004).
- **Radiation damage**: One major advantage of diamond dosimeters is their radiation hardness: they show no change in response for photon doses up to 100 kGy (Bauer et al., 1995) and electron doses up to 1 MGy (Mainwood, 2000).
- **Dose rate dependence**: The current as a function of dose rate shows a small deviation from linearity (Burgermeister, 1981; Planskoy, 1980).
- **Dependence on ambient parameters**: Diamond detectors show a small change in sensitivity related to an increase in temperature (Nakano et al., 2003).
- **Angular dependence**: Angular dependence is negligible (Nakano et al., 2003).
- **Sensitivity**: Because of its smaller atomic number, wider band gap (5.5 eV compared to 3.6 eV for Si), and constraints on thickness, the efficiency of a typical diamond detector is lower than that of a typical Si detector, as shown in Figure 4.2.20. On the other hand, the wider band gap implies a lower leakage current.

4.2.7.3. Typical Applications

The main limitation to the widespread use of diamond dosimeters is their high cost. They are typically used for online dose rate measurements, but can also be used as TL dosimeters for offline integrated dose measurement (Borchi et al., 1998; Cuttone et al., 1999; Mobit & Sandison 1999).
Because of their small size, diamond dosimeters are particularly suitable for measuring strong dose gradients, such as in stereotactic radiosurgery (Heydarian, Hoban, & Beddoe, 1996). They have also provided good results for less conventional uses, such as proton dosimetry (Cuttone et al., 1999; Onori et al., 2000) or $\alpha$-particle dosimetry (Keddy, Nam, & Burns, 1987).

4.2.8. CHEMICAL DOSIMETRY

Chemical dosimetry is based on measurements of chemical changes produced by radiation. As will be discussed, some chemical dosimeters can be used as absolute dosimeters.

The use of gel-based chemical dosimeters represents the only dosimetric technique capable of providing 3-D information in solid phantoms. Gels can be shaped in any form, acting concurrently as dosimeters and phantoms, and are nearly tissue equivalent.

Three types of chemical dosimeters are used: Fricke dosimeters, polymer gels, and alanine dosimeters.

4.2.8.1. Fricke Dosimetry

4.2.8.1.1. General Principles

A Fricke dosimeter is based on the measurement of the fraction of ferrous ions ($\text{Fe}^{2+}$) that have been oxidized to ferric ions ($\text{Fe}^{3+}$) through a series of chemical reactions triggered by radiation. More details about these reactions can be found in Greening (1985).

The original dosimeter, proposed by Fricke and coworkers in 1927, consists of a solution of ferrous sulfate in air-saturated dilute sulfuric acid. The yield of ferric ions is measured by spectrophotometry of the solution, which has absorption peaks in the UV region at 224 nm and 304 nm.

In principle, a Fricke dosimeter can be used as an absolute dosimeter, the absorbed dose in the solution being calculated as follows (Greening, 1985):

$$D = \frac{A - A_0}{\rho G(\text{Fe}^{3+})\varepsilon_m},$$

(4.2.5)

where $A$ and $A_0$ are the optical densities of the solution after the irradiation and before it, respectively; $\rho$ is the density of the solution; $G(\text{Fe}^{3+})$ is the radiation chemical yield of $\text{Fe}^{3+}$; that is, the amount of

Figure 4.2.20. Photon absorption efficiency for a Si detector and a diamond detector, both 300 $\mu$m in thickness (data from Nowotny, 1998).
substance that is produced by radiation per unit energy deposited; \( l \) is the length of the light path of the spectrometer; and \( \varepsilon_m \) is the molar absorption coefficient of \( \text{Fe}^{3+} \) solution: 
\[
\varepsilon_m = A/lc,
\]
where \( c \) is the concentration of the solute.

However, as the parameters affecting the dose response of a Fricke dosimeter vary greatly depending on the preparation, it is more common to use a Fricke dosimeter after a calibration against an absolute dosimeter rather than as an absolute dosimeter itself. Three-dimensional dose distributions can be calculated by means of optical tomography (Kelly, Jordan, & Battista, 1998).

In the 1980s, an alternative technique for measuring the amount of \( \text{Fe}^{2+} \) ions was proposed by Gore, Kang, and Schulz (1984). It consists of measuring the spin relaxation times of the dosimetric solution before and after the irradiation by means of magnetic resonance, based on the linear increase of the relaxation times \( T_2 \) and \( T_1 \) as a function of the concentration of \( \text{Fe}^{2+} \).

Usually the ferrous sulfate solution is dispersed in a gel matrix, typically gelatin or agarose (Schreiner, 2004). This strongly limits the diffusion of ions that causes a blur in the dose distribution when using an aqueous solution and allows 3-D dosimetry. The gel acts as a phantom and as a dosimeter at the same time, removing any problems related to discontinuity in the medium.

4.2.8.1.2. Characteristics of Fricke Dosimeters
As previously mentioned, the main advantage of Fricke gels is the ability to perform 3-D dosimetry. In addition, there are other advantages:

- **Tissue equivalence**: Fricke dosimeters are nearly tissue equivalent (Klassen et al., 1999; Kron, Metcalfe, & Pope, 1993).
- **Linearity**: Fricke dosimeters show good linearity up to a certain dose, depending on the concentration of iron ions. Saturation occurs when all \( \text{Fe}^{2+} \) ions have been converted into \( \text{Fe}^{3+} \). For typical Fricke dosimeters, saturation occurs at about 500 Gy for aqueous dosimeters (Podgorsak & Schreiner, 1992) and at about 50 Gy or less for gel dosimeters (Gambarini et al., 1994; Hazle et al., 1991; Schulz, Venkataramanan, & Huq, 1990). The minimum detectable dose is on the order of 1 Gy for magnetic resonance imaging (MRI) measurements (Olsson et al., 1989) and can be decreased by about one order of magnitude for spectrophotometric readouts (Mattsson, Johansson, & Svensson, 1982).
- **Dose rate dependence**: Fricke dosimeters are not dose rate dependent (Olsson et al., 1989; Schulz et al., 1990).
- **Reproducibility**: Reproducibility is reported to be within a few percent (Chu et al., 1998; Johansson Bäck et al., 1998).

The main limiting factor to the spatial resolution of a Fricke dosimeter is the diffusion of ferric ions after irradiation (Olsson et al., 1992; Schulz et al., 1990). This factor is dependent on the time interval between the irradiation and the readout and on the dose delivered. However, gels with a low diffusion of \( \text{Fe}^{3+} \) have recently been proposed (Chu, Jordan, Battista, Van Dyk, & Rutt, 2000).

4.2.8.1.3. Typical Applications
Fricke dosimeters are only used in beam calibration and for validating Monte Carlo algorithms in radiotherapy. They have proven valuable for both megavoltage and kilovoltage beams (Kron & Pope, 1994). Brachytherapy sources have been characterized (Schreiner et al., 1994) and less conventional uses, such as proton beam dosimetry (Bäck et al., 1999) and neutron dosimetry (Gambarini et al., 2002), have been reported.
4.2.8.2. Polymer Gels

4.2.8.2.1. General Principles

The use of polymer gel dosimeters was first proposed by Maryanski, Gore, Kennan, and Sculz (1993) in order to overcome some of the main limitations of Fricke dosimeters, the most important being the loss of spatial resolution due to ferric ion diffusion.

In a polymer gel dosimeter, monomers such as acrylamide are dispersed in a gel matrix (gelatin or agarose). Upon irradiation, the monomer polymerizes. The degree of polymerization can be determined by MRI (Maryanski et al., 1993), optical computed tomography (CT) (Xu, Wuu, & Maryanski, 2004), X-ray CT (Hilts, Jirasek, & Duzenli, 2005; Jenneson et al., 2004), ultrasound tomography (Mather & Baldock, 2003), proton spectroscopy (Murphy et al., 2000). There is a relation of proportionality between the degree of polymerization and the absorbed dose.

After the development of monomer-polymer gels known as BANG-type (bis-acrylamide, nitrogen, and gelatin) (Maryanski et al., 1996) and PAG-type (polyacrylamide gel), other gels were developed in order to overcome their main limitation, that is, their sensitivity to oxygen, which inhibits polymerization, causing a threshold effect (De Deene, Reynaert, & De Wagter, 2001; Hepworth, Leach, & Doran, 1999). In particular, in the so-called MAGIC gel (methacrylic and ascorbic acid in gelatin initiated by copper), oxygen is necessary for polymerization to take place (Fong et al., 2001).

4.2.8.2.2. Characteristics of Polymer Gel Dosimeters

Polymer gels share many advantages with Fricke gels. In particular good tissue equivalence makes polymer gels suitable for low energies (Pantelis et al., 2004) and for building phantoms of any shape for the calibration of radiotherapy beams and the validation of Monte Carlo protocols.

The main advantages of polymer gels are good tissue equivalence, independence of dose rate (Maryanski et al., 1996; McJury et al., 2000; Novotny et al., 2001), and relatively high sensitivity. Ibbott et al. (1997) reported that doses as small as 0.1 Gy can be detected, while greater sensitivity has been reported by MacDougall, Miquel, Wilson, Keevil, and Smith (2005).

Although the dose values above which linear behavior is lost depend strongly on the composition of the gel (Fong et al., 2001), it can be said that, in general, saturation occurs at higher doses for MAGIC-type gels than for BANG-type gels, the first being able to work at up to several tens of grays (De Deene et al., 2002; Fong et al., 2001; Trapp et al., 2004), while the latter are suitable for dose levels within 10 Gy (Hilts et al., 2005; Maryanski et al., 1996). Gels with added ascorbic acid have a much wider range of linearity (De Deene et al., 2002).

On the other hand, the main limitation to the use of polymer gel dosimeters is a variability in response that can be as large as 15% from one dosimeter to another (MacDougall et al., 2005). Another serious limitation to widespread use is their toxicity.

Also, polymerization occurs after irradiation, causing a change in response that eventually reaches a plateau as much as several days later (MacDougall et al., 2005; McJury et al., 1999a). Postirradiation polymerization depends on several factors, including light, heat, and impurities (McJury et al., 2000). Some diffusion of monomers may also introduce errors in measurements of steep dose gradients (Vergote et al., 2004).

4.2.8.2.3. Typical Applications

The application of polymer gels for the calibration of several types of radiotherapy sources has been evaluated. In particular, their use has been proposed for brachytherapy sources (Kipouros et al., 2003;
McJury et al., 1999b), for IMRT (Wuu, Xu, & Maryanski, 2004), and for stereotactic radiosurgery (McJury et al., 2000; Novotny et al., 2002). Moreover, the use of polymer gels for neutron dosimetry appears promising (Uusi-Simola et al., 2003).

4.2.8.3. Alanine/Electron Spin Resonance Dosimetry

4.2.8.3.1. General Principles

Alanine is an amino acid that can be used for high dose dosimetry in the form of pellets, films, or gels. Radiation causes the formation of alanine radicals, the concentration of which can be measured by means of electron spin resonance (ESR) or electron paramagnetic resonance (EPR) (Pilbrow, 1996).

4.2.8.3.2. Characteristics of Alanine Dosimeters

Alanine dosimeters are nearly tissue equivalent (Olsson, Bergstrand, Carlsson, Hole, & Lund, 2002) and have a fairly flat energy response over the typical energy range for radiotherapy (Zeng et al., 2004, 2005). They have a linear response over a wide range up to above 40 kGy (Olsson, Lund, & Erickson, 1996; Wieser & Girzikowsky 1996).

The main limitation of alanine/ESR dosimetry is its low sensitivity; high doses are necessary to obtain good reproducibility (De Angelis et al., 1999). However, it has been reported that preirradiation with about 50 Gy allows for subsequent detection of doses as low as 2 Gy (Olsson et al., 1996).

The response depends on both temperature and humidity (Arber & Sharpe, 1993; Bugay et al., 2000; Olsson et al., 2002), thus the calibration must take into account environmental conditions. Although some fading has been reported (Arber & Sharpe, 1993; Olsson et al., 1996), alanine dosimeters are suitable for long-term information storage.

Because of the necessity of using a minimum volume of about 0.05 cm³ in order to achieve an acceptable SNR ratio (Olsson et al., 2002), the spatial resolution of alanine dosimeters is not as good as that of other solid-state dosimeters. However, it is acceptable for most typical radiotherapy applications.

4.2.8.3.3. Typical Applications

Because of their characteristic of long-term stability after irradiation, the use of alanine dosimeters appears promising for dosimetry comparisons among different centers (Bartolotta et al., 1993; Gall et al., 1996). Also, because of their low sensitivity, they are suitable for high dose applications, and in particular, high dose rate brachytherapy (De Angelis et al., 1999; Guzman Calcina, et al., 2005; Olsson et al., 2002), proton beam therapy (Gall et al., 1996; Onori et al., 1997), and dose monitoring in radiation-harsh environments such as particle accelerators.

4.2.9. SCINTILLATION DOSIMETERS

4.2.9.1. General Principles

Scintillation dosimetry has been proposed relatively recently and is not yet a well-established technique. Scintillation detectors are based on the conversion of the kinetic energy of charged particles into detectable light and are optically coupled to a photomultiplier. A typical scintillation material must exhibit a high light yield proportional to the energy deposited, and should be transparent to the wavelength of its own emission (Knoll, 2000). Moreover, the decay time of the induced luminescence
must be short enough to allow fast signal pulses. This is particularly relevant when a fast response is needed (e.g., in IMRT). For a complete discussion of the characteristics of scintillation detectors, see Knoll (2000).

4.2.9.2. Characteristics of Scintillation Dosimeters

A typical scintillation dosimeter consists of a scintillation detector, a fiber-optic bundle, and a photomultiplier tube optically coupled to each other. The detectors are usually manufactured in small sizes (a few cubic millimeters) in order to perform dose measurements with high spatial resolution. Plastic scintillators are nearly tissue equivalent (Beddar, Mackie, & Attix, 1992a).

The background light and the Cherenkov light contribution can be removed by using two light guides, one coupled to the scintillator for the signal, and an entirely separate one for measuring the background. The response from the latter can be subtracted from the total signal arising from the scintillator (Beddar et al., 2003).

Scintillation dosimeters are linear with dose and are able to detect doses in the tens of grays (Beddar, Mackie, & Attix, 1992b; Letourneau, Pouliot, & Roy, 1999). They also have good reproducibility and their response is independent of dose rate (Beddar, Mackie, & Attix, 1992b).

4.2.9.3. Typical Applications

The use of scintillator dosimeters has been proposed for those cases in which fast response or high spatial resolution is required, such as in IMRT and stereotactic radiosurgery (Letourneau, Pouliot, & Roy, 1999). The use of plastic scintillators for a 2-D dosimeter was proposed by Kirov et al. (1999), and recently a system for 3-D dosimetry based on a liquid scintillator for eye plaque applicators has been developed (Kirov et al. 2005).

4.2.10. CALORIMETRY

In this section a brief description is given of the working principle and the use of calorimeters as absolute dosimeters. In principle, measuring the increase in temperature of a material after irradiation is the most direct way of measuring the energy deposited in that material. However, an apparatus based on this method has low sensitivity (temperature increases on the order of fractions of millikelvin have to be measured [Seuntjens & Palmans, 1999]), is not easily portable, requires a long time in order to reach thermal equilibrium, and requires high precision. Thus calorimetry is limited to primary standard laboratories or to a research environments, and calorimeters are not usually commercially available for clinical measurements, although some portable devices have recently been proposed (McEwen & Duane, 2000; Palmans et al., 2004).

A calorimeter consists of a known mass of material (usually water or graphite) equipped with a high precision thermometer probe. Upon irradiation, the increase in temperature is measured and the total energy deposited is calculated according to the following equation:

\[ E = \Delta Tm, \]  

(4.2.6)

where \( E \) is the energy deposited, \( c \) is the specific heat of the material, and \( m \) is the mass of the material. The dose deposited in the calorimeter is equal to \( E/m \), and thus is given by

\[ D = \Delta Tc. \]  

(4.2.7)
These equations are, of course, extremely simplified and hold if no chemical reactions take place upon irradiation and if the calorimetric material is thermally insulated from the environment. In order to achieve the latter condition, calorimeters are surrounded by insulating materials.

Originally calorimeters were based on graphite, because of the good tissue equivalence of that material, its ease of handling, and its high heat conductivity, which limits the time necessary to reach thermal equilibrium. More recently, water-based calorimeters have been built. The main reason for the increasing interest in water calorimetry is that dose in water is obtained directly, with no corrections being needed for the different weight of interaction mechanisms or for different energy absorption coefficients (Nutbrown et al., 2002).

Many issues with water calorimeters remain however. First, water has a higher heat capacity than graphite, so the same dose causes a smaller increase in temperature in water than in graphite. Second, the heating of other calorimeter components, having a lower heat capacity than water, can significantly affect the measurement if these components are too close to the point of measurement (Ross & Klassen, 1996). These include the thermometer probe, which must be as small as possible.

A more relevant issue is the chemical complexity of water. When water is irradiated, up to fifty reactions may take place, and the so-called heat defect (i.e., a correction for radiation-induced chemical changes) must be taken into account (Klassen & Ross, 1997). In order to limit the number of factors contributing to this heat defect, the purity of the water must be very high, or it can be saturated with nitrogen, hydrogen, or a mixture of oxygen and nitrogen (Seuntjens & Palmans, 1999).

Other physical factors that may affect the performance of a water calorimeter are convection, which can be a nonnegligible problem when measuring beams directed horizontally (Ross & Klassen, 1996), and thermal expansion. In order to limit the relative importance of the latter, the operating temperature of the calorimeter must be kept between 1°C and 4°C (Ross & Klassen, 1996).

A major technical issue that must be taken into account is the design of the temperature probe, which, as mentioned above, must be small enough to avoid introducing errors in the measurement due to radiation-induced heating of the probe. It must also be electrically insulated from the water (Ross & Klassen, 1996; Schulz, Wuu, & Weinhouls, 1987).

4.2.11. HOW TO CHOOSE A DOSIMETER

A dosimeter must be chosen according to the characteristics that make it suitable for a specific application. Table 4.2.2 summarizes the advantages and disadvantages of the dosimeters described in the previous sections.

4.3. SENSORS FOR RADIATION IMAGING

4.3.1. BACKGROUND

Wilhelm Conrad Röntgen discovered X-rays more than a century ago (November 8, 1895), and X-ray imaging dates back to December 22, 1895, when the world-famous radiograph of his wife’s hand was taken. From the very beginning, the primary detector for imaging has been film. The most important subsequent development was the introduction of the intensifier screen. The intensifier screen absorbs incoming X-rays and reemits the absorbed energy in the form of visible light, to which the film is sensitive. Thus, at the cost of diminished spatial resolution (the lateral spread of the visible light being
### Table 4.2.2. Advantages and disadvantages of the various dosimeters

<table>
<thead>
<tr>
<th>DEVICE</th>
<th>ADVANTAGES</th>
<th>DISADVANTAGES</th>
<th>MAIN USES</th>
<th>TYPICAL PRECISION ACHIEVABLE (95% CONFIDENCE LEVEL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ionization</td>
<td>Can be tissue equivalent</td>
<td>High voltage needed</td>
<td>Parallel plate: diagnostic beams, electron beams below 10 MeV</td>
<td>±0.5%</td>
</tr>
<tr>
<td>chamber</td>
<td>Absolute dosimetry possible</td>
<td>Temperature needed</td>
<td>Free air: calibration of photon beams below 300 keV</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Very limited radiation damage</td>
<td>Recombination at high pulse rates</td>
<td>Farmer: electron beams above 10 MeV</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Online response</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TL dosimeter</td>
<td>Portability</td>
<td>Sensitive to ambient parameters and handling</td>
<td>In vivo measurements</td>
<td>±2%–3% discs,</td>
</tr>
<tr>
<td></td>
<td>Reusability</td>
<td>Limited range of linearity (depending on TL material)</td>
<td>Depth dose measurements</td>
<td>±1%–2% powder</td>
</tr>
<tr>
<td></td>
<td>Low cost</td>
<td></td>
<td>(both for diagnostic and radiotherapy beams depending on TL material)</td>
<td>(Horowitz, 1981; Johns &amp; Cunningham, 1983)</td>
</tr>
<tr>
<td>Si diode</td>
<td>Almost independent of dose rate</td>
<td>Not tissue equivalent</td>
<td>In vivo entrance and exit dose measurements</td>
<td>±1%</td>
</tr>
<tr>
<td></td>
<td>Good linearity</td>
<td>Subject to radiation damage</td>
<td>In vivo entrance and exit dose measurements</td>
<td>(Loncol, Greffe, Vynckier, &amp; Scalliet, 1996)</td>
</tr>
<tr>
<td></td>
<td>High sensitivity</td>
<td>Angular dependence</td>
<td>Intracavity measurements</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Online response</td>
<td>Temperature dependence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Small size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiographic</td>
<td>Portable</td>
<td>Not tissue equivalent</td>
<td>2-D mapping of radiotherapy beams</td>
<td>±2%–3% (Zhu et al., 2002)</td>
</tr>
<tr>
<td>film</td>
<td>2-D dosimetry possible</td>
<td>Limited range of linearity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low energy dependence for electrons</td>
<td>Variation from batch to batch</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Small angular dependence</td>
<td>Nonreusable</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dose rate independent</td>
<td>Dependence on the conditions of the development process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOSFET</td>
<td>Portability</td>
<td>Subject to radiation damage</td>
<td>In vivo, online measurements</td>
<td>±1.5% (Jornet et al., 2004)</td>
</tr>
<tr>
<td></td>
<td>Implantable</td>
<td>Not tissue equivalent</td>
<td>Intracavity measurements</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High sensitivity</td>
<td>Not radiation transparent</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Online response</td>
<td>Directional dependence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wide range of linearity</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*continued on next page*
Table 4.2.2. Advantages and disadvantages of the various dosimeters (continued)

<table>
<thead>
<tr>
<th>DEVICE</th>
<th>ADVANTAGES</th>
<th>DISADVANTAGES</th>
<th>MAIN USES</th>
<th>TYPICAL PRECISION ACHIEVABLE (95% CONFIDENCE LEVEL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OSL dosimeters</td>
<td>Wide range of linearity</td>
<td>Not well-established technique</td>
<td>In vivo diagnostic and radiotherapy dosimetry</td>
<td>±0.2%–0.6% OSL; ±0.6%–1% RL (Aznar et al., 2004)</td>
</tr>
<tr>
<td></td>
<td>Near-real-time information</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diamond</td>
<td>Radiation hardness</td>
<td>High cost</td>
<td>Strong gradient radiotherapy beams</td>
<td>±1% (Fidanzio et al., 2000)</td>
</tr>
<tr>
<td></td>
<td>Tissue equivalence</td>
<td>Need for preirradiation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Small size</td>
<td>Need for high voltage</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Negligible directional dependence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricke dosimeters</td>
<td>Absolute dosimetry possible</td>
<td>Dependence on preparation</td>
<td>3-D beam calibration</td>
<td>±1.5% (MacDougall et al., 2002)</td>
</tr>
<tr>
<td></td>
<td>Tissue equivalence</td>
<td>Nonreusable</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-D dosimetry</td>
<td>Low sensitivity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good linearity</td>
<td>Diffusion</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dose rate independent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polymer gels</td>
<td>Tissue equivalence</td>
<td>Toxicity</td>
<td>3-D beam calibration</td>
<td>± 1.2–3.5% within the same batch (MacDougall et al., 2005)</td>
</tr>
<tr>
<td></td>
<td>3-D dosimetry</td>
<td>Strong dependence on preparation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good linearity</td>
<td>Sensitivity to oxygen (BANG gels)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-dependence of dose rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alanine</td>
<td>Wide range of linearity</td>
<td>Low sensitivity</td>
<td>Comparison between centers</td>
<td>±8%–2% for 2–10 Gy (Onori et al., 1997)</td>
</tr>
<tr>
<td></td>
<td>Long-term information storage</td>
<td>Poor spatial resolution</td>
<td>Harsh radiation environment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dose rate independent</td>
<td>Dependence on ambient parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scintillation detectors</td>
<td>Tissue equivalence</td>
<td>Not well-established technique</td>
<td>IMRT, stereotactic radiosurgery</td>
<td>±3% (Bambyndk et al., 2000; Petric, Robar, &amp; Clark, 2006)</td>
</tr>
</tbody>
</table>
roughly comparable to the screen thickness), the intensifier screen allows a dramatic dose reduction because the X-ray stopping power of plain film is only 1% to 3% that of a modern intensifier screen (Sabel & Aichinger, 1996).

There are many different reasons why film has survived for so long in X-ray imaging. An incomplete list would include the facts that film is extremely practical, easy to manufacture and use, reliable, and cost effective. In addition, its good spatial resolution makes it competitive for specific applications.

However, film does have a range of weaknesses. Among these, the most important is probably its poor dynamic range, which is intrinsically connected with the specific working principles of film and is thus inescapable. Basically a film consists of grains of silver bromide suspended in a gelatin. These grains offer a much larger cross section to the incoming radiation than does the gelatin, hence most interactions take place in the grains. To oversimplify somewhat, the resulting photoelectrons can be trapped in the grains, and when this happens the grains become “sensitized.” Film processing removes the nonsensitized grains and converts the sensitized ones into silver, thus forming the image (Mees & James, 1966). It is therefore clear that the range of the film is strictly bracketed between sensitizing almost no grains at all (“white” image) and nearly all of them (“black” image), with no physical possibility of reaching beyond these limits. The phenomenon is effectively expressed by the film “characteristic” or “HD curve” (from Hurter and Driffield, who described it in 1890). Figure 4.2.19 (see section 4.2.6.1.1) gives the optical density of film as a function of exposure (Johns & Cunningham, 1983). The maximum contrast achievable in film imaging is thus limited, making it difficult to visualize different details on the same image. This is a problem relevant to those fields where large absorption differences are encountered. For example, exposure differences of two orders of magnitude can be encountered in the field of mammography, according to Maidment, Fahrig, and Yaffe (1993). Moreover, an unavoidable amount of image noise is caused by the intrinsic granularity of film (Barnes, 1982), and the detection efficiency is not optimized.

Further relevant limitations of analog film imaging become clear when the advantages of digital imaging are described. Basically, when digital sensors are employed, the different processes of acquisition, display, analysis, and archiving of the image are physically separated and can thus be optimized individually (Brody, 1984). This makes possible a series of improvements, as follows. First, digital imaging detectors provide much wider dynamic ranges that can become virtually infinite in the case of photon counting devices. Second, the possibility of image manipulation allows for an impressive range of operations, from the use of different gray and color scales to enhance detail visibility and stretch the contrast, to the use of sophisticated analysis tools via many kinds of image filtration. Third, digital images can be stored and archived, and can be transmitted over communication networks for remote diagnoses. Furthermore, they enable the implementation of computer-aided diagnosis techniques, which are capable of providing excellent results (see, for instance, Ikeda et al., 2004) and may be crucially important in reducing the costs of screening surveys. In order to apply the same procedures to a film image, it must be digitized by means of a scanner, which, in addition to the time needed, can introduce further sources of noise and distortion.

The use of digital detectors was boosted by the introduction of computed tomography (CT) in the early 1970s (Ambrose, 1973; Ambrose & Hounsfieild, 1973; Hounsfieild, 1973). In CT, the use of digital detectors is mandatory, and the advantages of having access to 3-D information prevailed over all the disadvantages imposed by the limited detector and computer technology available at the time. For planar imaging, however, these limitations prevented an immediate acceptance of digital methods, because the digital methods were not considered competitive with the high resolution and large area
coverage provided by films and screen-film systems. The increasing use of CT images helped make clear the advantages of digital imaging, thus fostering their further development. The subsequent huge expansion of computer and detector technology eventually made today’s systems possible.

4.3.2. DETECTORS FOR MEDICAL AND BIOLOGICAL IMAGING

In this section, some of the main types of detectors available for medical and biological imaging will be individually reviewed. Some recent developments and experimental prototypes of interest will also be briefly discussed within each appropriate subsection.

In general, detectors for X-ray imaging are subdivided into direct and indirect detection devices. In the direct mode, the X-ray is directly converted into electric charge in the detector material, while in the indirect mode the photon is first converted into visible light by a luminescent material, after which the visible photons are converted into charge in the detector material. Luminescent materials are usually called “phosphors” when they are coupled to detectors operating in the integration mode (e.g., charge-coupled devices [CCDs]) and “scintillators” when they are coupled to photon counting devices, as occurs for example in most nuclear medicine applications. However, it should be noted that there is no neat distinction between phosphors and scintillators, and the two terms are often used as synonyms. An overview of the most commonly used luminescent materials and their properties is presented in section 4.3.2.4.1.

There are basically three optical coupling methods employed in the indirect detection mode: proximity, lens, and fiber-optic coupling. Proximity coupling simply involves keeping the phosphor as close as possible to the detector’s active surface, and provides the best results when that surface is directly coated with the phosphor material (although this can result in direct interaction of X-rays in the sensor material). However, because many devices (CCDs in particular) have limited sensitive areas, active surface magnification is often achieved by coupling a small detector area with a larger phosphor layer. In this case, relying on lens or fiber-optic coupling is mandatory. Lens coupling is simple and convenient, but often inefficient. Moreover, its use can result in geometric distortions (see, e.g., Liu et al., 2000). Thus more efficient and reliable coupling by fiber-optic bundles is often preferred. Although their use results in fixed pattern noise artifacts, as the boundaries of the individual bundles are usually discernible in the acquired image, these are easily removed by simple correction algorithms.

The basic operating principles of each detector device will be briefly reviewed and it will be made clear whether each specific device operates in the direct or indirect detection mode (or in both modalities).

4.3.2.1. COMPUTED RADIOLOGY: STORAGE PHOSPHOR SYSTEMS

“A new system of computed radiography that is based on new concepts and the latest computer technologies has been developed. This system eliminates the drawbacks of conventional screen-film radiography. The basic principle of the system is the conversion of the X-ray energy pattern into digital signals utilizing scanning laser stimulated luminescence (SLSL).” In this way, Sonoda, Takano, Miyahara, and Kato introduced the photostimulable phosphor (or storage phosphor, or image plate, or sometimes simply “computed radiography” [CR]) system in 1983 (Sonoda et al., 1983). This was one of the first attempts at film substitution, and time has proved it to be a very successful one: the system rapidly found extensive acceptance and is still widely used.
Different from conventional phosphors, photostimulable phosphors intentionally contain traps that capture the charges generated by X-ray interactions, so that exposure and readout phases are separated and the image is actually “stored” in the plate. The trapped charge is then released by external stimulation with a (red) laser, and its subsequent decay generates luminescent (blue) light that is collected by a light guide and detected by a photomultiplier tube (Fig. 4.3.1).

The critical parameter is the energy separation between the traps and the conduction band in the phosphor: this has to be small in order to facilitate laser stimulation, but cannot be too small because random charge release due to thermal excitation might affect image storage. Furthermore, during X-ray irradiation, part of the released charge is captured in the traps while part of it emits prompt light (which is usually discarded). This might seem to favor a design with increased trapping efficiency. However, such an increase would also imply that during laser stimulation, a large fraction of the released charge falls back into the traps instead of producing luminescence. It is easy to conclude that optimum balance can be achieved with a trapping efficiency of 50%, which means that, with respect to a conventional phosphor, only 25% of the generated charge actually results in luminescent light output. Although the basic working principle of the system is explained rather easily, some specific details of the photoluminescence mechanism have not yet been completely clarified (Blasse & Grabmaier, 1994; Seibert, 1997).

One of the technological challenges that had to be solved in CR system development was to prevent laser photons from contaminating the luminescent light readout. Although in principle it is not difficult to separate the laser red photons from the luminescence blue ones, it should be noted that when a low X-ray exposure has to be read, the ratio between blue and red photons can be as small as $10^{-8}$. This problem is solved by adding proper light filters to the readout chain and by selecting photomultiplier photocathodes with much higher efficiencies in the blue than in the red wavelengths. For commonly used bialkali photocathodes, this is typically 25% in the blue and 0.1% in the red wavelengths.

Originally CR systems were based on BaFBr:Eu$^{2+}$ (the dopant material is indicated after the colon). Basically, any BaFX:Eu$^{2+}$ compound can be used where X is bromine (Br), chlorine (Cl), iodine (I), or any mixture of these. In recent years, BaFBr$_{1-x}$I$_x$:Eu$^{2+}$ (where $x \leq 0.2$, usually 0.15) has found widespread acceptance, mainly because of the good match of its optimal stimulation wavelength with that emitted by conventional diode lasers.

Figure 4.3.1. Basic principle of the storage phosphor system. The excited electron is stored in a trap located at an energy $E_t$ below the conduction band. During the readout, it is released via laser stimulation and decays, emitting a luminescent photon of energy $E_f$. 
Regarding the spatial resolution of the system, the equivalent of the aperture is the cross section of the laser spot, while the spacing is determined by the scanning raster. Aside from the fact that it is not a real-time system, spatial resolution probably represents its main limitation because of the unavoidable process of light scattering inside the phosphor. Basically, since a flying laser spot is usually employed to scan the plate, this limitation arises from the scattering of laser light rather than luminescent light. The problem is also made more severe because of flare (i.e., multiple reflections from the plate to the face of the light guide and then back to the plate, resulting in luminescent light emitted from other, randomly located points) and halation (i.e., laser light reflected back and forth inside the protective layer deposited above the phosphor), which will eventually stimulate luminescent emissions from a ring of positions around the original point targeted by the laser.

Another limit is encountered in the noise performance of the system because it has a secondary quantum sink. To give an example, an absorbed 50 keV X-ray produces approximately 2000 hole-pairs in the phosphor plate; of these, about 7% are actually trapped. Fifty percent of the trapped electrons are released by the laser scan, and only 33% of the produced luminescent light is transported to the photomultiplier tube by the light guide. Here, only 25% of the photons are converted into photoelectrons, because of photocathode efficiency limitations. The system gain is thus \( G = 2000 \times 0.07 \times 0.5 \times 0.33 \times 0.25 = 5.8 \) (Rowlands, 2002). On a scale in which a photon counting detector represents the “optimum,” because the only source of noise arises from the statistical nature of X-ray absorption and emission, the smaller the gain, the worse the noise performance (e.g., a flat-panel systems has \( G = 1000 \)). However, the nonideal behavior encountered in the different stages outlined previously indicates that there is still room for possible developments (improved charge detrapping by laser stimulation on both sides of the plate, improved collection of the photostimulated luminescence, etc.). A comprehensive discussion of these can be found in Rowlands (2000).

Despite the extremely wide range of different detector systems developed during recent years, in terms of everyday practice, CR systems might be superseded only as a result of the greater acceptance of (much more expensive) flat-panel technology, and probably only to a limited extent. In some applications, such as emergency and bedside radiology, the CR system is still the most practical because of its ease of handling and portability, along with a wide dynamic range that largely solves the problem of over- or underexposure in conditions where it is very difficult, if not impossible, to rely on automatic exposure control.

4.3.2.2. Image Intensifiers

An image intensifier is a fast, reliable, and widely used real-time device that for many years has been the only tool for fluoroscopy. In the mid-1980s it was also used for CT applications in which high image quality was not the primary issue (e.g., radiotherapy planning; see, for instance, Arnot et al., 1984).

An image intensifier consists basically of an evacuated tube with an intensifier screen at either end and appropriate electron optics in between (Fig. 4.3.2). The input intensifier screen (typically CsI:Na because it can be deposited with a packing efficiency of nearly 100% with no binder and can be grown in columnar structures) is coupled to a photocathode (e.g., Cs,Sb). X-rays impinging on the intensifier screen generate visible light photons that are converted into photoelectrons by the photocathode. These electrons then travel the whole length of the tube driven by the electric field shaped by the electrodes of the electron optical system. The electrons are accelerated along their path by an applied voltage (20 kV–35 kV), striking the output phosphor (ZnS:Cu, CdS:Cu, ZnS:Ag, CdS:Ag), which
reconverts them into visible light with a high gain. The output screen is then optically coupled to a video system, originally a TV camera, but now in most cases a CCD.

All the components have nonunity efficiency. To give an example, for a 50 keV X-ray absorbed by the input screen, approximately 2000 light photons will be generated, and typically slightly less than half of these will reach the photocathode. Typical photocathode efficiencies are in the range of 10% to 15%, which means that slightly more than one hundred electrons will be released. Electron optical coupling is usually very efficient, possibly greater than 90%. A typical applied voltage is 25 kV, and a 25 kV electron produces approximately 2000 light photons when hitting the output phosphor, so, given the previous approximations, the result is the emission of approximately 200,000 light photons. Finally, the efficiency of the output phosphor has to be taken into account, resulting in an output of \(1 \times 10^5\) to \(1.5 \times 10^5\) visible photons per incident 50 keV X-ray.

In addition is the fact that in order to preserve the vacuum, an entrance window is always present, which means that a certain fraction (5%–10%, depending very much on the incident X-ray spectrum) of the incoming photons will not reach the input phosphor. Originally this entrance window was made of glass; currently aluminum is the preferred material. Actually beryllium is the optimum choice because of its low atomic number, but its use is restricted to nondestructive testing because of its extremely high toxicity.

The main advantages of the image intensifier can be summarized as follows:

- almost aberration-free electron-optic coupling;
- high gain achievable by electron coupling, which overcompensates for subsequent losses in the imaging system;
- Large area coverage (up to 60 cm in diameter); and
- low delivered dose.

The main disadvantage is the relatively low spatial resolution: in general, the lack of sharpness increases with size minimization; the resolution is lower than in most other imaging techniques. Image contrast is also perceptibly lower compared to that obtainable with other imaging devices.

![Figure 4.3.2. Schematic of an image intensifier.](image-url)
Furthermore, in practically all cases, although to a degree that can vary from one system to another, images suffer from spatial distortion and spatial nonuniformity (the “vignetting” effect, i.e., the image is brighter in the center than on the periphery).

4.3.2.3. Pixel Detectors

An impressive range of pixel detectors have been (and are being) developed for a variety of different applications. Of these, CCDs and the so-called flat-panel systems are currently the most widely used in a range of medical and biological imaging applications, thus they will be treated separately in the following subsections. Their widespread acceptance was mainly due to their compactness, cost-effectiveness, and simplicity of use. Recently, however, CMOS sensor technology has proven to be competitive, especially along with CCD methods. Thus some developments based on CMOS sensors will also be briefly discussed in a dedicated subsection.

Apart from these “dominating” technologies, several other techniques have been developed or are under development. A considerable number of the relevant prototypes are based on the single photon counting mode rather than on the integration method, which means that they are affected only by the Poisson noise due to X-ray interactions and are thus capable of providing optimized image contrast. Most of them are basically spinoffs from research carried out in the field of high-energy physics, primarily regarding the design of vertex detectors (i.e., the inner part of the “barrel” of detectors surrounding the interaction region in collider experiments).

In this field, the dominating material is silicon, the relevant technology of which is well established: its charge collection properties are known in detail and it allows for excellent shaping of the electric field by tailored impurity doping. Silicon detectors are made from crystalline silicon wafers, with thicknesses usually on the order of hundreds of micrometers. Structures are formed on one of the large surfaces in the shape of rectangular pixels or strips. In the former, each individual pixel on the silicon wafer can be connected to its own electronic readout channel by bump-bonding the wafer to a readout chip: in this way, a pixel detector is obtained. This arrangement is often referred to as the “hybrid pixel technique” (Wermes, 2004), as the sensor and the front-end chip are physically separated parts of the detector module (see Fig. 4.3.3). However, the limited thickness of the wafer and the low...
stopping power of silicon result in low detection efficiency, thus preventing, to a large extent, the utilization of such a device in medical imaging unless indirect detection methods are employed. Thus its use is limited to biological applications (see, for instance, Datte et al., 1999), while research on other, higher Z materials is carried out to make this device usable in medical applications.

Regarding silicon, other solutions have been proposed involving the use of strip detectors. In order to improve detection efficiency, the use of such devices in edge-on geometry has been investigated (Arfelli et al., 1998). In this modality, the radiation impinges on the side rather than on the surface of the chip, with the direction of the incoming photons parallel to the implanted strips. Thus a much greater thickness (a few centimeters instead of a few hundred micrometers) is available for photon conversion. This solution is well suited to scanned acquisition systems, in which it provides a range of additional advantages. In fact, the first mammography system based on this concept has recently been commercialized by Sectra, SE (http://www.sectra.com).

Another solution implies the use of wafers with strips implanted on both surfaces, in orthogonal directions, in order to retrieve 2-D information (Alfano et al., 1993). In this case the chips are read from the sides, rather than from below as occurs in the hybrid pixel geometry. In principle it is possible to increase the detection efficiency by stacking several devices one top of another (see Fig. 4.3.4). The same detector has also been used in the somewhat less demanding field of autoradiography (Bertolucci et al., 1996).

As already mentioned, another solution consists of implementing the hybrid pixel technique with higher stopping power sensor materials; the CERN-based MEDIPIX collaboration is an outstanding reference point for the research carried out in this field (http://medipix.web.cern.ch). For instance, the use of GaAs was investigated for potential applications in mammography (Amendolia et al., 2001) and in autoradiography (Abate et al., 2001). More recently, the use of CdTe has been suggested for a range of possible applications (Chmeissani et al., 2004). Another approach to the use of GaAs detectors, based on the integration readout mode rather than on photon counting, was proposed by Sellin et al. (2001).

![Figure 4.3.4. Schematic of a microstrip detector used in edge-on geometry (left) and a double-sided microstrip detector (right).](image)
The use of CdZnTe has also been proposed for mammography applications because of its high stopping power (Mainprize et al., 2002a). In this specific case, a CCD was used as the readout device by bump-bonding it to the CdZnTe sensor.

The reported examples provide a coarse sampling of the wide range of research that is being carried out on innovative imaging devices. The subject was touched upon mainly to show that this is a huge, continuously growing field.

For the sake of completeness, photodiode arrays must be mentioned here because they were one of the first solid-state pixel devices and are still widely used (an exhaustive description can be found in Knoll, 2000). Developments based on gas detectors will also be discussed.

Gas detectors are an offspring of research carried out in high-energy physics and have been used in a range of medical and biological applications. X-ray absorption results in the creation of an electron-ion pair (instead of an electron-hole pair), and the created charge is driven to the readout electrodes by an electric field applied to the gas. The main problem with gas detectors in medical and biological imaging is their low X-ray stopping power, requiring the use of gases with the highest $Z$, xenon in particular. This is especially relevant in medical imaging, where the detection efficiency is a primary issue affecting the delivered dose, but it is a limiting factor in biological imaging as well. Nevertheless, several approaches using gas detection have been explored, especially those based on multiwire proportional chambers (MWPCs), because of their spatial resolution. These were invented by 1992 Nobel Prize Laureate G. Charpak (Charpak & Sauli, 1979). Some examples of autoradiography applications based on these detectors can be found in Angelini, Bellazzini, Brez, Massaia, and Torquati (1988) and in Dominik et al. (1989).

In medical applications, gas detectors have been mainly arranged in a geometry resembling an edge-on strip detector in such a way that a relevant gas depth is made available for photon conversion. For example, the NIKOS project for synchrotron radiation (SR) coronary angiography, historically the first SR-based program that carried out a relevant number of investigations on human patients, employed this kind of detector design (Dill et al., 1998). Much more recently, Despres et al. (2005) proposed a similar design for orthopedic X-ray imaging. In some cases, implementation of the time delay integration (TDI) readout mode has been proposed for gas detectors in a way substantially corresponding to that employed with CCDs. This approach was proposed by DiBianca and Barker (1985) and further developed by Wagenaar and Terwilliger (1995).

Some further examples of promising developmental detectors for radiation imaging will be discussed at the end of the following subsection because they are substantially upgrades of the basic CCD architecture.

### 4.3.2.3.1. Charge-Coupled Devices

Charge-coupled devices developed at Bell Laboratories in 1969 (Boyle & Smith, 1970) are probably the most compact and practical detectors currently on the market. They prevailed over competing technologies because a smaller pixel size was achievable and because they proved to be much less affected by fixed pattern noise, which was considered at the time to be the most limiting constraint associated with the use of MOS-based devices (Fry, Noble, & Rycroft, 1970). Subsequently, impressive improvements in terms of quantum efficiency, fill factor, dark current, charge transfer efficiency, smear and lag suppression, readout rate, full well capacity, and noise performance were achieved for CCDs.

A CCD is basically a monolithic silicon chip subdivided into columns by implanted potential barriers (channel stops). The surface of the chip is covered with an insulating layer (silicon dioxide or nitride) on top of which is arranged an array of metallic electrodes (gates). The voltage applied to the
electrodes subdivides each column into pixels by creating a regular array of potential wells in which the charge created by radiation interaction is stored. By properly clocking the gates, it is possible to shift the stored charge down the columns, physically separating the different charge packets from one another. This is done simultaneously for all columns in such a way that an entire row is shifted downward by one position. The bottom row of the device is an analog output shift register having the input of an amplifier on one side, usually integrated into the chip itself. The charge packets are then shifted horizontally and individually clocked out. Once this stage is terminated, the vertical shift procedure outlined previously is repeated and the successive row of charge packets is horizontally shifted, and so on, until the entire device is read out (see Fig. 4.3.5).

This “full-frame” readout mechanism has the clear advantage that the entire surface of the CCD is available for X-ray exposure. But it also has a severe drawback in that the pixels have to be processed one at a time, which dramatically affects the readout time. Although in principle clock-out frequencies of several megahertz can be used, this adversely affects the noise performance of the device—optimum results are obtained with frequencies below 500 kHz. Moreover, the smear effect (i.e., irradiation during the readout phase resulting in extra charge created in the wrong position) has to be taken into account, which in some cases makes the use of shutters necessary to prevent detector exposure during the readout phase.

A possible way of reducing the readout time is segmentation of the CCD into halves or quadrants, each one with its own amplifier, or the realization of interline or frame transfer readout configurations, in which part of the CCD area is shielded from radiation and used to store the collected charge. In this way it is possible to read out the shielded area while the unshielded area collects the successive frame. In the interline configuration, CCD columns are alternately shielded, while in the frame transfer method, half of the area is shielded and used for charge storage. Since these configurations dramatically reduce the fill factor, their use is discouraged in X-ray medical and biological imaging, although in frame transfer it is possible to expose only the unshielded area, and the system is (at least partly) buttable.

When scanned acquisition procedures are used, the problem is effectively solved by reading out the device in the TDI mode. While the beam and the detector assembly are scanned across the sample in one direction, the charge in the CCD is shifted with the same speed, but in the opposite direction.

Figure 4.3.5. Schematic of the basic CCD architecture.
In this way the collected charge is kept stationary with respect to a given projection path in the sample and is continuously accumulated until it reaches the readout region. A complementary advantage of the TDI modality is that all detector pixels in a column contribute to the final image element, thus the effect of defective detector pixels (if any) on the final image is minimized. To give an example, the SenoScan mammography system, by Fischer Imaging (http://www.fischerimaging.com), based on the pioneering work of Maidment, Fahrig, and Yaffe (1993), works in the TDI mode.

Modern CCD architectures rely on the “buried channel” arrangement, which means that the depleted region is created within the silicon bulk, well below the silicon–insulator interface (Gruner, Tate, & Eikenberry, 2002). This avoids charge trapping at the interface above, thus making the charge lifetime much longer. This is of primary importance because the charge stored in the individual pixel has to be shifted many times along the chip before being actually read out. The buried channel arrangement, together with the well capacity of modern CCDs, which can be as large as several hundred thousand electrons, makes the development of devices with outstandingly high noise performances possible (in some cases, less than 10 electrons root-mean-square [rms]).

However, it is important to stress that, because of the extremely thin depleted region resulting in an extremely small detection efficiency, CCDs are used to detect X-rays only in the indirect conversion configuration at low X-ray energies. This implies a trade-off between resolution and efficiency, as often occurs in indirect detection techniques, which might to some extent negate one of the main advantages of CCD cameras, that is, the high spatial resolution. Some possible implementations based on deep depletion devices realized in high-resistivity silicon or, even better, higher stopping power alternative semiconductors have been devised. In such cases, however, one of the primary benefits of CCD realization is lost, namely the possibility of relying on well-established fabrication schemes, because the device has to be custom designed and fabricated. Thus alternative solutions (diode arrays, pixel arrays, etc.) are often preferred.

Alternative architectures have also been explored. In pn-CCDs, for example, the MOS structures are replaced with reverse-biased pn-diode architectures. This allows for the achievement of relatively large (hundreds of cubic micrometers) depleted volumes, consequently increasing the detection efficiency for direct X-rays. Moreover, excellent noise and speed performances have been obtained (Soltau et al., 1996).

Another innovative device arising from pn-CCD technology is the controlled-drift detector (CDD), the working principle of which is somewhere between a drift chamber and a CCD. During the irradiation phase, potential wells are established in the device by superimposing a periodic perturbation on the classic linear slope of the drift potential. Thus the charge created by irradiation is stored locally in these wells. During the readout, the perturbation is removed and the collected charge packets can drift out of the active surface one after the other (see Fig. 4.3.6). Particular care must be taken in the design of the potentials in order to keep the charge packets separated from one another, including during the readout (drift) phase, but an excellent time resolution can be achieved (Castoldi et al., 2002). In some cases, CCDs have also been used as readout devices by bump-bonding them to other, higher stopping power devices such as photodiode arrays (Mainprize et al., 2002b).

4.3.2.3.2. CMOS Image Sensors

For many years considered markedly inferior to CCDs, CMOS sensors are currently experiencing a resurgence in development work, and several companies have commercialized devices for X-ray imaging applications based on this technology (see, e.g., http://www_cmosxray.com, http://www.rad-icon.com, http://www.exxim-cc.com). This work has been triggered mainly by two factors where CMOS
technology is distinctly superior to CCD technology, namely low cost and low power consumption, the latter being especially important in the framework of deep space exploration experiments. Although, initially, smaller pixel dimensions were more easily achievable with CCD technology, current developments in CMOS pixel sensors have overcome this limitation. Moreover, in the design of CMOS pixel sensors, additional functions such as analog-to-digital (A/D) conversion can be integrated on the chip. And finally, as each pixel is, in practice, read out individually, there is no signal degradation due to charge transfer.

Although substantially reduced, both readout and fixed pattern noise are still perceptibly higher than in CCDs. Furthermore, it should be noted that in CMOS sensors the fill factor is always less than 100%, although this can also be the case in some CCD designs.

Pixel implementation in CMOS technology is achieved through two approaches: the passive and active schemes. The passive scheme is based on the pioneering work of Weckler (1967) in which the charge created by X-ray interactions in a photodiode (the individual pixel) is read out when the “access” transistor connecting the pixel to a bus is activated (see Fig. 4.3.7). The bus is kept at a constant voltage by a charge integrating amplifier. When the photodiode is accessed, this charge is converted into a voltage signal on the bus. This is the simplest implementation, and the one allowing the largest fill factor, but also the one with the lower noise performance, being of the order of 250 electrons rms compared with the typical 20 electrons rms of modern CCD cameras (Fossum, 1997).

This low noise performance suggested the insertion of an amplifier directly within each pixel, opening the way to active CMOS sensors. The amplifier is activated only during the readout phase, and thus it does not substantially affect the low power consumption characteristics of the device. This enhances the system performance (an overall noise level of 75 to 100 electrons rms is achievable); on the other hand, it also reduces the maximum achievable fill factor, which in some cases can be as small as 30%, as shown in Figure 4.3.8. A detailed description of a possible design of such a device can be found in, for instance, Kleinfelder et al. (2002).

Further improvements in noise performance can be achieved by moving from the active photodiode style to an active “photogate” style. With this device, noise performances comparable to those of

![Figure 4.3.6. The potentials in the CCD during the acquisition (left) and readout phase (right) (from Castoldi et al., 1997).](image)
CCDs are obtained. A detailed description of this device is beyond the scope of this chapter, but can be found in Mendis et al. (1997).

4.3.2.3.3. Flat-Panel Detectors

Progress achieved in the field of large-area active matrix arrays, basically related to the development of liquid crystal displays, has made flat-panel detector technology available at affordable costs. This technology has the capability of providing large area, high quantum efficiency, fast readout, reliable detectors applicable to almost all radiological modalities (radiology, fluorography, fluoroscopy). Thus flat-panel detectors are rapidly becoming the new standard in radiology.

Basically, two strategies have been followed in the development of flat-panel detectors: indirect, mostly based on amorphous silicon (a-Si) (Antonuk et al., 1991; Fujieda et al., 1991); and direct (Zhao & Rowlands, 1995), mainly employing amorphous selenium (a-Se).

Indirect detection systems take advantage of the possibility provided by a-Si of realizing large area devices. Competing technologies (e.g., CCDs) are based on crystalline silicon (c-Si), in which silicon wafers, obtained by the proper cutting of c-Si grown in cylindrical shapes, are severely limited in their maximum achievable area.
Large matrices of imaging pixels arranged in regular arrays, each one consisting of a sensor (photodiode) coupled to a thin-film transistor (TFT), can be realized in a-Si. Each pixel element is connected vertically and horizontally to the electronics located on the edges of the matrix via data/control/bias lines and can thus be individually addressed and read out (see Fig. 4.3.9). In order to be used for X-ray imaging, the a-Si matrix is coated with a scintillating material to convert X-rays into visible light, which is then converted into electrons by the individual photodiodes. This coating can be realized in conventional phosphor screens such as Gd$_2$O$_2$S or in columnar CsI phosphors for improved spatial resolution (see the following section for scintillating material descriptions). For use in megavoltage (portal) imaging, a metal plate can be added to improve the detection efficiency.

It should be noted that because each individual pixel must contain both the actual sensor and the transistor(s), the fill factor is always less than one, and the problem becomes more acute as the pixel size is made smaller (e.g., in mammography applications).

The same active matrix described previously (or a simplified version of it in which the photodiodes are replaced by simple electrodes with integrated storage capacitance) can be used in direct-detection flat-panel systems. In this case a photoconductor, typically a-Se, is directly evaporated onto the matrix array. The active matrix can also be realized in other materials, for example, polycrystalline silicon or CdSe, that have higher carrier mobilities (Zhao & Rowlands, 1995). The incoming radiation is then directly converted into charge in the a-Se layer without the intermediate conversion into visible light, thus providing significant advantages. A common electrode is applied on top of the photoconductor layer to produce the electric field necessary to drift the created charge toward the electrodes.

The excellent photoconductor properties of a-Se were already well known in X-ray imaging, as, for example, in xeroradiography, which was mainly based on this material (Boag, 1973). In that case, the latent image on the photoconductor layer was read out by means of fine toner particles, resulting in an “edge enhancement” effect on the obtained images. Lately other readout schemes based, for example, on electrometer probes (Neitzel, Maack, & Gunther-Kohfahl, 1994) have been employed.

**Figure 4.3.9.** Schematic of an a-Si active matrix array. As in CMOS sensors, the transistor is integrated in each single pixel and, as a consequence, the fill factor is less than 100%.
Currently these non-real-time readout schemes have been superseded by the active matrix technology described earlier.

Both direct and indirect flat-panel detectors have proven capable of providing good performances in nearly all radiological fields, such as chest radiography (Liu & Shaw, 2004), angiography (Granfors et al., 2003), and fluoroscopy (Hunt, Tounsiignant, & Rowlands, 2004). However, some open issues are still encountered in the highly demanding field of mammography. Although both indirect (Senographe 2000D; http://www.gehealthcare.com) and direct (Selenia Digital Mammography System; http://www.hologic.com) systems have already completed the U.S. Food and Drug Administration (FDA) approval process and are currently commercially available, most clinical trials have reported a reduced sensitivity compared to screen-film mammography (Pisano & Yaffe, 2005). However, these reported differences are small and in most cases have not been considered statistically significant by the researchers. Moreover, they are usually accompanied by a slight increase in specificity. It is hoped that a much larger survey, under way at the time of writing, will be capable of providing definitive answers (Galen, Staab, & Pisano, 2002). Meanwhile, research is currently in progress to evaluate and improve the performances of flat-panel detectors in this field (Jee et al., 2003; Saunders et al., 2005).

4.3.2.4. Scintillators and Scintillation Light Readout

Strictly speaking, a scintillator is not a “complete” detector by itself: it converts X-rays, gamma rays, or charged particles into light that must then be detected by a successive device. However, there is such a wide range of different scintillating materials with different properties and for different purposes that the subject deserves to be treated in dedicated sections. Thus this section is split in two subsections, the first dealing with specific scintillator materials, and the second dealing with the devices used to read out the scintillating light. In practice, a range of possible detectors results from combining an item taken from the first subsection with one taken from the second.

The scintillator/readout combination is widespread in biomedical imaging with ionizing radiation. It is the basis of most nuclear imaging techniques, although direct detection techniques have been proposed (e.g., by Kastis et al., 2002), where a CdZnTI pixel detector is involved. Because of the relatively high-energy X-rays usually employed, a range of CT and micro-CT techniques are also based on this detection scheme. Moreover, scintillators are the basis for all X-ray imaging techniques based on indirect detection.

Specific examples related to all these applications will be given later. Here, a general overview of the most commonly used scintillating material/readout devices is provided.

4.3.2.4.1. Scintillators

The basic requirements for a scintillating material in biomedical imaging are the following:

- It must have a high density and atomic number in order to achieve the maximum detection efficiency.
- It must efficiently transmit the scintillating light in order to avoid signal losses. This is achieved by using ionic crystals, or by partly covalent crystals with a band gap;
- It must provide high light yield, because the generation of a large number of scintillation photons means maximizing the information and minimizing the noise. Thus the band gap must be small.
- In most applications, speed takes primary importance. Therefore, short decay time and reduced afterglow are also fundamental parameters.
In indirect detection X-ray imaging, and especially in screen-film imaging, the green-light emitting Gd$_2$O$_2$S:Tb (again, the dopant material is after the colon) is widely used. Another material, blue-emitting CaWO$_4$, has also been used. For some applications, YTaO$_4$ is used in an undoped form, emitting at 350 nm, while doping with Nb shifts the emission to longer wavelengths. Although $Z$ and $\rho$ are the key parameters in detection efficiency evaluation, in X-ray imaging it is of primary importance to take into account the K-edge energy of the different materials.

Obviously, the thicker the scintillating screen, the higher the detection efficiency, but this severely affects the spatial resolution because of the lateral spread of light in the scintillator. In fact, a trade-off between the two is needed, and usually a broadening of the point spread function (PSF) on the order of the scintillator thickness is encountered.

This limitation was to a large extent removed by growing the scintillating material in columns, which act almost like individual fiber optics, thus strongly limiting the lateral light spread. This is basically achievable with CsI:Tl and CsI:Na. Currently the former is more widely used because it is less hygroscopic and matches the efficiency curve of a-Si better than the latter. A typical individual CsI:Tl column is 3 $\mu$m in diameter and approximately 5 mm in height, which allows a high detection efficiency (approximately 83% at 60 keV). Current problems with CsI:Tl are its afterglow and the phenomenon of hysteresis, which is an increase in the light yield with radiation damage. Moreover, separation between the columns is obtained by cracking, and as a consequence, light channeling may not be optimized (Spekowius et al., 1995).

In nuclear imaging the most commonly used materials are NaI, Bi$_4$Ge$_3$O$_{12}$ (BGO), Gd$_2$SiO$_5$ (GSO), and Lu$_2$SiO$_5$:Ce (LSO). Almost all of these have limitations: NaI in speed and sensitivity; BGO in speed, resolution, and brightness; and GSO in brightness. Also, because of intellectual property restrictions, LSO is currently used by only one manufacturer.

Recently a new crystal called LYSO (Lu$_{2(1-x)}$Y$_x$SiO$_5$:Ce), which has characteristics very close to those of LSO, was developed and tested by a company in collaboration with the European Organization for Nuclear Research (CERN)-based collaboration CrystalClear (http://crystalclear.web.cern.ch). LYSO was also used in combination with LuAlO$_3$:Ce (LuAP) in the development of an innovative positive emission tomography (PET) scanner for small animals. Furthermore, the CrystalClear collaboration is carrying out considerable work on PbWO$_4$ (PWO), which is a very attractive material because of its relatively low production cost. However, its use is currently restricted to scientific applications because of its low light yield.

Other materials, such as YAlO$_3$:Ce (YAP:Ce), are mainly used in systems for small animal examinations. These materials have lower densities, resulting in reduced detection efficiency.

Finally, easy-to-machine ceramic scintillators such as Y$_2$O$_3$:Eu$^{3+}$, Gd$_2$O$_2$:Eu$^{3+}$ codoped with Pr, and especially Gd$_2$O$_2$:Pr codoped with Ce and F (GOS) have recently become common, particularly in CT, where they appear to be the optimal scintillating material. Because they are translucent, they are subject to resolution reduction due to lateral light scattering. However, this problem is kept under control with the use of rather thin devices, without an excessive detection efficiency reduction, as they are characterized by relatively high $\rho$ and $Z$ values. Moreover, the emitted light matches well the sensitivity curves of the photodiodes. Their main problem—the afterglow—is strongly reduced by the codoping procedure, although this also affects the scintillation efficiency. Research is also in progress on other ceramic materials such as (Lu$_2$O$_2$:Eu,Tb and SrHfO$_3$:Ce).

The main properties of the most commonly used scintillating materials are summarized in Table 4.3.1.
<table>
<thead>
<tr>
<th>Material</th>
<th>Density (g/cm³)</th>
<th>$\rho Z$ (10⁶)</th>
<th>Attenuation Length at 511 keV (mm/prob. phot. eff.)</th>
<th>Hygroscopicity</th>
<th>Light Yield (photons/MeV)</th>
<th>Afterglow (% after 3 ms/100 ms)</th>
<th>Decay Time (ns)</th>
<th>Emission λ max. (nm)</th>
<th>DE/E (%), PMT Readout</th>
</tr>
</thead>
<tbody>
<tr>
<td>CsI:Na</td>
<td>4.51</td>
<td>38</td>
<td>22.9/21</td>
<td>Yes</td>
<td>40,000</td>
<td>800 – 6 × 10⁻⁶</td>
<td>630</td>
<td></td>
<td>420</td>
</tr>
<tr>
<td>CsI:Tl</td>
<td>4.51</td>
<td>38</td>
<td>22.9/21</td>
<td>Slightly</td>
<td>66,000</td>
<td>&gt;6 × 10⁻⁶</td>
<td>2/03</td>
<td>550</td>
<td>6.6 (PMT) / 4.3 (SDD)</td>
</tr>
<tr>
<td>CaWO₄</td>
<td>6.1</td>
<td>89</td>
<td>13.6/32</td>
<td>No</td>
<td>20,000b</td>
<td></td>
<td>420</td>
<td></td>
<td>IM</td>
</tr>
<tr>
<td>YTaO₃: Nb</td>
<td>7.5</td>
<td>96</td>
<td>11.8/29</td>
<td>No</td>
<td>40,000b</td>
<td></td>
<td>410</td>
<td></td>
<td>IM</td>
</tr>
<tr>
<td>Gd₂O₃:S:Tb</td>
<td>7.3</td>
<td>103</td>
<td>12.7/27</td>
<td>No</td>
<td>60,000b</td>
<td>1 × 10⁻⁶</td>
<td>545</td>
<td></td>
<td>IM</td>
</tr>
<tr>
<td>Gd₂O₃:S:Pr,Ce,F</td>
<td>7.3</td>
<td>103</td>
<td>12.7/27</td>
<td>No</td>
<td>35,000b</td>
<td>4 × 10³</td>
<td>510</td>
<td></td>
<td>IM</td>
</tr>
<tr>
<td>Gd₂O₃:S:Pr(UFC)</td>
<td>7.3</td>
<td>103</td>
<td>12.7/27</td>
<td>No</td>
<td>50,000b</td>
<td>3 × 10³</td>
<td>510</td>
<td></td>
<td>IM</td>
</tr>
<tr>
<td>Y₁.₃₄Gd₀.₆₆O₃:Eu,Pr</td>
<td>5.9</td>
<td>44</td>
<td>17.8/16</td>
<td>No</td>
<td>42,000b</td>
<td>1 × 10⁶</td>
<td>610</td>
<td></td>
<td>IM</td>
</tr>
<tr>
<td>Gd₅Ga₂O₁₅:Cr,Ce</td>
<td>7.1</td>
<td>58</td>
<td>14.8/18</td>
<td>No</td>
<td>40,000b</td>
<td>140 × 10³</td>
<td>730</td>
<td></td>
<td>IM</td>
</tr>
<tr>
<td>CdWO₄</td>
<td>7.9</td>
<td>134</td>
<td>11.1/29</td>
<td>No</td>
<td>20,000b</td>
<td>5 × 10³</td>
<td>495</td>
<td></td>
<td>6.8</td>
</tr>
<tr>
<td>Lu₂O₃:Eu,Tb</td>
<td>9.4</td>
<td>211</td>
<td>8.7/35</td>
<td>No</td>
<td>30,000b</td>
<td>&gt;10⁶</td>
<td>611</td>
<td></td>
<td>IM</td>
</tr>
<tr>
<td>CaHfO₃:Ce</td>
<td>7.5</td>
<td>139</td>
<td>11.6/30</td>
<td>No</td>
<td>=10,000b</td>
<td>40</td>
<td>390</td>
<td></td>
<td>IM</td>
</tr>
<tr>
<td>SrHfO₃:Ce</td>
<td>7.7</td>
<td>122</td>
<td>11.5/28</td>
<td>No</td>
<td>=20,000b</td>
<td>40</td>
<td>390</td>
<td></td>
<td>IM</td>
</tr>
<tr>
<td>BaHfO₃:Ce</td>
<td>8.4</td>
<td>142</td>
<td>10.6/30</td>
<td>No</td>
<td>=10,000b</td>
<td>25</td>
<td>400</td>
<td></td>
<td>IM</td>
</tr>
<tr>
<td>NaI:Tl</td>
<td>3.67</td>
<td>24.5</td>
<td>29.1/17</td>
<td>Yes</td>
<td>41,000</td>
<td>230</td>
<td>410</td>
<td></td>
<td>5.6</td>
</tr>
<tr>
<td>LaCl₃:Ce</td>
<td>3.86</td>
<td>23.2</td>
<td>27.8/14</td>
<td>Yes</td>
<td>46,000</td>
<td>25 (65%)</td>
<td>330</td>
<td></td>
<td>3.3</td>
</tr>
<tr>
<td>LaBr₃:Ce</td>
<td>5.3</td>
<td>25.6</td>
<td>21.3/13</td>
<td>Yes</td>
<td>61,000</td>
<td>35 (90%)</td>
<td>358</td>
<td></td>
<td>2.9</td>
</tr>
</tbody>
</table>

*continued on next page*
Table 4.3.1. Physical characteristics of some scintillator materials (continued)

<table>
<thead>
<tr>
<th>MATERIAL</th>
<th>DENSITY (g/cm³)</th>
<th>ρZₑₓ (10⁶)</th>
<th>ATTENUATION LENGTH AT 511 keV (mm/ prob. phot. eff.)</th>
<th>HYGROSOPICITY</th>
<th>LIGHT YIELD (photons/MeV)</th>
<th>DECAY TIME (ns)</th>
<th>AFTERGLOW (% after 3 ms/100 ms)</th>
<th>EMIS-SION λ max. (nm)</th>
<th>DE/E (FWHM) (%)</th>
<th>PMT READOUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bi₄Ge₃O₁₂ (BGO)</td>
<td>7.1</td>
<td>227</td>
<td>10.4/40</td>
<td>No</td>
<td>9000</td>
<td>300</td>
<td>480</td>
<td>9000</td>
<td>4.80</td>
<td>9.0</td>
</tr>
<tr>
<td>Lu₂SiO₅:Ce (LSO)</td>
<td>7.4</td>
<td>143</td>
<td>11.4/32</td>
<td>No</td>
<td>26,000</td>
<td>40</td>
<td>420</td>
<td>9000</td>
<td>4.20</td>
<td>7.9</td>
</tr>
<tr>
<td>Gd₂SiO₅:Ce (GSO)</td>
<td>6.7</td>
<td>84</td>
<td>14.1/25</td>
<td>No</td>
<td>8000</td>
<td>60</td>
<td>440</td>
<td>9000</td>
<td>4.40</td>
<td>7.8</td>
</tr>
<tr>
<td>YAlO₃:Ce (YAP)</td>
<td>5.5</td>
<td>7</td>
<td>21.3/4.2</td>
<td>No</td>
<td>21,000</td>
<td>30</td>
<td>350</td>
<td>9000</td>
<td>4.30</td>
<td>4.3</td>
</tr>
<tr>
<td>LuAlO₃:Ce (LuAP)</td>
<td>8.3</td>
<td>148</td>
<td>10.5/30</td>
<td>No</td>
<td>12,000</td>
<td>18</td>
<td>365</td>
<td>9000</td>
<td>4.30</td>
<td>15</td>
</tr>
<tr>
<td>Lu₂Si₂O₇:Ce (LPS)</td>
<td>6.2</td>
<td>103</td>
<td>14.1/29</td>
<td>No</td>
<td>30,000</td>
<td>30</td>
<td>380</td>
<td>9000</td>
<td>4.30</td>
<td>≈10</td>
</tr>
</tbody>
</table>

a Measured with a PMT and silicon drift detector (SDD).
b Measured at approximately 60 keV to 80 keV; all others 662 keV.
c Proprietary codopant.
IM, integrating mode.
van Eijk, 2002.
4.3.2.4.2. Scintillator Output Readout

The most common device used to convert the feeble light pulse produced by a scintillator into an easily detectable signal is the photomultiplier tube. This is basically an evacuated tube with a photocathode at the top. This converts the light produced by the scintillator into photoelectrons (typically a few hundred). Focusing electrodes drive these electrons to a series of dynodes that provide the actual multiplying stages. Usually the interdynode potential is on the order of a few hundred volts and the multiplication factor (the number of secondary electrons emitted per primary incident electron) of each dynode ranges from between 4 and 6 to about 10 in optimal conditions. In this way, the few hundred electrons are converted into approximately $10^7$ to $10^{10}$ charge carriers, which are then collected at the anode (see Fig. 4.3.10).

One of the main characteristics of photomultiplier tubes is their speed. If illuminated by an appropriately short light pulse, they produce an electron pulse a few nanoseconds long within a delay time of a few tens of nanoseconds. This time interval can be made one order of magnitude smaller when continuous channels, rather than multiplying structures based on dynodes, are used (De Vries & van Eijk, 1985). In this device, the multiplication effect is caused by the electrons hitting the wall of the tube rather than a series of discrete dynodes.

The classic design of a photomultiplier tube, in which all electrons generated by the photocathode are driven to the same chain of dynodes by the focusing electrodes, prevents any possible return to the original position in which the light hit the photocathode surface (and thus the position in which the original X-ray or gamma ray hit the scintillator, creating that light). On the other hand, many applications, and nuclear medicine in particular, need this information. For this reason, position sensitive photomultipliers have been developed, in which a multiplying structure capable of preserving the spatial separation of charge clouds generated by different regions of the photocathode is introduced.

The simplest solution is achieved by using continuous channels with very small diameters (tens of micrometers), arranging them in clusters to form what is called a multichannel plate (see Fig. 4.3.11). More sophisticated solutions involve the development of fine mesh structures in which each layer of horizontally displaced multiple dynodes is followed by a layer in which multiple holes are arranged to

Figure 4.3.10. Schematic of a photomultiplier tube.
prevent the different charge clouds from mixing, or arrangement of the dynodes themselves in channel structures. Although in most cases these structures are capable of confining the electron avalanches properly, a nonuniform response is often observed with respect to light hitting different positions on the photocathode surface. This type of device has been successfully implemented in a range of nuclear imaging applications (Cherry et al., 1997; Meikle et al., 2002; Miyaoka, Kohlmyer, & Lewellen, 2001; Weisenberger et al., 1998).

A possible replacement for photomultiplier tubes is provided by silicon photodiodes. In general these are cheaper, more compact, and consume less power than photomultiplier tubes. Moreover, they have a broader spectral response extending to longer wavelengths, which can be important when scintillators with a significant yield at long wavelengths (CsI:Tl, BGO) are used.

Because they have no inherent multiplying effect, the number of electron-hole pairs created is at maximum equal to the number of impinging scintillation photons. Thus their use in pulse mode usually results in lower energy resolution compared to that provided by photomultiplier tubes. On the other hand, their use in current mode is much more reliable, especially for high-rate applications. This has made them the light detector of choice in many CT applications based on the use of scintillators.

In order to operate silicon photodiodes in pulsed mode, the system can be cooled to reduce the leakage current, but consequently increasing the SNR. Another possibility for reducing the leakage current consists of employing materials with a wider band gap, such as Hgl₂ (Wang, Patt, & Iwanczyk, 1996).

Alternatively, it is possible to increase the collected charge rather than to reduce the leakage current. This may be achieved by using avalanche photodiodes. In these devices, a high applied voltage accelerates the charges to energies large enough to produce additional electron-hole pairs. However, good temperature and voltage stability is required because the gain factor is strongly dependent on these parameters.

The most common configuration is known as “reach-through,” where the electric field is shaped in such a way that it increases slowly along the drift region (a large fraction of the silicon bulk) and then suddenly jumps to a much higher value in the region immediately adjacent to the collection electrode (see Fig. 4.3.12). In this way, the multiplication takes place in the last part of the device, as far as possible from the entrance window. This reduces the lateral spread of the charge and makes the system more controllable. Several nuclear imaging instruments employing these devices have been devised (Lecomte et al., 1996; Ziegler et al., 2001).
A form of intermediate solution between those outlined previously is provided by hybrid photomultiplier tubes (sometimes called hybrid photodiodes). In these devices, a silicon detector instead of a dynode structure is used to multiply the electrons. Its construction is rather delicate: the silicon device has to be inserted into the vacuum tube, with consequent risk of contaminating the photocathode (making it inefficient) or the surface of the silicon chip itself (thus increasing the surface leakage current). Nevertheless, the number of produced charge carriers per incident electron is much higher (approximately 3000) than in conventional photomultiplier tube dynodes (typically 5, maximum 25 for specific high-gain, negative electron affinity dynode materials), and this results in increased spectral resolution. Moreover, a position-sensitive silicon detector can be used, in which case a position-sensitive hybrid photomultiplier tube results.

Gas detectors, like multiwire proportional chambers (MWPCs) have also been proposed for possible coupling with scintillators. They can actually provide a much cheaper solution (about one-third of the cost) with respect to photomultiplier tubes. Usually they have to be coupled to a crystal that produces short wavelength photons (BaF, for instance, where $\lambda = 200$ nm) and filled with a specific gas mixture capable of converting the short wavelength photon into a photoelectron. Another advantage is that large MWPCs can be easily constructed to cover remarkably large areas.

### 4.3.3. SPECIFIC APPLICATIONS

This section provides a brief and necessarily incomplete overview of a range of applications with the main goal of stressing in each case the specific detector performance requirements and providing a few examples of the most commonly used devices.
4.3.3.1. Computed Tomography

In CT, relatively high (80 kVp to 140 kVp) X-ray energies are used. Consequently X-ray detection is largely based on inorganic scintillators, classically CdWO$_4$, BGO, and CsI:Tl. The most common scintillating light readout scheme is based on arrays of photodiodes operated in the integration mode.

However, these scintillators have several disadvantages: relatively low light yields, imperfect matching with the sensitivity curve of the photodiodes, relatively long afterglow, sensitivity to radiation damage, and toxicity (CdWO$_4$). Therefore alternative solutions have been found. For example, one of the Siemens (SOMATON) CT scanners employed high-Z gas (xenon) kept at high pressure (van Eijk, 2002), but a loss in detection efficiency was encountered. Another development was provided by innovative ceramic technologies involving the introduction of codopants to reduce the afterglow: Y$_2$O$_3$:Eu$^{3+}$ and Gd$_2$O$_3$:Eu$^{3+}$ codoped with Pr, Gd$_2$O$_2$S:Pr codoped with Ce and F (Grabmaier & Rossner, 1993; Greskovitch & Duclos, 1997). As a consequence, ultrafast ceramic scintillating detectors rapidly became the standard in CT. Moreover, some direct detection methods have also been proposed, in some cases relying on the photon counting readout mode (Pani et al., 2004; Shikhaliev, Xu, & Molloi, 2005).

4.3.3.2. Mammography

Mammography is probably one of the most challenging radiological fields, as both high spatial resolution and high detection efficiency—two classically counteracting issues—are of primary importance. Although classic screen-film mammography is a precise and reliable technique, it has some shortcomings in terms of sensitivity and specificity. These might be overcome by higher image contrast (which provided the first incentive toward digital radiology) and higher resolution. However, increased spatial resolution often implies reduced detection efficiency, which is a major problem in this field because of the high radiosensitivity of the breast. Several asymmetric double-screen–double-emulsion combinations have been tested (Haus, 1990) and were proven to provide a dose reduction of up to 50%, but at the price of a reduced spatial resolution.

A further major incentive toward digital mammography arises from dynamic range requirements: according to Maidment, Fahrig, and Yaffe (1993), the optimal latitude should be 1:100, while the screen-film system is capable of providing a range of only 1:25.

The optimal pixel size for mammography is a rather controversial issue (Chan et al., 1994); at present, 50 μm is the usually chosen compromise. The first step into digital mammography was provided by photostimulable phosphor systems (Hillen, Schiebel, & Zaengel, 1987). Some of the advantages were immediately clear, including dynamic range, windowing and contrast enhancement, and image manipulation, in addition to some of the limitations, such as spatial resolution.

Flat-panel detectors are rapidly becoming the new standard. However, as mentioned previously, they show reduced sensitivity (although often accompanied by improved specificity) with respect to the classic screen-film technique (Pisano & Yaffe, 2005).

Charge-coupled device detectors are good candidates because of their high spatial resolution, but their relatively small area requires the use of tiling techniques. Because it is not easy to arrange tiled CCDs to cover the required area of 18 cm × 24 cm, slot systems are usually preferred (the Fisher system, based on CCDs operated in the TDI mode, has already been mentioned; a description can be found in Tesic, Fisher Piccaro, & Munier, 1999). Smaller systems can be used for stereotactic breast biopsy (Karellas et al., 1992; Roehrig, Fajardo, & Yu, 1993). A completely different solution, based
on single photon counting edge-on microstrip detectors, has more recently been made commercially available (see section 4.3.2.3 regarding the SECTRA system).

4.3.3.3. Chest Radiography
The classic screen-film combination, with Gd₂O₂S:Tb as the typical scintillator material (Blasse & Grabmaier, 1994) and standard dimension of 35 cm × 43 cm, is still widely used in this application. The first digital approach to chest radiography was provided by storage phosphor systems, usually based on BaFBr₁₋ₓIx:Eu (x ≤ 0.2) as the active material (Sonoda et al., 1983). In the early 1990s, a solution based on the electrostatic readout of a drum covered with amorphous selenium was introduced by Neitzel, Maack, and Gunther-Kohfahl (1994), which may be regarded as anticipatory of current flat-panel technologies.

More recently, flat-panel technology has received widespread acceptance in this field, employing direct conversion in an a-Se readout by thin-film transistor arrays (Eastman Kodak) as well as indirect conversion obtained by depositing a scintillating layer of Gd₂O₂S:Tb on top of an array of a-Si photodiodes, again with readout by thin-film transistors (Agfa). The typical pixel size of these devices is about 150 μm, which is considered sufficient for this kind of imaging. Other companies (Canon, GE, Varian, Trixell) rely on the use of CsI:Tl as the scintillating material because, as already mentioned, it can be grown in columnar structures providing reduced lateral light spread and thus increased spatial resolution (Jing et al., 1994; Zhao, Ristic, & Rowlands, 2004). Experimentation based on scanned systems, in some cases employing the TDI principle, has also been carried out.

4.3.3.4. Dental Imaging
Although dental imaging is not one of the most demanding fields in radiology, significant research has been carried out in this field. Here, small (2 cm × 3 cm or 3 cm × 4 cm) active area detectors are required. Originally the task was easily accomplished using the classic intraoral film package. A small intraoral storage phosphor system (van der Stelt, 2001) may be used, as well as direct X-ray detection in small CCDs.

4.3.3.5. Fluoroscopy
The X-ray image intensifier has for many years been the workhorse in fluoroscopy, and is still widely used (Hell, Knupfer, & Mattern, 2000). Its main advantages and limitations have already been discussed, but in the context of its everyday use, a further drawback is its rather bulky nature, which can make it quite impractical in some cases. Flat-panel technology is currently encountering wide acceptance in this field, and large (40 cm × 40 cm) systems based on the CsI:Tl/a-Si indirect detection modality have recently been developed.

4.3.3.6. Single Photon Emission Computed Tomography
Single photon emission computed tomography (SPECT) is a technique completely different from the ones discussed up to this point. It is based on the detection of radiation emitted from the human body in the form of high-energy gamma rays rather than X-rays. Its nontomographic version, often called scintigraphy, is based on a simplified version of the instrumentation discussed in the following paragraphs, and hence it will not be treated separately.

The first important factor in SPECT is that resolution and SNR are mostly driven by a collimator system rather than by the detection device itself. Planar collimators are used in scintigraphy, and
conical diverging ones are the standard for SPECT. In the latter, multiple camera heads are often used to increase the sensitivity, and their rotation around the patient provides the 3-D information. As in almost all other imaging fields, the sensitivity and resolution are counteracting issues. Pinhole collimators provide maximum resolution, but minimum sensitivity and field of view, and a range of solutions has been proposed, including multiple and coded-aperture pinholes. The problem is less dramatic when lower energy X-ray photons (for instance, 27 keV to 35 keV as emitted by $^{125}$I) are used. In this case, reabsorption within the human body becomes a limiting issue; in practice, good results are obtained only in small animal systems.

In the classical arrangement, monolithic NaI:Tl crystal plates are read out by photomultipliers arranged in hexagonal packing schemes (Jasczak, Coleman, & Lim, 1980; Short, 1984). More recent systems are arranged in rectangular packing with thicker (25 mm versus 6 mm to 12 mm) NaI:Tl crystals.

Recently, new materials have been introduced, notably LaCl$_3$:Ce and LaBr$_3$:Ce. These are capable of providing high light yield, fast response, good time resolution, excellent energy resolution (because of the small nonproportionality of the emitted light to the absorbed gamma-ray energy), while maintaining a detection efficiency comparable to that of NaI:Tl.

4.3.3.7. Positron Emission Tomography (PET)

PET differs from SPECT in that positron ($\beta$)-emitting radioactive tracers are inserted into the patient. After a short distance in the tissue, the emitted positron is annihilated with an electron and two 511 keV back-to-back gamma rays are emitted.

The back-to-back nature of the emission is advantageous in that collimation is not required because position information is retrieved by the coincident detection of the two photons. The system has a high sensitivity ($10^2$ to $10^3$ higher than SPECT), but a series of key technical challenges have to be faced. Depth of interaction encoding (i.e., implementing systems capable of detecting at which depth within the scintillator crystal the gamma-ray interaction has taken place) is highly desirable in order to avoid parallax errors limiting the spatial resolution. Furthermore, with high-density materials, 511 keV photons have a high likelihood of being Compton scattered inside the detector, and being absorbed in a neighboring detector element, thus affecting the resolution.

Unlike SPECT, PET has at least two unavoidable intrinsic limits. The first is the range of the positron before annihilation (the image is actually a map of the $\beta$ annihilation positions rather than the emission centers). The second is a small noncollinearity ($\pm 0.25^\circ$) in the back-to-back emission due to the residual momentum of the positron or the electron. These limits can only be partially compensated by means of correction algorithms.

Bi$_4$Ge$_3$O$_{12}$ is the most commonly used scintillator because of the relatively high probability (40%) of photoelectric absorption of 511 keV photons. NaI:Tl is also used because it allows large, curved crystals to be produced. However, the detection efficiency is lower than for BGO.

Considerable research has been carried out on new scintillating materials because BGO has non-optimal light yield, speed, and energy resolution. The introduction of LSO removed some of these limitations, but as previously mentioned, its use is currently limited to only one manufacturer because of intellectual property rights. GSO is also a good candidate for solving these problems but suffers from low light yield. Further solutions based on scintillating crystals are mentioned in section 4.3.4.1. It is worth noting that alternative detection methods have also been proposed, employing, for example, liquid xenon as a scintillator and gas chambers to read the signal (Chepel et al., 1997; Collot, Jan,
Finally, in small animal examinations, the use of lower density materials like YbO$_2$:Ce (YAP:Ce) is often proposed (Del Guerra et al., 2000; Weber et al., 2000).

### 4.3.3.8. Multimodal Imaging

The term “multimodal” refers to the combined use of at least two different imaging modalities. The main aim is to provide functional and morphological information at the same time. Originally this was achieved by PET/SPECT (functional) examination, quickly followed and perhaps anticipated by NMR scans (morphological). However, this approach works only in an extremely limited number of cases (mainly some applications in brain imaging), because for most other organs the results are affected by physiological motion (heartbeat, respiratory cycle, bladder filling/emptying, etc.). In this framework, optimal results are achieved by combined SPECT/CT or PET/CT scans, although problems arise from the relevant differences in the acquisition time of the two approaches. When the SPECT/CT combination is used, further complications arise because of the smaller difference in energy between the X-rays employed by the two techniques.

### 4.3.3.9. Molecular Imaging

Molecular imaging is quickly gaining predominance because of our increasing knowledge of the roles of specific genes and proteins in the development of various diseases. The adjective “molecular,” however, does not imply new kinds of imaging techniques: it arises principally from a shift in the emphasis given to the information extracted from the examination (Cherry, 2004). For ionizing radiation, the techniques employed are in fact principally PET and SPECT. Since it provides solely morphological information, CT can play only a “supporting” role in terms of providing high-resolution structural information for the interpretation and correction of data provided by other techniques (which can be SPECT and PET, but also other imaging techniques not based on ionizing radiation). Actually, in some cases, a further role of CT can be the assessment of changes in vascularization associated with or caused by a specific drug. However, the approach is to a large extent limited to the imaging of small animals, because in practice it requires very high resolution, in vivo micro-CT. In this framework, some high-efficiency detectors based on PbI and HgI$_2$ have been devised (Street et al., 2002). The problem of detection efficiency for small animal micro-CT is not trivial, as the dose levels associated with current micro-CT techniques can kill a mouse after about ten scans, while in many cases the primary aim is to follow up the effects of a specific drug on the same animal over a relatively long time span.

The main requirements of molecular imaging can be summarized as high resolution, high sensitivity, and high accuracy in the extraction of quantitative information. The current goal is to be able to image approximately $10^9$ cells, something about the size of a grain of rice. Since it is a technique that is used to a large extent in small animal research, trade-offs can be achieved in resolution and sensitivity, which are the two classically counteracting issues in nuclear imaging.

### 4.3.3.10. Portal Imaging

Portal imaging refers to the acquisition of images during radiotherapy treatments, making use of the treatment beam itself. Because of the high energies employed, the conditions are far from optimal for imaging because image contrast is strongly suppressed. Nevertheless, this technique is absolutely essential to the quality of the treatment.

Two modalities must be distinguished: localization and verification imaging. In localization, a small portion (5% to 10%) of the total dose is delivered to the patient. This allows the checking of
patient positioning as well as possible corrections before the main dose is delivered. Verification imaging, on the other hand, provides a way to assess the treatment itself.

Like practically all imaging methods, portal imaging has been based on films. Originally the film was sandwiched between two metal plates in which the high-energy, high-power treatment beam created a number of electrons that impinged upon the film. The back plate, sometimes made of plastic, was used to take advantage of the backscattered electrons. A first improvement was achieved with the so-called enhanced contrast localization system, in which a low-speed, fine-grain film was sandwiched between two phosphor screens, with a thin copper plate on top.

Such systems provide an image quality basically fulfilling the requirements of the technique. In this case, rather than being driven by image quality issues, the quest for digital devices was mainly driven by other factors, notably the necessity of processing the film, which means that several minutes are required before the image is actually accessible (Antonuk, 2002). In some cases this can make the localization image useless due to patient motion. Furthermore, a real-time imaging technique would make online adjustments of the patient position feasible during the verification stage. Finally, the overall amount of time required by film imaging techniques reduces the overall number of treatments per day, imposing a limit on the maximum number of verification images taken of the same patient during the treatment. In addition, there are the usual advantages associated with digital imaging (i.e., dynamic range, image processing, etc.). Specifically designed flat-panel detectors for megavoltage imaging can essentially fulfill all these requirements (see, for instance, Antonuk et al., 1998).

4.3.3.11. Autoradiography

Autoradiography allows the visualization of the radioactivity distribution within a (thin) sample. This is the only example discussed in this chapter in which not only electromagnetic radiation, but also charged particles (β) are detected. In the classical scheme, a high-sensitivity film and an intensifier screen are placed on top of the sample; sometimes a second screen is placed below the sample to increase the sensitivity, although at the price of reduced spatial resolution. The film is then read via microdensitometers (although alternative readout solutions, usually based on CCD cameras, have been proposed, as for instance by Lear, Plotnick, & Rumley, 1987). In the early 1960s, the use of the electron microscope was also proposed to achieve an extremely high resolution on particularly small and thin samples (Caro & Van Tubergen, 1962).

However, it is clear that in this application the use of intrinsically digital detector devices can result in considerable benefits, especially in terms of sensitivity (and consequently reduced exposure times). A range of possible solutions has been proposed, beginning with storage phosphor systems (Johnson, Pickett, & Barker, 1990) and including a number of different devices. For instance, the use of MWPC (Angelini et al., 1988), double-sided microstrip detectors (Bertolucci et al., 1996), detectors based on microchannel plates (Lees, Fraser, & Dinsdale, 1997), and CCDs (Ott, MacDonald, & Wells, 2000) have all been proposed. More recently, use of the MEDIPIX hybrid pixel detector has also been proposed for autoradiography (Mettivier, Montesi, & Russo, 2003). Georges Charpak developed and commercialized an autoradiography system based on a parallel plate avalanche chamber (Biospace Instruments, http://www.biospace.fr). The same company commercialized the “Micro-Imager,” based on a scintillating foil connected to a CCD via an image intensifier, still widely used in the screening of gene expression (see, for instance, Salin et al., 2002).

The limited number of examples presented is probably sufficient to illustrate that devices suited for specific autoradiography applications can be obtained from many of the existing detector technologies.
The fact that inanimate samples are being imaged makes this application somewhat less demanding than most radiological ones, although a number of technical problems, especially regarding sensitivity and resolution, still have to be addressed.

4.4. SENSORS FOR SPECTROSCOPY

4.4.1. SPECIAL REQUIREMENTS FOR SPECTROSCOPY

Accurate spectral measurement requires a high degree of accuracy in the measurement of the charge deposited. For a radiation sensor, the degree of accuracy for spectral measurements is typically quoted in terms of the energy resolution, which is defined as the ability of the sensor to resolve two proximal energies. Consider a detector irradiated with a monoenergetic radiation beam and in which all interactions deposit the full amount of energy. In a perfect system, all charge pulses generated will produce the same pulse height, but in practice this is not the case, and a distribution of events is observed, as shown in Figure 4.4.1. As a general rule, a detector can resolve two energies if they are separated by more than one value of full width at half maximum (FWHM).

A variety of factors can limit the achievable accuracy, including random noise within the detector or electronics, but typically radiation detectors are quantum limited, meaning that the dominant cause of error is the statistical variation in the amount of charge generated for different events. The energy required for ionization is not a constant amount for every interaction; the ionization potential quoted earlier is the mean amount of energy to create an ion pair. Therefore a Gaussian distribution of pulse heights will be observed. The random nature of radiation means that Poisson statistics govern these interactions. This means that the degree of variation observed in the charge generated is proportional to the square root of the mean number of the charge, and therefore is inversely proportional to the energy resolution component due purely to statistical fluctuations. Therefore, for a given energy of radiation, the most accurate sensor for spectroscopy in a quantum-limited environment will be the one with the smallest ionization potential, as this generates the greatest amount of charge per kiloelectronvolt of energy deposited.

![Figure 4.4.1. Definition of energy resolution.](image-url)
In practice, the observed energy distribution is not quite Gaussian because other factors contribute to imperfections in the system. Common examples are impurities in the detector material restricting charge flow or random noise from slight defects in detector construction or from the controlling electronics (see Knoll, 2000, for further details). The FWHM of the individual sources of error add in quadrature. This departure of the measured fluctuations in the detector from Poisson statistics is quantified by the Fano factor, $F$. This is defined as

$$F = \frac{\text{observed variance in } N}{\text{Poisson variance}}$$

and leads to an increase in the observed energy resolution $R$ as follows:

$$R \propto \sqrt{\frac{F}{N}}.$$

In a scintillation detector, the mechanism for achieving good energy resolution is slightly more complicated than in a gas or a semiconductor, there being a two-stage procedure: (1) the conversion of ionizing radiation to light, and (2) the conversion of light to an electrical signal. The efficiency of both of these stages dictates spectroscopic performance. The conversion efficiency is typically quantified by the light yield, that is, the number of light photons produced per kiloelectronvolt of energy. On the other hand, the conversion of light to an electrical signal depends upon a number of factors (discussed in section 4.3.2.4). A further key point for spectroscopic work using scintillators is the linearity of light production in the scintillator. Ideally the light yield should be proportional to the deposited energy for a wide energy range. In practice, some degree of nonlinearity is observed, which is more marked in some crystals than others. For further details see Moszynski (2003).

### 4.4.2. SELECTING A DETECTOR FOR SPECTROSCOPIC MEASUREMENTS

When selecting an appropriate detector system for spectroscopy the following questions should be considered:

- What energy and type of radiation will be measured?
- What level of energy resolution is acceptable?
- What fluence rate will be encountered at the detector?
- What other detector properties are also required, such as ability to form an image, high detection efficiency, coincidence/timing information, etc.?
- What are the practical considerations; for example, are there any size restrictions or cost limitations?

Generally, all but the first question have conflicting answers, so some compromise must be reached. No detector system is perfect in all aspects, thus it is important to quantify the required parameters and assign an order of priority. For example, the following sections will show that detectors with the best energy resolution require a cooling system (electronic or cryogenic) and thus can be quite bulky. In addition, achieving the highest energy resolution with such a system imposes constraints on the pulse processing electronics, which in turn limits the count rate that can be accurately measured. Also, requirements such as high imaging spatial resolution only further complicate matters. Thus compromise is necessary.
4.4.2.1. **Practical Requirements for Accurate Spectral Measurement**

Once a detector has been selected it is important to ensure that it is operated in a way that achieves optimal spectroscopic performance. This is largely governed by the pulse processing electronics. A summary of the relevant details is presented here; for further details, see Knoll (2000).

The overall purpose of the pulse processing electronics is to accurately measure the voltage pulse produced by an event in the detector. Figure 4.4.2 shows a typical spectroscopic system configuration.

### 4.4.2.1.1. Amplification

The purpose of the amplification stage is threefold: (1) to amplify the detector pulse so that its amplitude can be accurately measured (usually from a few millivolts to within the 0.1 V to 10 V range); (2) to shape the pulse in order to broaden the peak and thus facilitate pulse height analysis; and (3) to shape the pulse in order to eliminate slow exponential decay of the pulse to the baseline. Point 3 is necessary to ease constraints on the event rate within the detector. Many pulse-shaping amplifiers include baseline restorer circuitry to ensure that each pulse starts at a fixed level.

In many cases, such as germanium detectors, the detector pulse has a very small amplitude and thus some preamplification of the generated pulse is required. However, in detectors such as NaI(Tl)–PMT scintillation detectors, the pulse is sufficiently large due to intrinsic amplification.

For optimal energy resolution, a linear pulse-shaping amplifier is required. A user-defined gain setting allows the amplifier to be used for a range of source energies. A variety of pulse shapes are employed; for spectroscopic work, a semi-Gaussian pulse shape is often encountered (Fig. 4.4.3). If event timing information is also required, a bipolar output is useful, with accurate timing information obtained from the baseline crossover point, although some compromise in the energy resolution is observed.

A pulse-shaping amplifier integrates the detector pulse for a given duration, set by the pulse-shaping time (often user defined). For an accurate energy measurement this time should be long enough to ensure collection of the entire pulse. Too short a time will be detrimental to the energy resolution; for example, sodium iodide was measured as having an energy resolution of 3.8% at 662 keV with a long shaping time of 50 μs, compared to only 5.9% at a much shorter time of 1.2 μs (Moszynski, 2003). Shaping time requirements are detector dependent and are determined by either the charge collection time at the electrodes or the speed of the preamplifier. A germanium detector, for example, achieves the best spectroscopic performance with a shaping time of approximately 6 μs, whereas a NaI(Tl)–PMT combination requires approximately 2 μs.

![Figure 4.4.2. A typical spectroscopic system configuration with a demonstration of the pulses at each stage.](image-url)
4.4.2.1.2. Multichannel Pulse Height Analysis

The analog-shaped pulse from the amplifier is passed through an analog-to-digital (A/D) converter, sampled, and digitized. The sampling frequency of the A/D converter needs to be sufficiently high to ensure an accurate representation of the analog trace is obtained. From the digital trace, pulse height information can be measured, histogrammed into multiple channels (energy bins), and stored to a memory buffer.

4.4.2.1.3. Digital Signal Processing Units

Digital signal processing (DSP) units are available (e.g., the DSPEC from ORTEC) that can replace the analog approach described. The preamplified signal is sampled and digitized. The digitized trace can then be processed and analyzed. Such systems tend to offer greater flexibility in terms of signal processing than is generally available with an analog unit.

4.4.2.1.4. Dead Time and Pulse Pile-up

A direct relationship between the measured pulse height and the energy deposited in the detector relies on the voltage across the electrodes being at zero immediately prior to the pulse. This is not the case if the tail of a previous pulse has not yet decayed to the baseline, as illustrated in Figure 4.4.4. This situation, known as pulse pile-up, causes the measured energy to appear greater than the true value and occurs when the event rate in the detector is too high. The outcome is to effectively impose a maximum usable count rate for a detector.

The time required to process the pulse is quantified by the dead time and is usually quoted as a percentage of the real time. Real time is composed of detector live time, during which the detector can actively process a pulse, plus the dead time. Charge drift times in gas or semiconductor detectors are generally negligible compared to electronics effects. However, afterglow effects in certain scintillators can be significant. It is generally recommended that the dead time be less than 10% for optimal energy resolution. This corresponds to a maximum count rate of approximately 33,000 counts per second, for a pulse-shaping time of 3 μs. The maximum count rate decreases proportionately for longer pulse-shaping times.

Pulse pile-up rejection circuitry, in which two amplifiers are employed, is available to help minimize this problem. A fast pulse-shaping amplifier with a short shaping time allows adjacent pulses

![Figure 4.4.3. A semi-Gaussian pulse shape for good energy resolution (unipolar) and timing applications (bipolar).](image-url)
to be identified, but this will provide a relatively poor energy measurement, so a slower spectroscopy amplifier is used for accuracy. Alternatively, deconvolution techniques can be implemented in real time to preserve energy resolution at higher count rates (Guo, Gardner, & Mayo, 2005).

4.4.3. DETECTOR TYPES AND MATERIALS FOR SPECTROSCOPY

4.4.3.1. X- and Gamma-ray Spectroscopy

Table 4.4.1 lists a range of common detector materials with their reported energy resolution values for photon interactions. The second column of the table shows the statistical component of the energy resolution quantified for each material. For semiconductors and gas detectors, a small ionization potential (in common terminology the band gap and the $w$-value, respectively) provides better energy resolution. For scintillators, a more representative parameter is the light yield, defined as the number of light photons generated per unit energy deposited. A high light yield generally infers that it should be (theoretically) possible to achieve good energy resolution.

A generalized summary of the table may be made by stating that semiconductor detectors typically show the best spectroscopic performance, followed by scintillators and then gas detectors. Some overlap occurs between detector types, but it is a reasonably well-followed rule. The best energy resolution of any commercially available detector is provided by high-purity germanium detectors, although their appeal is offset by their cost and bulky cooling systems. The intense commercial development of silicon over many years means that costs are low and they are widely available in a variety of forms. However, low-$Z$ and limited thickness ($<1$ mm is common) restrict their usefulness for direct X- and gamma-ray detection to low energies ($<20$ keV for good detection efficiency). Consequently, silicon devices for X- and gamma-ray detection are generally coupled to scintillators and thus suffer a reduction in spectroscopic performance. Of the room-temperature semiconductors, CdZnTe appears the most promising in terms of spectroscopic performance, and considerable effort is being expended to improve production quality (eV products; Verger et al., 2004).

The high light yield scintillators, in particular NaI(Tl) and CsI(Tl), demonstrate reasonable energy resolution, and their low cost and high $Z$ make them advantageous in certain applications. Sodium iodide was for many years the scintillator with the highest light yield, and so has long been
Table 4.4.1. Energy resolution values of common detector materials

<table>
<thead>
<tr>
<th>DETECTOR MATERIAL</th>
<th>STATISTICAL PARAMETER</th>
<th>ENERGY RESOLUTION</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semiconductor</td>
<td>Band gap (eV per electron-hole pair)</td>
<td>All values quoted at 122 keV</td>
<td>Value quoted for lithium drifted silicon detectors</td>
</tr>
<tr>
<td>Silicon</td>
<td>1.12</td>
<td>150–300 eV at 5.9 keV (ORTEC products)</td>
<td></td>
</tr>
<tr>
<td>Germanium</td>
<td>0.67</td>
<td>=0.5 keV (ORTEC products)</td>
<td></td>
</tr>
<tr>
<td>Cadmium telluride</td>
<td>1.5</td>
<td>1.8 keV (Niraula et al., 2002)</td>
<td>This performance is achieved with processing and correction; 3% to 10% can be expected for unprocessed signals (Feichtinger et al., 2004).</td>
</tr>
<tr>
<td>Cadmium zinc telluride</td>
<td>1.57</td>
<td>1.2 keV (Verger et al., 2004)</td>
<td></td>
</tr>
<tr>
<td>Mercuric iodide</td>
<td>2.15</td>
<td>1.2 keV FWHM for 59.5 keV</td>
<td>These measurements were taken at room temperature; resolution improved to 5.65 keV at 59.5 keV at 273 K.</td>
</tr>
<tr>
<td>Gallium arsenide</td>
<td>1.42</td>
<td>8.8 keV (Zat’ko et al., 2004)</td>
<td></td>
</tr>
<tr>
<td>Thallium bromide</td>
<td>2.68</td>
<td>6.0 keV (Onodera et al., 2004)</td>
<td></td>
</tr>
<tr>
<td>Scintillator</td>
<td>Light yield (photons/MeV)</td>
<td>All values quoted at 662 keV</td>
<td>Arguably the most commonly used scintillator for spectrometry since its discovery in 1948 (Hofstadter, 1948). Hygroscopic.</td>
</tr>
<tr>
<td>NaI(Tl)</td>
<td>38,000</td>
<td>5.9% (=39 keV) (Moszynski, 2003)</td>
<td>A dense, high-Z scintillator optimized for detection efficiency.</td>
</tr>
<tr>
<td>CsI(Tl)</td>
<td>52,000</td>
<td>4.9% (=32 keV) (Ikagawa et al., 2005)</td>
<td></td>
</tr>
<tr>
<td>BGO</td>
<td>8200</td>
<td>12% (=79 keV) (Hu et al., 2004)</td>
<td></td>
</tr>
</tbody>
</table>

continued on next page
widely used. Its hygroscopic nature requires encapsulation, thus for low-energy photons or particles, cesium iodide is often the scintillator of choice. Some new chloride and bromide scintillators, namely LaCl₃, LaBr₃, and CeBr₃, look particularly promising for spectroscopy. Reported energy resolutions at 662 keV of 2.9% and 3.4% for LaBr₃ and CeBr₃, respectively (Shah et al., 2004; van Loef et al., 2001), indicate that they could rival the spectroscopic performance of many semiconductors.

If high count rate is a requirement, there are a number of scintillators with very short decay times that are capable of high-speed operation. Plastic organic scintillators are notable for their short decay times, although their low $Z$/density tends to limit their application to particles or low-energy photons. For higher energy photons, BaF₂ and YAP:Ce offer some useful properties (Kerek et al., 1998). Gas proportional counters or ionization chambers operated in pulse mode can provide reasonable energy resolution, often better than a scintillator, but rarely competitive with a semiconductor. The main drawback of gas detectors for photon work is the low detection efficiency compared with the other materials, which tends to limit their application to energies less than 20 keV if sensitivity is a factor, although the useful range can be extended using high pressures (for further details, see Grey, Sood, & Manchanda, 2004). Despite this, gas detectors have other advantages and provide large area coverage at relatively low cost.

### 4.4.3.2. Charged Particle Spectroscopy

Both silicon and scintillation detectors are commonly employed for charged particle detection. For spectroscopic work there is much commonality between photon and charged particle studies, but there are a few important differences. First, low-$Z$ materials tend to be preferable for electron studies, which is opposite to the situation for photons. The reason for this is that the probabilities of bremsstrahlung production and backscatter increase significantly with the atomic number of the absorber, both of

#### Table 4.4.1. Energy resolution values of common detector materials (continued)

<table>
<thead>
<tr>
<th>Detector Material</th>
<th>Statistical Parameter</th>
<th>Energy Resolution</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scintillator</td>
<td>Light yield (photons/MeV)</td>
<td>All values quoted at 662 keV</td>
<td>Light output shows a very linear response with energy. Both this material and the one below are new materials showing very promising spectroscopic performance.</td>
</tr>
<tr>
<td>YAP:Ce</td>
<td>24,000</td>
<td>4.36% (=29 keV) (Kapusta et al., 1999)</td>
<td></td>
</tr>
<tr>
<td>LaBr₃</td>
<td>61,000</td>
<td>2.9% (=19 keV) (van Loef, Dorenbos, &amp; van Eijk, 2001)</td>
<td>Both this material and the one below are new materials showing very promising spectroscopic performance.</td>
</tr>
<tr>
<td>CeBr₃</td>
<td>68,000</td>
<td>3.4% (=22 keV) (Shah et al., 2004)</td>
<td></td>
</tr>
<tr>
<td>Gas</td>
<td>$w$-value (eV/electron-ion pair)</td>
<td>All values quoted at 662 keV</td>
<td>Theoretical limit of 1.3% FWHM at 140 keV and 0.6% at 662 keV (Bolotnikov &amp; Ramsey, 1998)</td>
</tr>
<tr>
<td>Xenon</td>
<td>21.5</td>
<td>14.4 keV at 662 keV (Dmitrenko et al., 2000)</td>
<td></td>
</tr>
</tbody>
</table>
which affect the spectral shape. Consequently, common detectors for electron studies are silicon—
either diodes or lithium drifted devices—low-Z organic scintillators such as anthracene or plastics, or liquid scintillators. Second, the detector entrance window needs careful consideration. Alpha- and beta-particle ranges are considerably shorter than equivalent energy X- or gamma rays, so detector encapsulation or surface layers become important issues. This requirement makes hygroscopic scintillators such as NaI(Tl) less suitable. Third, an advantage of the short range of the particles is that the detector can be thin or low Z and still achieve good detection efficiency. Consequently, silicon detectors, which have a typical thickness of less than 1 mm, exhibit good particle absorption across a normal energy range. A final point concerns high-resolution measurements of low-energy particles. A particle interaction can deposit such small amounts of energy in the detector that they are comparable with the detector noise. Thus particle detectors for high-resolution spectroscopy are operated to provide the lowest possible noise, and a minimum energy threshold is specified. Silicon detectors, for example, are commonly cooled to minimize thermal noise.

Tables 4.4.2 and 4.4.3 provide typical energy resolution values that can be expected for common
detector materials for alpha and beta detection. Similar trends are seen in X- and gamma-ray detectors, with the semiconductors demonstrating superior spectroscopic performance. Of the scintillators, anthracene is possibly the most common for beta detection because of its high light output, although requirements for other properties mean that a whole range of detector materials is encountered.

**Table 4.4.2. Reported energy resolutions of alpha particles from $^{241}$Am, a common source for calibrating alpha detectors**

<table>
<thead>
<tr>
<th>Detector material</th>
<th>Energy resolution (using $^{241}$Am at 5.48 MeV)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVD diamond</td>
<td>0.4%</td>
<td>Manfredotti (2005)</td>
</tr>
<tr>
<td>CsI(Tl)</td>
<td>6.0%</td>
<td>Bhattacharjee et al. (2002)</td>
</tr>
<tr>
<td>YAG(Ce)</td>
<td>8.4%</td>
<td>Bhattacharjee et al. (2002)</td>
</tr>
<tr>
<td>Silicon (surface barrier)</td>
<td>0.4%</td>
<td>Rahab et al. (2001)</td>
</tr>
<tr>
<td>Cadmium zinc telluride</td>
<td>1.8%</td>
<td>Pearson, Regan, &amp; Divoli (2001)</td>
</tr>
</tbody>
</table>

Note: ZnS is a common material for alpha-particle detection, but is only available as a polycrystalline powder. Consequently it is usually in the form of a thin sheet held together by a binding material. It is generally unsuitable for accurate energy measurement.

**Table 4.4.3. Reported energy resolutions of beta particles**

<table>
<thead>
<tr>
<th>Detector material</th>
<th>Energy resolution</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silicon CCD</td>
<td>3 keV at 219 keV ($^{111}$In)</td>
<td>Hofsaess et al. (2003)</td>
</tr>
<tr>
<td>Si (Li)</td>
<td>2 keV at 0.976 MeV</td>
<td>ORTEC</td>
</tr>
<tr>
<td>ZnSe(Te)</td>
<td>≈30 keV (3%–6%) at 0.976 MeV</td>
<td>Ryzhikov et al. (2005)</td>
</tr>
<tr>
<td>Anthracene (organic scintillator)</td>
<td>&lt;10%</td>
<td></td>
</tr>
<tr>
<td>Plastic scintillator</td>
<td>≈5%</td>
<td>Sanchez, Ono, &amp; Miyata (2002)</td>
</tr>
</tbody>
</table>

Note: A wide range of plastic scintillators exists, each exhibiting different light yield properties. Consequently the energy resolution varies according to the material chosen.
There is some interest in developing detectors that are sensitive to different forms of radiation (e.g., simultaneous measurement of both beta and gamma interactions). For such applications it is possible to use either two separate detectors on either side of the sample, or a single detector such as silicon that can provide a good signal from both particles and low-energy gammas, or composite detectors if higher energy gamma rays are used. A composite detector consists of two detector materials, each one optimized for radiation of a particular type and energy. A recent example of this is the CsI(Tl)/ZnSe(Te) composite of Ryzhikov et al. (2005).

4.4.3.3. Considerations for Achieving Optimal Energy Resolution

The values provided in Tables 4.4.2 and 4.4.3 are indicative of the optimal energy resolutions that can be achieved with those materials. The following sections discuss the various factors to be considered to achieve optimal energy resolution.

4.4.3.4. Charge/Light Collection

Section 4.4.1 describes the dependence of energy resolution on the number of charge carriers (or light photons in the case of scintillators). In order to approach the statistical limit, charge (light) collection should be maximized. This corresponds to the shaping time of the pulse-shaping amplifier.

4.4.3.5. Semiconductor/Gas Detectors

For the case of semiconductor detectors or gas detectors, the electrode structure and charge transport within the material affect energy resolution. The spectroscopic performance can be considerably restricted by electrode design (Baciak & He, 2003; Lacy et al., 2004). Pixelating the electrodes reduces the energy resolution of the detector. This is largely due to spread of charge over an area greater than the pixel. Summing the charge across the struck pixel and all neighboring pixels can help recover the energy resolution (He et al., 2000). However, for each detector type and geometry, certain pixel arrangements will provide better energy resolution than others (Gros d’Aillon et al., 2005).

Poor charge transport prevents the full amount of charge from contributing to the pulse, and therefore the energy measurement. This can lead to ionic recombination caused by the interaction of mobile charge carriers with the detector material, especially at low bias voltages or large detector volumes. A possible cause of poor charge transport is impurities in the detector material, leading to charge trapping and an apparent low energy measurement. Crystal growth methods ultimately dictate the charge transport properties of a semiconductor crystal, and hence the spectroscopic performance. High-purity germanium crystals can now be produced with excellent quality, but for newer materials such as CdZnTe, the quality of crystal growth is still being explored by a variety of methods (Funaki et al., 1999; Schlesinger et al., 1999; Szeles & Eissler, 1998; Verger et al., 2004).

Charge trapping effects are frequently seen in room temperature semiconductors, for example, cadmium telluride. This is because these materials tend to exhibit poor charge mobility (usually holes). The reader is directed to Knoll (2000) for further details. Such spectra where this is evident show a characteristic charge trapping “tail” on the low-energy side of the peak.

Various authors have reported methods for correcting for the charge trapping effect. One such technique is to use a capacitive Frisch grid, which is a modified electrode structure that minimizes the effect of hole motion such that the signal is primarily due to electrons, thus overcoming the problem of poor hole mobility (McNeill & McGregor, 2004; Montemont et al., 2001). Other suggested meth-
ods include a biparametric approach that corrects the measured pulse height using pulse rise time information (Verger et al., 2001) and pulse shape discrimination (Ho et al., 1998).

4.4.3.6. Scintillation Detectors

Efficient light collection in scintillation counters, and therefore good spectroscopic performance, is governed by the following factors:

1. Matching the emission spectrum of the scintillator to the absorption spectrum of the light counter
2. An efficient light counter
3. Transparency of the scintillator to its own emissions
4. A refractive index closely matched to the entrance window of the light counter to minimize reflections
5. Maximizing the amount of light that passes through the exit window of the scintillator.

Points 1 and 2 are key to the spectroscopic performance of a scintillation detector. A light counter exhibiting poor quantum efficiency at the wavelength emitted by the scintillator will produce poor energy resolution. Some improvements in energy resolution can be expected using recently developed photodetectors, such as avalanche silicon photodiodes (Moszynski et al., 2002), HgI₂ photodetectors (Wang, Iwanczyk, & Patt, 1994), and TlBr (Hitomi et al., 2000) because of the high quantum efficiencies of these materials. However, for low-energy photons or particles, electronic noise in photodiodes tends to dominate energy resolution.

For spectroscopic work, the choice of a bright scintillator largely dictates the light counter selection according to the required spectral response. For example, the emission spectrum of NaI(Tl) (a blue–violet emitter) is commonly matched with a bialkali photocathode on a photomultiplier tube, although for other applications it is often the case that the application warrants a particular light detector which in turn governs scintillator choice. Again, for example, high-resolution imaging applications could require a CCD, which has a broad absorption spectrum typically peaking at around 600 nm to 700 nm, at the red end of the visible spectrum. CCDs are typically coupled with Gd₂O₂S:Eu (commonly known as gadox–europium doped) or CsI(Tl), both of which emit some of their light at these wavelengths.

Point 5 can be achieved by applying a reflective coating to all surfaces of the scintillator other than the exit window and by choosing a suitable scintillator geometry (Naydenov, 2005). Figure 4.4.5 shows two scintillator configurations. For the one on the left, the majority of interactions occur just above the exit window. The light traveling in other directions is reflected from coated surfaces to contribute to the signal. In contrast, the scintillator on the right is required for an imaging device with good detection efficiency. A columnar scintillator structure with no reflective coatings ensures that light spread is minimized to maintain a good spatial resolution in the image, whereas good detection efficiency is ensured by a thick scintillator. However, only a small fraction of the light will reach the exit window by either direct transmission or total internal reflection, and thus the energy resolution will be poor. For example, a 3 mm thick CsI(Tl) matrix consisting of orthogonal arrays of crystal pillars with 0.6 mm × 0.6 mm cross sections showed an energy resolution of 22% FWHM at 122 keV, whereas 6% to 8% can be expected in optimal conditions (Vittori, Malatesta, & de Notaristefani, 1998).

4.4.3.7. Noise Reduction

Although radiation measurements are often quantum limited, section 4.4.1 showed that other noise sources are detrimental to the spectroscopic performance and should be minimized. This is particularly
relevant for low-energy measurements where small signals are encountered. Electronic noise in the readout of the detector signal will appear as random, low voltage pulses and can dominate the lower end of the measured spectrum. Often these pulses can be removed from the spectrum using a low-level discriminator on the multichannel analyzer; however, for low-energy depositions, the pulses generated can be of comparable amplitude to the noise. This is sometimes seen in the measurement of low-energy radiation using a photodiode coupled to a scintillator. Low-energy measurements require the careful selection of a low-noise system.

While many noise sources are caused by detector design and production quality, certain noise sources depend on operation. While the band gap of germanium is sufficiently low that it must be operated at low temperatures, all semiconductor detectors show significant improvement in spectroscopic performance if cooled. For example, a GaAs detector was reported to show an FWHM of 6.8 keV at 59.5 keV at 20°C compared to 5.65 keV at 0°C (Zat’ko et al., 2004). In the case of scintillation detectors, light leakage into the crystal can add a significant background noise source to worsen performance. And in all cases, environmental effects such as electric or magnetic fields or vibration can also have detrimental effects on performance.

4.4.4. CORRECTION TECHNIQUES TO IMPROVE SPECTRAL MEASUREMENT

Providing the response of the system and any experimental factors that can distort the spectrum are well known, some attempt can be made at extracting a “pure” spectrum from noisy experimental data that have been distorted by the detector system. The measured spectrum is effectively the convolution of the pure spectrum with the response of the system. Provided the system response can be well described, deconvolution techniques can be applied, often following data smoothing to reduce statistical noise. Potential problems with this approach are that the system response is not accurately known across the entire measurement range and that there are differences in experimental conditions between measuring the system response and the spectral measurements that are to be corrected. Consequently, maximum likelihood techniques can be used to estimate system response. The reader is directed to Gelfgar, Kosarev, and Podolyak (1993) and Pentilä et al. (2005) for further details. However, in all cases the ability to apply corrections does not detract from the need for good quality raw data.

Figure 4.4.5 Light collection efficiency in a scintillator according to geometry and reflective coatings.
4.4.4.1. Highest Achievable Energy Resolution
The best energy resolution of any detector type is provided by cryogenic microcalorimeters. These devices operate at very low temperatures (typically less than 100 mK) and measure the elevation of temperature generated by a radiation interaction in an absorber. An energy resolution of 160 eV (FWHM) at the 59.5 keV gamma line of $^{241}$Am has been reported (Loidl et al., 2004), which compares very favorably to a value of approximately 450 eV that can be expected from the best semiconductor. The energy resolution is so superior because the excitation energies required to cause a temperature elevation in an absorber are in the microelectronvolt to millielectronvolt range: this causes considerably less statistical fluctuation than the electronvolts required for ionization in other detector types. The low-temperature operation is necessary to minimize the heat capacity of the absorber and to limit the effect of thermal fluctuations, but it creates certain practical difficulties. Such systems are not currently commercially available; details can be found in Loidl et al. (2004) and Booth, Cabrera, and Fiorini (1996).

4.4.4.2. Energy Measurement Using Integrating Detectors
Discussion so far has focused on detectors operating in pulse mode, which is the conventional approach. However, spectroscopy is possible with an integrating imaging sensor. CCDs and CDDs are integrating sensors that have both achieved good energy resolution (Castoldi et al., 2004; Harris et al., 2001). The technique sums the charge deposited in a pixel cluster on a calibrated image to create an energy value. It relies on the event rate in the sensor being sufficiently low compared to the integrating frame rate so that no overlap occurs between pixel clusters from separate events. A CDD operating at a frame rate of 100 kHz can achieve an energy resolution at room temperature better than 300 eV FWHM at 5.9 keV. A value of approximately 150 eV is expected with moderate (Peltier) cooling (Castoldi et al., 2004).

4.4.4.3. Detector Selection for Spectroscopy: Case Studies

4.4.4.3.1. Case 1: High-Resolution Spectral Measurements in an X-ray Scatter Experiment
A collimated beam from a W-anode X-ray source irradiates a thin crystalline sample. We wish to accurately measure the energy spectrum diffracted to a small scatter angle, with an energy resolution better than 1 keV. The X-ray source is operated at 60 kV, emitting a continuous polyenergetic X-ray beam between 10 keV and 60 keV. The X-ray fluence rate at the detector is known to be approximately 10 kcounts s$^{-1}$ mm$^{-2}$.

*Detector selection.* The main priority for this work is high energy resolution. The spectroscopic requirements effectively restrict the choice to a germanium detector. X-ray energy is relatively low, so there are no special demands on germanium crystal thickness, therefore a planar geometry would be the preferred choice. However, it is important to consider the X-ray fluence rate. Germanium detectors require a shaping time of approximately 6 μs for optimum energy resolution. This corresponds to a maximum event rate in the detector of approximately 15 kcounts s$^{-1}$, for a dead time of less than 10%. Therefore it is necessary to collimate the scattered beam at the detector to an area of 1.5 mm$^2$ in order to keep the event rate within the preferred limit and maximize spectroscopic performance.

4.4.4.3.2. Case 2: Measuring a Radioisotope Distribution in a Biological Sample
A radiolabeled tracer is injected into a thick biological tissue section. We wish to dynamically image the distribution of the radiotracer through the tissue section. The radioisotope $^{99m}$Tc emits gamma rays at an energy of 140 keV. The area to be imaged is 3 cm × 3 cm and a pixel size of 4 mm × 4 mm
is required. In order to reduce the image contrast degradation effects of scattered radiation, an energy resolution of 10% is preferred.

Detector selection. An energy resolution of 10% opens up the detector choice to most semiconductor detectors, some high light yield scintillators, and even some gas detectors. The need to image dynamically imposes a strict demand on high detection efficiency, which at 140 keV effectively eliminates gas detectors and low-Z materials such as silicon. The imaging requirement of \(4 \text{ mm} \times 4 \text{ mm}\) is not overly demanding, but will cause some degradation of energy resolution and restrict scintillator choice to those with the highest light yields, such as NaI(Tl) or CsI(Tl).

Thus the choices are the high-Z, high light yield scintillators and high-Z semiconductors, all of which are capable of meeting the required specifications. In this situation it would be wise to consider the practical issues, which are usually dominated by cost. The most cost-effective solution would be a scintillator backed by photomultiplier tubes in an Anger gamma camera arrangement. However, if a key practical consideration is compactness, then a room temperature semiconductor such as CdZnTe may be a good solution.

ACKNOWLEDGEMENT

The authors are grateful to Dr. Beate Planskoy for her valuable comments and fruitful discussions on the theoretical and technical aspects of dosimetry.

REFERENCES


ICRU. (1979).


**ABOUT THE AUTHORS**

**Professor Robert Speller** obtained his first degree and PhD in physics from Imperial College London. He has over 40 years experience in radiation physics and currently holds a chair in the Department of Physics Applied to Medicine at University College London (UCL) and is head of the Radiation Physics Research Group. He has particular interests in scattered radiation fields and sensors and has published widely in these fields.

**Dr. Alessandro Olivo** did both his undergraduate and postgraduate studies in Trieste, Italy, obtaining his “Laurea” in physics *summa cum laude* in 1995 and his PhD in 2000. Following his PhD award he spent five years at ELETTRA, the synchrotron facility in Trieste, developing new X-ray imaging techniques and collaborating in the design of the first station for *in vivo* mammography with synchrotron radiation. Having been awarded a Marie Curie Fellowship, he moved to UCL in 2005, where he invented a new method to produce X-ray phase contrast images with conventional sources. Thanks to these results, he was awarded a Career Acceleration Fellowship by EPSRC to build a research group, which is currently developing the new methods in areas such as medical imaging, homeland security, material science, and others.
**Dr. Silvia Pani** obtained her “Laurea” degree (MPhys) and her PhD in physics from the University of Trieste (Italy) in 1996 and 2001, respectively. She then worked on the design and commissioning of the first synchrotron beamline dedicated to *in vivo* breast imaging at the synchrotron facility, ELETTRA. From 2004 to 2007 she worked at UCL as a Marie Curie Fellow and at St. Bartholomew’s and the London NHS trust on the development of novel X-ray techniques for tissue analysis, breast and dental imaging, and the identification of illegal substances in parcels. In 2008 she became a lecturer in applied radiation physics at the University of Surrey, Guildford, UK.

**Dr. Gary Royle** is currently senior lecturer at the Department of Medical Physics and Bioengineering at UCL. He obtained a physics degree from Imperial College London before completing a PhD in medical physics at UCL. He has over twenty years of experience in the field of medical radiation physics, primarily devoted to the detection and analysis of ionizing radiation for medical applications. This work has lead to one hundred and fifty publications and three patents. He has also taught extensively in this field and for many years ran a masters degree program in medical radiation physics.