Chapter 11

Semiconductor Diode Detectors

In many radiation detection applications, the use of a solid detection medium is of great advantage. For the measurement of high-energy electrons or gamma rays, detector dimensions can be kept much smaller than the equivalent gas-filled detector because solid densities are some 1000 times greater than that for a gas. Scintillation detectors offer one possibility of providing a solid detection medium, and their application to the detection and measurement of various radiations has been described in Chapter 10.

One of the major limitations of scintillation counters is their relatively poor energy resolution. The chain of events that must take place in converting the incident radiation energy to light and the subsequent generation of an electrical signal involves many inefficient steps. Therefore, the energy required to produce one information carrier (a photoelectron) is of the order of 100 eV or more, and the number of carriers created in a typical radiation interaction is usually no more than a few thousand. The statistical fluctuations in so small a number place an inherent limitation on the energy resolution that can be achieved under the best of circumstances, and nothing can be done about improving the energy resolution beyond this point. As detailed in Chapter 10, the energy resolution for sodium iodide scintillators is limited to about 6% when detecting 0.662 MeV gamma rays and is largely determined by the photoelectron statistical fluctuations.

The only way to reduce the statistical limit on energy resolution is to increase the number of information carriers per pulse. As we show in this chapter, the use of semiconductor materials as radiation detectors can result in a much larger number of carriers for a given incident radiation event than is possible with any other common detector type. Consequently, the best energy resolution from radiation spectrometers in routine use is achieved using semiconductor detectors. The fundamental information carriers are *electron–hole pairs* created along the path taken by the charged particle (primary radiation or secondary particle) through the detector. The electron–hole pair is somewhat analogous to the ion pair created in gas-filled detectors. Their motion in an applied electric field generates the basic electrical signal from the detector.

Devices employing semiconductors as the basic detection medium became practically available in the early 1960s. Early versions were called *crystal counters*, but modern detectors are referred to as *semiconductor diode detectors* or simply *solid-state detectors*. Although the latter term is somewhat ambiguous in the sense that technically scintillation counters can also be thought of as solid-state detectors, it has come into widespread use to characterize only those devices that are based on electron–hole pair collection from semiconductor media.

In addition to superior energy resolution, solid-state detectors can also have a number of other desirable features. Among these are compact size, relatively fast timing characteristics, and an effective thickness that can be varied to match the requirements of the

application. Drawbacks may include the limitation to small sizes and the relatively high susceptibility of these devices to performance degradation from radiation-induced damage.

Of the available semiconductor materials, silicon predominates in the diode detectors used primarily for charged particle spectroscopy and discussed in this chapter. Germanium is more widely used in the gamma-ray measurements described in Chapter 12, whereas devices that use other semiconductor materials are covered in Chapter 13.

Several comprehensive books are available on the topic of solid-state detectors, including Refs. 1–6. Each of these contains a rather complete citation of the literature up to the time of publication, and the other references in these chapters are largely limited to those that have appeared more recently.

I. SEMICONDUCTOR PROPERTIES

A. Band Structure in Solids

The periodic lattice of crystalline materials establishes allowed energy bands for electrons that exist within that solid. The energy of any electron within the pure material must be confined to one of these energy bands, which may be separated by gaps or ranges of forbidden energies. A simplified representation of the bands of interest in insulators or semiconductors is shown in Fig. 11.1. The lower band, called the *valence band*, corresponds to those outer-shell electrons that are bound to specific lattice sites within the crystal. In the case of silicon or germanium, they are parts of the covalent bonding that constitute the interatomic forces within the crystal. The next higher-lying band is called the *conduction band* and represents electrons that are free to migrate through the crystal. Electrons in this band contribute to the electrical conductivity of the material. The two bands are separated by the *bandgap*, the size of which determines whether the material is classified as a semiconductor or an insulator. The number of electrons within the crystal is just adequate to fill completely all available sites within the valence band. In the absence of thermal excitation, both insulators and semiconductors would therefore have a configuration in which the valence band is completely full and the conduction band completely empty. Under these circumstances, neither would theoretically show any electrical conductivity.

In a metal, the highest occupied energy band is not completely full. Therefore, electrons can easily migrate throughout the material because they need achieve only small incremental energy to be above the occupied states. Metals are therefore always characterized by very high electrical conductivity. In insulators or semiconductors, on the other hand, the electron must first cross the bandgap to reach the conduction band and the

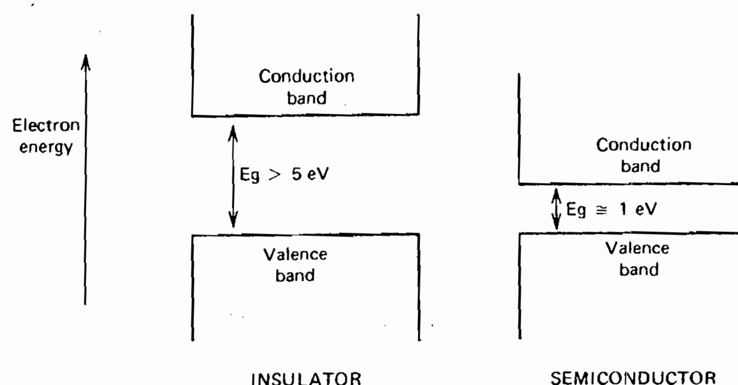


Figure 11.1 Band structure for electron energies in insulators and semiconductors.

conductivity is therefore many orders of magnitude lower. For insulators, the bandgap is usually 5 eV or more, whereas for semiconductors, the bandgap is considerably less.

B. Charge Carriers

At any nonzero temperature, some thermal energy is shared by the electrons in the crystal. It is possible for a valence electron to gain sufficient thermal energy to be elevated across the bandgap into the conduction band. Physically, this process simply represents the excitation of an electron that is normally part of a covalent bond such that it can leave the specific bonding site and drift throughout the crystal. The excitation process not only creates an electron in the otherwise empty conduction band, but it also leaves a vacancy (called a *hole*) in the otherwise full valence band. The combination of the two is called an *electron–hole pair* and is roughly the solid-state analogue of the ion pair in gases. The electron in the conduction band can be made to move under the influence of an applied electric field. The hole, representing a net positive charge, will also tend to move in an electric field, but in a direction opposite that of the electron. The motion of both of these charges contributes to the observed conductivity of the material.

The probability per unit time that an electron–hole pair is thermally generated is given by

$$p(T) = CT^{3/2}\exp\left(-\frac{E_g}{2kT}\right) \tag{11.1}$$

where

T = absolute temperature

E_g = bandgap energy

k = Boltzmann constant

C = proportionality constant characteristic of the material

As reflected in the exponential term, the probability of thermal excitation is critically dependent on the ratio of the bandgap energy to the absolute temperature. Materials with a large bandgap will have a low probability of thermal excitation and consequently will show the very low electrical conductivity characteristic of insulators. If the bandgap is as low as several electron volts, sufficient thermal excitation will cause a conductivity high enough for the material to be classified as a semiconductor. In the absence of an applied electric field, the thermally created electron–hole pairs ultimately recombine, and an equilibrium is established in which the concentration of electron–hole pairs observed at any given time is proportional to the rate of formation. From Eq. (11.1), this equilibrium concentration is a strong function of temperature and will decrease drastically if the material is cooled.[†]

After their formation, both the electron and the hole take part in a random thermal motion that results in their diffusion away from their point of origin. If all electrons (or holes) were initially created at a single point, this diffusion leads to a broadening distribution of the charges as a function of time. A cross section through this distribution would be approximated by a Gaussian function with a standard deviation σ given by

$$\sigma = \sqrt{2Dt} \tag{11.2}$$

[†]Because the ionization potential for gases is typically 15 eV or more, the probability of a thermally generated ion pair is negligibly small in gas ionization chambers, even at room temperature.

where D is the diffusion coefficient and t is the elapsed time. Values for D can be predicted from the relationship

$$D = \mu \frac{kT}{e} \tag{11.3}$$

where μ is the mobility of the charge carrier, k is the Boltzmann constant, and T is the absolute temperature. At 20°C (293K), the numerical value of kT/e is 0.0253 V.

C. Migration of Charge Carriers in an Electric Field

If an electric field is applied to the semiconductor material, both the electrons and holes will undergo a net migration. The motion will be the combination of a random thermal velocity and a net *drift velocity* parallel to the direction of the applied field. The motion of the conduction electrons is a relatively easy process to visualize, but the fact that holes also contribute to conductivity is less obvious. A hole moves from one position to another if an electron leaves a normal valence site to fill an existing hole. The vacancy left behind by the electron then represents the new position of the hole. Because electrons will always be drawn preferentially in an opposite direction to the electric field vector, holes move in the same direction as the electric field. This behavior is consistent with that expected of a point positive charge, because the hole actually represents the absence of a negatively charged electron.

At low-to-moderate values of the electric field intensity, the drift velocity v is proportional to the applied field. Then a *mobility* μ for both electrons and holes can be defined by

$$v_h = \mu_h \mathscr{E} \tag{11.4}$$

$$v_e = \mu_e \mathscr{E} \tag{11.5}$$

where $\mathscr{E}$ is the electric field magnitude. In gases, the mobility of the free electron is much larger than that of the positive ion, but in semiconductor materials the mobility of the electron and hole are roughly of the same order. Numerical values for common semiconductor materials are given in Table 11.1.

At higher electric field values, the drift velocity increases more slowly with the field. Eventually, a *saturation velocity* is reached which becomes independent of further increases in the electric field. Figure 11.2 shows the dependence of the drift velocity on field magnitude for silicon and germanium.

Many semiconductor detectors are operated with electric field values sufficiently high to result in saturated drift velocity for the charge carriers. Because these saturated velocities are of the order of 10^7 cm/s, the time required to collect the carriers over typical dimensions of 0.1 cm or less will be under 10 ns. Semiconductor detectors can therefore be among the fastest-responding of all radiation detector types.

In addition to their drift, the charge carriers will also undergo the influence of diffusion mentioned in the previous section. Without diffusion, all charge carriers would travel to the collecting electrodes following exactly the electric field lines that connect their point of origin to their collection point. The effect of diffusion is to introduce some spread in the arrival position that can be characterized as a Gaussian distribution whose standard deviation can be predicted by combining Eqs. (11.2), (11.3), and (11.4)

$$\sigma = \sqrt{\frac{2kTx}{e\mathscr{E}}} \tag{11.6}$$

where x represents the drift distance. In small-volume detectors, a typical value for σ would be less than 100 μm. This diffusion broadening of the charge distribution limits the precision to which position measurements can be made using the location at which charges are collected at the electrodes in semiconductor detectors.

Table 11.1 Properties of Intrinsic Silicon and Germanium

	Si	Ge
Atomic number	14	32
Atomic weight	28.09	72.60
Stable isotope mass numbers	28-29-30	70-72-73-74-76
Density (300 K); g/cm^3	2.33	5.32
Atoms/cm^3	4.96×10^{22}	4.41×10^{22}
Dielectric constant (relative to vacuum)	12	16
Forbidden energy gap (300 K); eV	1.115	0.665
Forbidden energy gap (0 K); eV	1.165	0.746
Intrinsic carrier density (300 K); cm^{-3}	1.5×10^{10}	2.4×10^{13}
Intrinsic resistivity (300 K); $\Omega \cdot$ cm	2.3×10^5	47
Electron mobility (300 K); cm^2/V $\cdot$ s	1350	3900
Hole mobility (300 K); cm^2/V $\cdot$ s	480	1900
Electron mobility (77 K); cm^2/V $\cdot$ s	2.1×10^4	3.6×10^4
Hole mobility (77 K); cm^2/V $\cdot$ s	1.1×10^4	4.2×10^4
Energy per electron–hole pair (300 K); eV	3.62	
Energy per electron–hole pair (77 K); eV	3.76	2.96
Fano factor (77 K)	0.143 (Ref. 7)	0.129 (Ref. 9)
	0.084 (Ref. 8)	0.08 (Ref. 10)
	0.085 to 0.137 (Ref. 12)	< 0.11 (Ref. 11)
		0.057 to 0.064 (Ref. 12)
	0.16 (Ref. 13)	0.058 (Ref. 14)

Source: G. Bertolini and A. Coche (eds.), *Semiconductor Detectors*, Elsevier-North Holland, Amsterdam, 1968, except where noted.

The collection time of the charges is also spread out by diffusion by an amount that can be estimated from the spatial broadening divided by the drift velocity, or generally less than 1 ns for small volumes. In many cases, these diffusion effects are negligible, and the charges can be pictured as moving along the electric field lines, all with the same drift velocity. However, the consequences of diffusion can become significant for large-volume detectors or when position or timing measurements of high precision are involved.

D. Effect of Impurities or Dopants

1. *INTRINSIC SEMICONDUCTORS*

In a completely pure semiconductor, all the electrons in the conduction band and all the holes in the valence band would be caused by thermal excitation (in the absence of ionizing radiation). Because under these conditions each electron must leave a hole behind, the number of electrons in the conduction band must exactly equal the number of holes in the valence band.

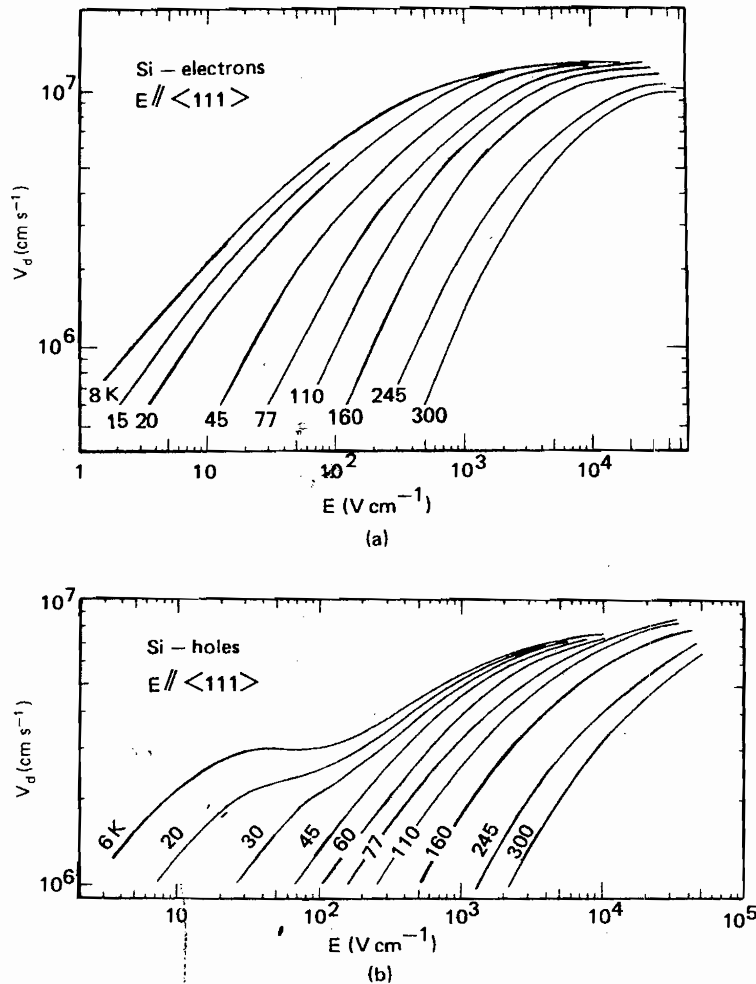


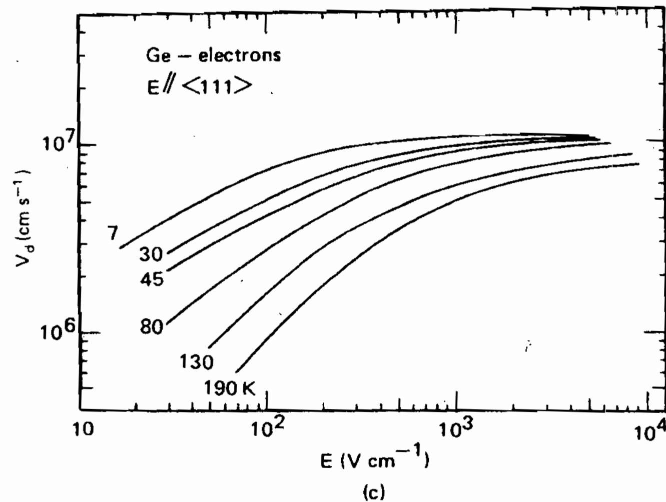
Figure 11.2 Drift velocity as a function of electric field applied parallel to the ⟨111⟩ crystallographic direction. Absolute temperature is the parameter for the different curves. (*a*) Electrons in silicon; (*b*) holes in silicon; (*c*) electrons in germanium; (*d*) holes in germanium. (From Ottaviani et al.[15])

Such material is called an *intrinsic* semiconductor. Its properties can be described theoretically, but in practice it is virtually impossible to achieve. The electrical properties of real materials tend to be dominated by the very small levels of residual impurities; this is true even for silicon and germanium, which are the semiconductors available in the highest practical purities.

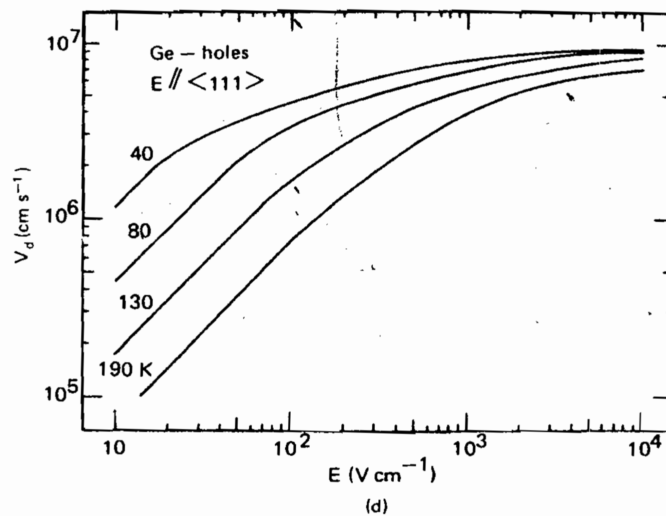
In the discussions that follow, we let n represent the concentration (number per unit volume) of electrons in the conduction band. Also, p represents the concentration of holes in the valence band. In the intrinsic material (subscript i), the equilibrium established by the thermal excitation of electrons from the valence to conduction band and their subsequent recombination leads to equal numbers of electrons and holes, or

$$n_i = p_i \tag{11.7}$$

The quantities n_i and p_i are known as the intrinsic carrier densities. From Eq. (11.1), it is clear that these densities will be lowest for materials with large bandgap energy and when the material is used at low temperature. Intrinsic hole or electron densities at room temperature are 1.5×10^{10} cm^{-3} in silicon, and 2.4×10^{13} cm^{-3} in germanium.



(c)



(d)

Figure 11.2 *(Continued)*

In a metallic conductor, only the flow of negatively charged electrons contributes to its electrical conductivity. In contrast, the flow of both negatively charged electrons and positively charged holes contribute to the conductivity of an intrinsic semiconductor. The value of the conductivity (or its inverse, the resistivity ρ) is determined by the intrinsic carrier density n_i and the mobilities μ_h and μ_e of the holes and electrons. If we have a slab of a semiconductor with thickness t and surface area A, the current I that will flow when a voltage V is applied across the thickness is

$$I = \frac{AV}{\rho t} \qquad \text{or} \qquad \rho = \frac{AV}{It}$$

The current is made up of two separate components: the current due to the flow of holes I_h and that due to the flow of electrons I_e. Note that, although the two types of charge carriers move in opposite directions, the separate currents are additive because of the opposite charges of holes and electrons. Thus an external measurement of the current by itself cannot distinguish between the flow of holes or electrons. The total observed current will

be their sum, with each term given by the product of the area, intrinsic carrier density, electronic charge e, and the drift velocity of the charge carrier. Thus

$$I = I_e + I_h = An_i e(v_e + v_h)$$

From Eqs. 11.2 and 11.3

$$I = An_i e\mathscr{E}(\mu_e + \mu_h) = An_i e\frac{V}{t}(\mu_e + \mu_h)$$

Combining

$$\rho = \frac{1}{en_i(\mu_e + \mu_h)} \tag{11.8}$$

Inserting numerical values for intrinsic silicon at room temperature:

$$\rho = \frac{1}{(1.6 \times 10^{-19}C)(1.5 \times 10^{10}/\text{cm}^3)(1350 + 480)\ \text{cm}^2/V \cdot s}$$

$$\rho = 2.3 \times 10^5 \frac{V \cdot s \cdot \text{cm}}{C} = 230{,}000\ \Omega \cdot \text{cm}$$

Presently available silicon material of the highest purity falls short of achieving this resistivity value because of the effect (discussed in the following sections) of residual impurities.

2. *n*-TYPE SEMICONDUCTORS

To illustrate the effect of doping on semiconductor properties, we use crystalline silicon as an example. Germanium and other semiconductor materials behave in a similar way. Silicon is tetravalent and in the normal crystalline structure forms covalent bonds with the four nearest silicon atoms. A sketch of this situation is shown in Fig. 11.3*a*, where each of the dashes represents a normal valence electron involved in a covalent bond. Thermal excitation in the intrinsic material consists of breaking loose one of these covalent electrons, leaving behind an unsaturated bond or hole.

We now consider the effect of the small concentration of impurity that may be present in the semiconductor either as a residual amount after the best purification processes, or as a small amount intentionally added to the material to tailor its properties. We first assume that the impurity is pentavalent or is found in group V of the periodic table. When present

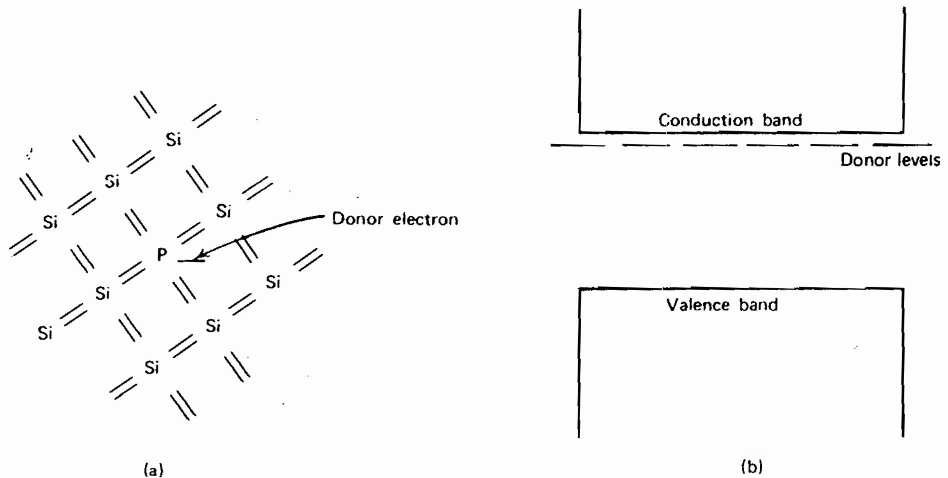


Figure 11.3 (*a*) Representation of a donor impurity (phosphorus) occupying a substitutional site in a silicon crystal. (*b*) Corresponding donor levels created in the silicon bandgap.

in small concentrations (of the order of a few parts per million or less) the impurity atom will occupy a substitutional site within the lattice, taking the place of a normal silicon atom. Because there are five valence electrons surrounding the impurity atom, there is one left over after all covalent bonds have been formed. This extra electron is somewhat of an orphan and remains only very lightly bound to the original impurity site. It therefore takes very little energy to dislodge it to form a conduction electron without a corresponding hole. Impurities of this type are referred to as *donor impurities* because they readily contribute electrons to the conduction band. Because they are not part of the regular lattice, the extra electrons associated with donor impurities can occupy a position within the normally forbidden gap. These very loosely bound electrons will have an energy near the top of the gap as shown in Fig. 11.3*b*. The energy spacing between these donor levels and the bottom of the conduction band is sufficiently small so that the probability of thermal excitation given by Eq. (11.1) is high enough to ensure that a large fraction of all the donor impurities are ionized. In nearly all cases, the concentration of impurity N_D is large compared with the concentration of electrons expected in the conduction band for the intrinsic material. Therefore, the number of conduction electrons becomes completely dominated by the contribution from the donor impurities, and we can write

$$n \cong N_D \tag{11.9}$$

The added concentration of electrons in the conduction band compared with the intrinsic value increases the rate of recombination, shifting the equilibrium between electrons and holes. As a result, the equilibrium concentration of holes is decreased by an amount such that the equilibrium constant given by the product of n and p is the same as for the intrinsic material:

$$np = n_i p_i \tag{11.10}$$

For example, in room-temperature silicon, the intrinsic carrier densities are about 10^{10} cm^{-3}. If a donor impurity is present at a concentration of 10^{17} atoms/cm^3 (about 2 parts per million), the density of conduction electrons n will be 10^{17} cm^{-3} and the concentration of holes p will be 10^3 cm^{-3}. Because the *total* number of charge carriers of both types is now much greater (10^{17} cm^{-3} versus 2×10^{10} cm^{-3}), the electrical conductivity of a doped semiconductor is always much larger than that of the corresponding pure material.

Even though conduction electrons now greatly outnumber the holes, charge neutrality is maintained because of the presence of ionized donor impurities. These sites represent net positive charges that exactly balance the excess electron charges. They are not, however, to be confused with holes because the ionized donors are fixed in the lattice and cannot migrate.

The net effect in n-type material is therefore to create a situation in which the number of conduction electrons is much greater and the number of holes much smaller than in the pure material. The electrical conductivity is then determined almost exclusively by the flow of electrons, and holes play a very small role. In this case, the electrons are called the *majority carriers* and holes the *minority carriers.*

The resistivity of doped material can be calculated from the dopant concentration and the mobility of the majority carrier, since only the flow of majority carriers is important in measured currents. As an example, assume we have silicon with a donor density of 10^{13}/cm^3, which will also be the concentration of conduction electrons. Then the resistivity will be

$$\rho = \frac{1}{eN_D\mu_e} \tag{11.11}$$

$$\rho = \frac{1}{(1.6 \times 10^{-19} C)(10^{13}/\text{cm}^3)(1350\ \text{cm}^2/V \cdot s)}$$

$$\rho = 463\ \Omega \cdot \text{cm}$$

3. p-TYPE SEMICONDUCTORS

The addition of a trivalent impurity such as an element from group III of the periodic table to a silicon lattice results in a situation sketched in Fig. 11.4*a*. If the impurity occupies a substitutional site, it has one fewer valence electron than the surrounding silicon atoms and therefore one covalent bond is left unsaturated. This vacancy represents a hole similar to that left behind when a normal valence electron is excited to the conduction band, but its energy characteristics are slightly different. If an electron is captured to fill this vacancy, it participates in a covalent bond that is not identical to the bulk of the crystal because one of the two participating atoms is a trivalent impurity. An electron filling this hole, although still bound to a specific location, is slightly less firmly attached than a typical valence electron. Therefore, these *acceptor impurities* also create electron sites within the normally forbidden energy gap. In this case, the acceptor levels lie near the bottom of the gap because their properties are quite close to sites occupied by normal valence electrons.

Normal thermal excitation in the crystal ensures that there will always be some electrons available to fill the vacancies created by the acceptor impurities or to occupy the acceptor sites shown in Fig. 11.4*b*. Because the energy difference between typical acceptor sites and the top of the valence band is small, a large fraction of all the acceptor sites are filled by such thermally excited electrons. These electrons come from other normal covalent bonds throughout the crystal and therefore leave holes behind in the valence band. To a good approximation, an extra hole is created in the valence band for every acceptor impurity that is added. If the concentration N_A of acceptor impurities is made to be large compared with the intrinsic concentration of holes p_i, then the number of holes is completely dominated by the concentration of acceptors, or

$$p \cong N_A \tag{11.12}$$

The increased availability of holes enhances the recombination probability between conduction electrons and holes and therefore decreases the equilibrium number of conduction electrons. Again, the same equilibrium constant discussed earlier holds, and $np = n_i p_i$. In *p*-type material, holes are the majority carrier and dominate the electrical conductivity. The filled acceptor sites represent fixed negative charges that balance the positive charge of the majority holes.

The equilibrium that is at work between electrons and holes is illustrated by the nomogram in Fig. 11.5. Two logarithmic scales are shown, on the left the concentration of con-

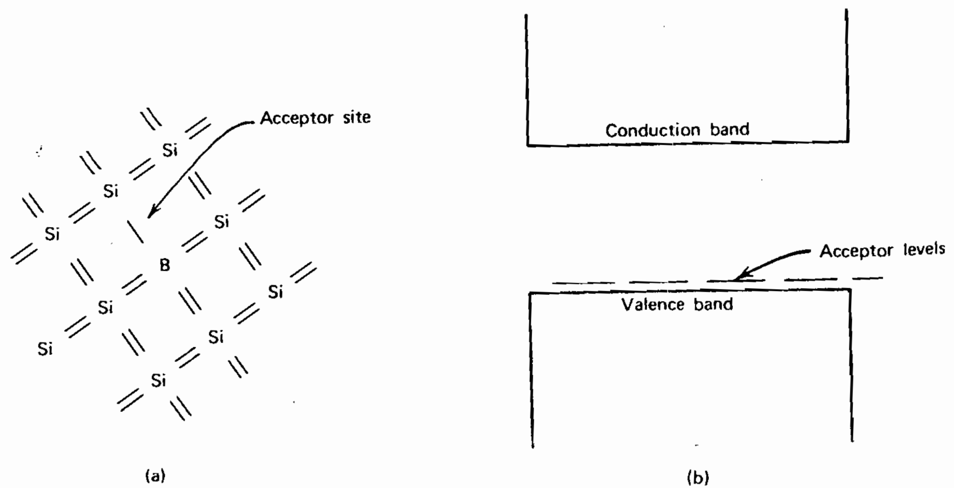


Figure 11.4 (*a*) Representation of an acceptor impurity (boron) occupying a substitutional site in a silicon crystal. (*b*) Corresponding acceptor levels created in the silicon bandgap.

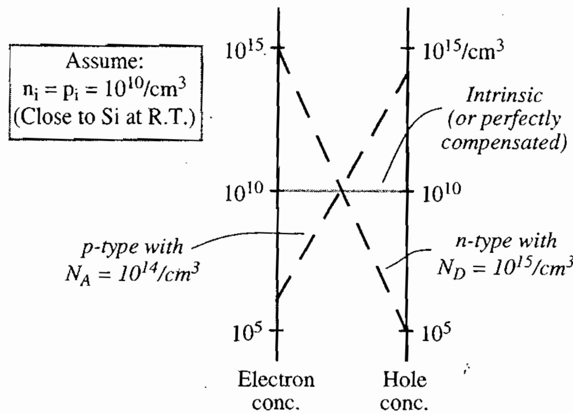


Figure 11.5 Nomogram showing the relationship between electron and hole concentrations in a semiconductor. Lines connecting points on the two logarithmic scales always pass through the center of the diagram for any type or degree of doping.

duction electrons, and on the right the concentration of holes. In intrinsic or perfectly compensated (see following section) semiconductors, the concentrations are equal and have a value given by the intrinsic carrier density. In silicon at room temperature, that value is $1.5 \times 10^{10}/cm^3$. Intrinsic (or perfectly compensated) material can thus be represented as a horizontal line connecting the points on the scales at about $10^{10}/cm^3$. Doping the material will tip the equilibrium so that one carrier dominates, but the product of the concentrations of electrons and holes must still be the same as shown in Eq. (11.10). Since the scales are logarithmic on the nomogram, the two concentrations will be linked by a line that passes through the center of the previous line. The same statement can be made for any level of doping, whether with donors or acceptors, so this point at the center of the nomogram acts as a "pivot point" that is common to all lines. Material that is n-type will result in lines with negative slope, while p-type material corresponds to lines with positive slope.

Figure 11.5 also illustrates that the total concentration of charge carriers (electrons plus holes) is at a minimum for the intrinsic material. This condition corresponds to a minimum in electrical conductivity if the mobilities of the two carriers are about the same. Tipping the equilibrium in either direction due to the presence of either donors or acceptors will raise the majority carrier concentration by an absolute amount that is always greater than the amount by which the minority carrier concentration is decreased. The effect is shown in Fig. 11.6, where the minimum conductivity occurs for the pure or intrinsic material, and either excess donors or acceptors result in a higher conductivity. Thus one measure of the impurity level of semiconductor materials is the electrical conductivity, or

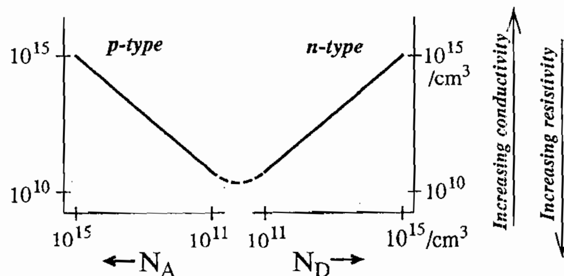


Figure 11.6 Plot using logarithmic scales of the conductivity of a semiconductor as a function of the net concentration of acceptors (N_A) or donors (N_D).

its inverse, the resistivity. For the intrinsic material in which all the charge carriers are produced through thermal excitation, the resistivity value can be calculated from Eq. (11.8) using the intrinsic carrier density and mobilities. The corresponding values for silicon and germanium are listed in Table 11.1. In practice, these theoretical values of resistivity are never observed because of the unavoidable residual impurities. Using the most advanced purification methods available at this writing, silicon resistivity of about 50,000 Ω-cm can be achieved, compared with a theoretical value of over 200,000 Ω-cm.

4. COMPENSATED MATERIAL

If donor and acceptor impurities are present in a semiconductor in equal concentration, the material is said to be *compensated.* Such material has some of the properties of an intrinsic semiconductor because electrons contributed by donor impurities are removed to some extent by their capture at the site of acceptor impurities. Despite the potential confusion with purified intrinsic material, compensated regions in semiconductors are commonly given the designation i because of their near intrinsic properties.

In practice, it is impossible to achieve exact compensation at the time of fabrication of the doped material because any small imbalance in the acceptor or donor concentration quickly leads to n-type or p-type behavior. At present, the only practical means for achieving compensation over large volumes in silicon or germanium is through the lithium ion drifting process after the crystal has been fabricated. This procedure is discussed in Chapter 13.

5. HEAVILY DOPED MATERIAL

Thin layers of semiconductor material that have an unusually high concentration of impurity are often given a special notation. Thus, n^+ and p^+ designate heavily doped n- and p-type layers that, as a result, have very high conductivity. These layers are often used in making electrical contact with semiconductor devices, because the very low minority carrier density allows their application as "blocking" contacts described later in this chapter.

E. Trapping and Recombination

Once electrons and holes are formed in a semiconductor, they will tend to migrate either spontaneously or under the influence of an applied electric field until they are either collected at an electrode or recombination takes place. There are theoretical predictions[16] that the average lifetime of charge carriers before recombination in perfectly pure semiconductors could be as large as a second. In practice, lifetimes at least three or four orders of magnitude smaller than a second are actually observed that are dominated entirely by the very low level of impurities remaining in the material. Some of these impurities, such as gold, zinc, cadmium, or other metallic atoms occupying substitutional lattice positions, introduce energy levels near the middle of the forbidden gap. They are therefore classified as "deep impurities" (as opposed to acceptor or donor impurities that, because the corresponding energy levels lie near the edges of the forbidden band, are called *shallow impurities*). These deep impurities can act as *traps* for charge carriers in the sense that if a hole or electron is captured, it will be immobilized for a relatively long period of time. Although the trapping center ultimately may release the carrier back to the band from which it came, the time delay is often sufficiently long to prevent that carrier from contributing to the measured pulse.

Other types of deep impurities can act as *recombination centers.* These impurities are capable of capturing both majority and minority carriers, causing them to annihilate. An impurity level near the center of the forbidden gap might, for example, first capture a conduction electron. At a slightly later time, a hole from the valence band might also be captured, with the electron then filling the hole. The impurity site is thus returned to its original state and is capable of causing another recombination event. In most crystals,

recombination through such centers is far more common than direct recombination of electrons and holes across the full bandgap.

Both trapping and recombination contribute to the loss of charge carriers and tend to reduce their average lifetime in the crystal. For the material to serve as a good radiation detector, a large fraction (preferably 100%) of all the carriers created by the passage of the incident radiation should be collected. This condition will hold provided the collection time for the carriers is short compared with their mean lifetime. Collection times of the order of 10^{-7} to 10^{-8} s are fairly common, so that carrier lifetimes of the order of 10^{-5} s or longer are usually sufficient.

Another widely quoted specification is the *trapping length* within the material. This quantity is simply the mean distance traveled by a carrier before trapping or recombination and is given by the product of the mean lifetime and the average drift velocity. In order to have an acceptable detector, the trapping length should be long compared with the physical dimensions over which the charge must be collected.

In addition to impurities, structural defects within the crystal lattice can also lead to trapping and charge carrier loss. These imperfections include point defects such as vacancies or interstitials that tend to behave as acceptors or donors, respectively. Carrier loss may also occur at line defects or dislocations that may be produced in stressed crystals. A dislocation represents the slippage of one crystal plane with respect to another, and its intersection with the surface of the crystal leads to a pit upon chemical etching. The density of these etched pits is often quoted as a measure of the crystalline perfection of a semiconductor sample.

II. THE ACTION OF IONIZING RADIATION IN SEMICONDUCTORS

A. The Ionization Energy

When a charged particle passes through a semiconductor with the band structure shown in Fig. 11.1, the overall significant effect is the production of many electron–hole pairs along the track of the particle. The production process may be either direct or indirect, in that the particle produces high-energy electrons (or *delta rays*) that subsequently lose their energy in producing more electron–hole pairs. Regardless of the detailed mechanisms involved, the quantity of practical interest for detector applications is the average energy expended by the primary charged particle to produce one electron–hole pair. This quantity, often loosely called the *ionization energy* and given the symbol ϵ, is experimentally observed to be largely independent of both the energy and type of the incident radiation. This important simplification allows interpretation of the number of electron–hole pairs produced in terms of the incident energy of the radiation, provided the particle is fully stopped within the active volume of the detector.

When radiation interacts in a semiconductor, the energy deposition always leads to the creation of equal numbers of holes and electrons. This statement holds regardless of whether the host semiconductor is pure or intrinsic, or doped as *p*-type or *n*-type. Just as equal numbers of free electrons and positive ions are created in a gas, every conduction electron produced in a semiconductor must also create a hole in the valence band, leading to an exact balance in the initial number of created charges. It should also be emphasized that the doping levels typical in *p*- or *n*-type semiconductors are so low that these atoms play no significant role in determining the nature of the radiation interactions in the material. Thus *p*-type or *n*-type silicon of equal thickness will present identical interaction probabilities for gamma rays, and the range of charge particles in either type will also be the same.

The dominant advantage of semiconductor detectors lies in the smallness of the ionization energy. The value of ϵ for either silicon or germanium is about 3 eV (see Table 11.1), compared with about 30 eV required to create an ion pair in typical gas-filled detectors. Thus, the number of charge carriers is 10 times greater for the semiconductor case, for a given energy

deposited in the detector. The increased number of charge carriers has two beneficial effects on the attainable energy resolution. The statistical fluctuation in the number of carriers per pulse becomes a smaller fraction of the total as the number is increased. This factor often is predominant in determining the limiting energy resolution of a detector for medium to high radiation energy. At low energies, the resolution may be limited by electronic noise in the preamplifier, and the greater amount of charge per pulse leads to a better signal/noise ratio.

More detailed examination shows that ϵ depends on the nature of the incident radiation. Most detector calibrations are carried out using alpha particles, and the values for ϵ shown in Table 11.1 are based on this mode of excitation. All experimental values obtained using other light ions or fast electrons seem to be fairly close,[17–19] but differences as large as 2.2% have been reported[20] between proton and alpha particle excitation in silicon. These observed differences point up the need to carry out detector calibration using a radiation type that is identical to that involved in the measurement itself if precise energy values are required.

A much larger difference is measured for ϵ when heavy ions or fission fragments are involved. The value of ϵ is significantly higher than for alpha particle excitation, leading to a lower than anticipated number of charge carriers. The physical origins of this *pulse height defect* are discussed later in this chapter.

The ionization energy is also temperature dependent. For the most significant detector materials, the value of ϵ increases with decreasing temperature. As shown in Table 11.1, ϵ in silicon is about 3% greater at liquid nitrogen temperature compared with room temperature.[18]

There is also evidence[21–23] that the ionization energy can be dependent on the energy of the radiation, especially in the soft X-ray energy range. It appears that its value increases with decreasing X-ray energy below about 1 keV, and that the Fano factor (described below) also increases significantly with decreasing energy in this range.

B. The Fano Factor

In addition to the mean number, the fluctuation or variance in the number of charge carriers is also of primary interest because of the close connection of this parameter with energy resolution of the detector. As in gas counters, the observed statistical fluctuations in semiconductors are smaller than expected if the formation of the charge carriers were a Poisson process. The Poisson model would hold if all events along the track of the ionizing particle were independent and would predict that the variance in the total number of electron–hole pairs should be equal to the total number produced, or E/ϵ. The Fano factor F is introduced as an adjustment factor to relate the observed variance to the Poisson predicted variance:

$$F \equiv \frac{\text{observed statistical variance}}{E/\epsilon} \tag{11.13}$$

For good energy resolution, one would like the Fano factor to be as small as possible. Although a complete understanding of all the factors that lead to a nonunity value for F does not yet exist, rather sophisticated models have been developed[24] that at least qualitatively account for experimental observations. Some numerical values for silicon and germanium are given in Table 11.1.

There is considerable variation in reported experimental values, particularly for silicon. The Fano factor is usually measured by observing the energy resolution from a given detector under conditions in which all other factors that can broaden the full-energy peak (such as electronic noise or drift) can be estimated and taken into account. The assumption is then made that the residual width can be attributed to statistical effects only. If nonstatistical residual factors remain, however, the Fano factor will appear to be larger than it actually is. This may explain the historical trend toward lower values as measurement procedures are refined. It has also been postulated[13] that the value of the Fano factor may depend on the nature of the particle that deposits the energy. Some measurements[25]

suggest that its value in silicon may also vary significantly with radiation energy, especially in the energy range typical of X-ray photons.

III. SEMICONDUCTORS AS RADIATION DETECTORS

A. Pulse Formation

When a particle deposits energy in a semiconductor detector, equal numbers of conduction electrons and holes are formed within a few picoseconds along the particle track. The detector configurations that are discussed in the following sections all ensure that an electric field is present throughout the active volume, so that both charge carriers feel electrostatic forces that cause them to drift in opposite directions. The motion of either the electrons or holes constitutes a current that will persist until those carriers are collected at the boundaries of the active volume. With the simplifying assumption that all charge carriers are formed at a single point, the resulting currents can be represented by the plot at the top of Fig. 11.7. Since the charge collection times are not likely to be the same because of differences in drift distance and carrier mobilities, one of the two currents will persist for a longer time than the other. When these currents are integrated on a measuring circuit with long time constant, the measured induced charge has the time characteristics shown in the lower plot in Fig. 11.7. This time profile will also be that of the rise of the pulse produced by a conventional preamplifier used to process the pulses from the detector. This pulse profile is similar to that derived in Chapter 5 for ion chambers (see Fig. 5.16), with the exception of the time scale. In gases, the collection time for the positive charges (ions) is greater by orders of magnitude than that for the negative (free electron) charges, so the ion motion adds a very long component to the pulse rise and, as a practical matter, does not contribute to the output pulse. In silicon or germanium semiconductor detectors, the hole mobility is within a factor of about 2 or 3 of the electron mobility, so the collection times are much closer to being equivalent. As a result, while pulse-type ion chambers almost never include the motion of the ions in the output pulse, standard silicon and germanium semiconductor detectors rely on complete integration of the currents due to both the electrons and holes. Both carrier types must therefore be completely collected for the resulting pulse to be a faithful measure of the energy deposited by the particle.

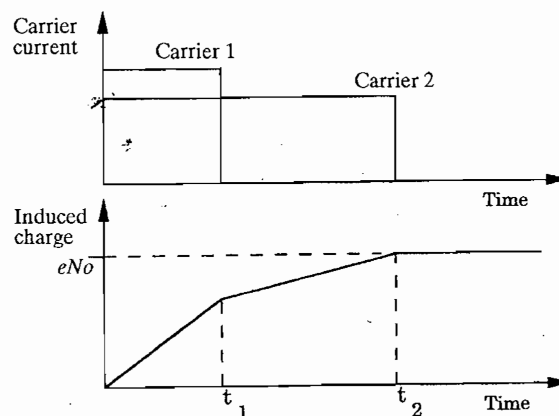


Figure 11.7 The upper plot shows an idealized representation of the electron and hole currents flowing in a semiconductor following the creation of N_0 electron–hole pairs. In the lower plot, t_1 represents the collection time for the carrier type (either electrons or holes) that is collected first, and t_2 is the collection time for the other carrier. If both are fully collected, a charge of eN_0 is induced to form the signal, where e is the electronic charge.

B. Electrical Contacts

In order to construct a practical radiation detector, some means must be provided to collect the electrical charges created by the radiation at either boundary of the semiconductor material. An *ohmic* contact is a nonrectifying electrode through which charges of either sign can flow freely. If two ohmic contacts are fitted on opposite faces of a slab of semiconductor and connected to a detection circuit, the equilibrium charge carrier concentrations in the semiconductor will be maintained. If an electron or hole is collected at one electrode, the same carrier species is injected at the opposite electrode to maintain the equilibrium concentrations in the semiconductor.

The steady-state leakage currents that are observed using ohmic contacts are too high, even with the highest resistivity material available (see following section), to permit their general application to silicon or germanium detectors. Instead, *noninjecting* or *blocking* electrodes are universally employed to reduce the magnitude of the current through the bulk of the semiconductor. If blocking electrodes are used, charge carriers initially removed by the application of an electric field are not replaced at the opposite electrode, and their overall concentration within the semiconductor will drop after application of an electric field. The leakage current can thus be reduced to a sufficiently low value to allow the detection of the added current pulse created by the electron–hole pairs produced along the track of an ionizing particle.

The most appropriate type of blocking contacts are the two sides of a *p-n* semiconductor junction. It is very difficult to inject electrons from the *p* side of this junction because holes are the majority carrier and free electrons are relatively scarce. At the opposite side, electrons are the majority carrier and holes cannot readily be injected. In this chapter, we discuss detectors that are created by placing the *p*- and *n*-type materials in direct contact, forming a *p-n* junction. In Chapters 12 and 13, detectors in which the *p* and *n* regions are separated by an intrinsic or compensated region (the *i* region) are described.

C. Leakage Current

In order to create an electric field large enough to achieve an efficient collection of the charge carriers from any semiconductor detector, an applied voltage of typically hundreds or thousands of volts must be imposed across the active volume. Even in the absence of ionizing radiation, all detectors will show some finite conductivity and therefore a steady-state *leakage current* will be observed. Random fluctuations that inevitably occur in the leakage current will tend to obscure the small signal current that momentarily flows following an ionizing event and will represent a significant source of noise in many situations. Methods of reducing the leakage current are therefore an important consideration in the design of semiconductor detectors.

The resistivity of the highest purity silicon currently available is about 50,000 Ω-cm. If a 1-mm thick slab of this silicon were cut with 1-cm^2 surface area and fitted with ohmic contacts, the electrical resistance between faces would be 5000 Ω. An applied voltage of 500 V would therefore cause a leakage current through the silicon of 0.1 A. In contrast, the peak current generated by a pulse of 10^5 radiation-induced charge carriers would only be about 10^{-6} A. It is therefore essential to reduce this bulk leakage current greatly through the use of blocking contacts. In critical applications, the leakage current must not exceed about 10^{-9} A to avoid significant resolution degradation.

At these levels, leakage across the surface of the semiconductor can often become more significant than bulk leakage. Great care is taken in the fabrication of semiconductor detectors to avoid contamination of the surfaces, which could create leakage paths. Some configurations may also use grooves in the surface or guard rings to help suppress surface leakage (see Chapter 5).

D. The Semiconductor Junction

1. BASIC JUNCTION PROPERTIES

The radiation detectors described in this chapter are based on the favorable properties that are created near the junction between *n*- and *p*-type semiconductor materials. Charge carriers are able to migrate across the junction if the regions are brought together in good thermodynamic contact. Simply pressing together two pieces of the material will not suffice because gaps will inevitably be left that will be large compared with the interatomic lattice spacing. In practice, the junction is therefore normally formed in a single crystal by causing a change in the impurity content from one side of the junction to the other.

As an illustration, assume that the process begins with a *p*-type crystal that has been doped with a uniform concentration of acceptor impurity. In the concentration profile at the top of Fig. 11.8, this original acceptor concentration N_A is shown as a horizontal line. We now assume that the surface of the crystal on the left is exposed to a vapor of an *n*-type impurity that diffuses some distance into the crystal. The resulting donor impurity profile is labeled N_D on the figure and falls off as a function of distance from the surface. Near the surface, the donor impurities can be made to outnumber the acceptors, converting the left portion of the crystal to *n*-type material.

The approximate variation of equilibrium charge carrier concentration is also plotted at the top of Fig. 11.8 and labeled as *p* (hole concentration) and *n* (conduction electron concentration). These profiles are subsequently altered in the vicinity of the *p*-*n* junction because of the effects of charge carrier diffusion. In the *n*-type region at the left, the density of conduction electrons is much higher than in *p*-type. The junction between the two regions therefore represents a discontinuity in the conduction electron density. Wherever such a sharp gradient exists for any carrier that is free to migrate, a net diffusion from regions of high concentration to those of low concentration must take place. Thus, there will be some net diffusion of conduction electrons into the *p*-type material, where they will quickly combine with holes. In effect, this annihilation represents the capture of the conduction electron by one of the vacancies existing in the covalent bonds in the *p*-type material. The diffusion of conduction electrons out of the *n*-type material leaves behind immobile positive charges in the form of ionized donor impurities. A similar and symmetric argument leads to the conclusion that holes (the majority in the *p*-type material) must also diffuse across the junction because they also see an abrupt density gradient. Each hole that is removed from the *p* side of the junction leaves behind an acceptor site that has picked up an extra electron and therefore represents a fixed and immobile negative charge. The combined effect is to build up a net negative space charge on the *p* side and a positive space charge on the *n* side of the junction.

The accumulated space charge creates an electric field that diminishes the tendency for further diffusion. At equilibrium, the field is just adequate to prevent additional net diffusion across the junction, and a steady-state charge distribution is therefore established.

The region over which the charge imbalance exists is called the *depletion region* and extends into both the *p* and *n* sides of the junction. If the concentrations of donors on the *n* side and acceptors on the *p* side are equal, the diffusion conditions are approximately the same for both holes and electrons, and the depletion region extends equal distances into both sides. Usually, however, there is a marked difference in the doping levels on one side of the junction compared with the other. For example, if the donor concentration in the *n*-type material is higher than that of acceptor atoms in the *p*-type, the electrons diffusing across the junction will tend to travel a greater distance into the *p*-type material before all have recombined with holes. In this case, the depletion region would extend farther into the *p* side.

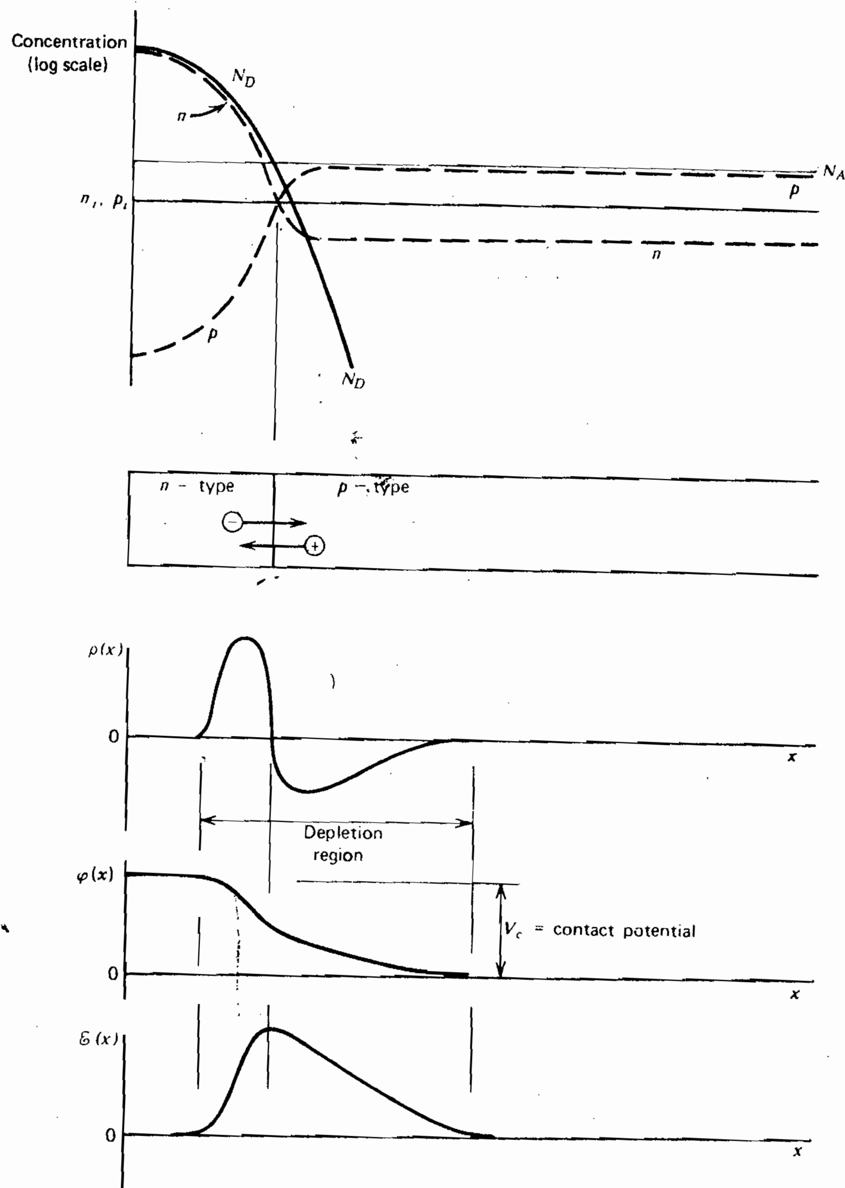


Figure 11.8 The assumed concentration profiles for the *n-p* junction shown at the top are explained in the text. The effects of carrier diffusion across the junction give rise to the illustrated profiles for space charge $\rho(x)$, electric potential $\varphi(x)$, and electric field $\mathscr{E}(x)$.

The buildup of net charge within the region of the junction leads to the establishment of an electric potential difference across the junction. The value of the potential φ at any point can be found by solution of Poisson's equation

$$\nabla^2\varphi = -\frac{\rho}{\epsilon} \tag{11.14}$$

where ϵ is the dielectric constant of the medium, and ρ is the net charge density. In one dimension, Eq. (11.14) takes the form

$$\frac{d^2\varphi}{dx^2} = -\frac{\rho(x)}{\epsilon} \tag{11.15}$$

so that the shape of the potential across the junction can be obtained by twice integrating the charge distribution profile $\rho(x)$. Graphical examples are shown in Fig. 11.8. At equilibrium, the potential difference across the junction (called the *contact potential*) amounts to nearly the full bandgap value of the semiconductor material. The direction of this potential difference is such that it opposes the further diffusion of electrons from left to right and holes from right to left in Fig. 11.8.

Where a difference in electrical potential exists, there must also be an electric field $\mathscr{E}$. Its magnitude is found by taking the gradient of the potential

$$\mathscr{E} = -\text{grad } \varphi \tag{11.16}$$

which, in one dimension, is simply

$$\mathscr{E}(x) = -\frac{d\varphi}{dx} \tag{11.17}$$

The electric field will extend over the width of the depletion region, in which charge imbalance is significant and the potential has some gradient. Its variation is also sketched in Fig. 11.8.

The depletion region exhibits some very attractive properties as a medium for the detection of radiation. The electric field that exists causes any electrons created in or near the junction to be swept back toward the n-type material, and any holes are similarly swept toward the p-type side. The region is thus "depleted" in that the concentration of holes and electrons is greatly suppressed. The only significant charges remaining in the depletion region are the immobile ionized donor sites and filled acceptor sites. Because these latter charges do not contribute to conductivity, the depletion region exhibits a very high resistivity compared with the n- and p-type materials on either side of the junction. Electron–hole pairs that are created within the depletion region by the passage of radiation will be swept out of the depletion region by the electric field, and their motion constitutes a basic electrical signal.

The thermal generation of charge carriers will continue to take place in the depletion region, contributing a component sometimes called the *generation current* to the observed leakage current. These charges are swept away typically within a few nanoseconds, however, a time that is many orders of magnitude shorter than the time required to establish thermal equilibrium. Thus, the steady-state concentration of carriers is strongly reduced in the depletion region because the removal of charges is a much faster process than their creation. The small concentration of carriers created by an ionizing particle is therefore easily detected above this highly suppressed, thermally generated concentration.

2. REVERSE BIASING

Thus far, we have discussed a semiconductor diode junction to which no external voltage is applied. Such an unbiased junction will function as a detector, but only with very poor performance. The contact potential of about 1 V that is formed spontaneously across the junction is inadequate to generate a large enough electric field to make the charge carriers move very rapidly. Therefore, charges can be readily lost as a result of trapping and recombination, and incomplete charge collection often results. The thickness of the depletion region is quite small, and the capacitance of an unbiased junction is high. Therefore, the noise properties of an unbiased junction connected to the input stage of a preamplifier are quite poor. For these reasons, unbiased junctions are not used as pulse mode radiation detectors, but instead, an external voltage is applied in the direction to cause the semiconductor diode to be reverse biased.

The p-n junction is most familiar in its role as a diode. The properties of the junction are such that it will readily conduct current when voltage is applied in the "forward" direction, but it will conduct very little current when biased in the "reverse" direction. In the

configuration of Fig. 11.8, first assume that a positive voltage is applied to the p side of the junction with respect to the n side. The potential will tend to attract conduction electrons from the n side as well as holes from the p side across the junction. Because, in both cases, these are the majority carriers, conductivity through the junction is greatly enhanced. The contact potential shown in Fig. 11.8 is reduced by the amount of the bias voltage that is applied, which tends to lessen the potential difference seen by an electron from one side of the junction to the other. This is the direction of forward biasing, and only small values of the forward bias voltage are needed to cause the junction to conduct large currents.

If the situation is reversed, and the p side of the junction is made negative with respect to the n side, the junction is reverse biased. Now the natural potential difference from one side of the junction to the other is enhanced, as shown in Fig. 11.9c. Under these

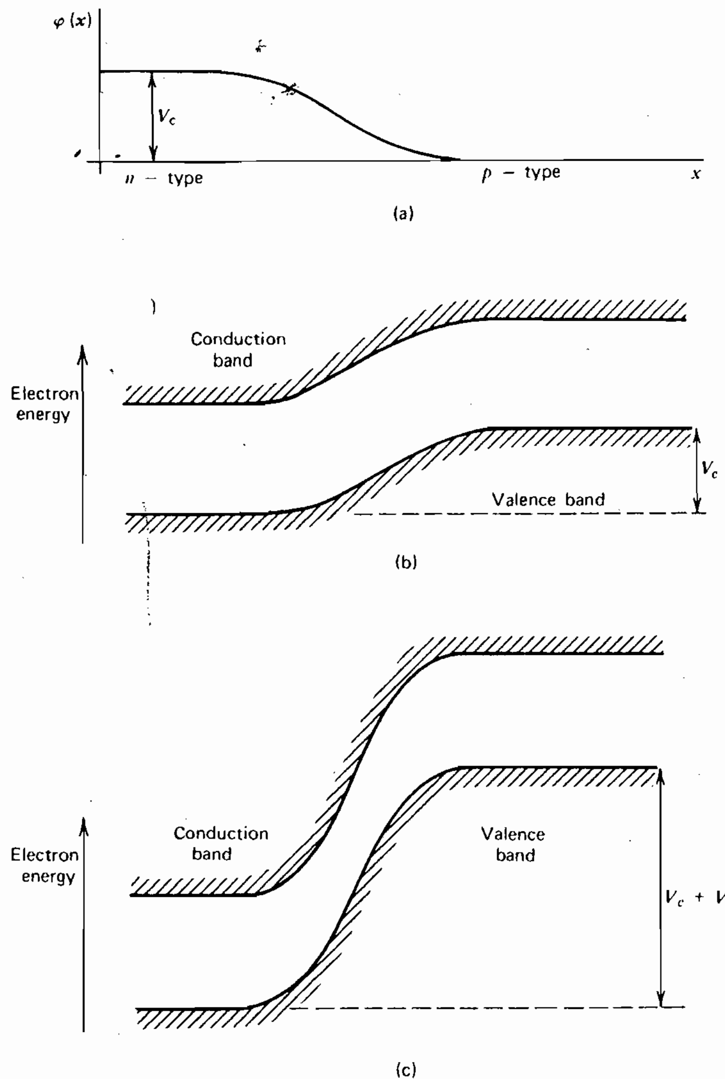


Figure 11.9 (a) The variation of electric potential across an n-p junction from Fig. 11.8. (b) The resulting variation in electron energy bands across the junction. The curvature is reversed because an increase in electron energy corresponds to a decrease in conventional electric potential $\varphi(x)$ defined for a positive charge. (c) The added displacement of the bands caused by application of a reverse bias V across the junction.

circumstances, it is the minority carriers (holes on the n side and electrons on the p side) that are attracted across the junction and, because their concentration is relatively low, the reverse current across the diode is quite small. Therefore, the p-n junction serves as a rectifying element, allowing relatively free flow of current in one direction while presenting a large resistance to its flow in the opposite direction. If the reverse bias is made very large, a sudden breakdown in the diode will occur and the reverse current will abruptly increase, often with destructive effects.

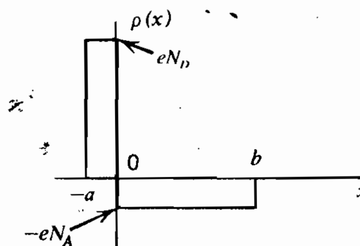
3. PROPERTIES OF THE REVERSE BIAS JUNCTION

When a reverse bias is applied to the junction, virtually all the applied voltage will appear across the depletion region, because its resistivity is much higher than that of the normal n- and p-type material. Because the effect of the reverse bias is to accentuate the potential difference across the junction, Poisson's equation [Eq. (11.14)] demands that space charge must also increase and extend a greater distance on either side of the junction. Thus, the thickness of the depletion region is also increased, extending the volume over which radiation-produced charge carriers will be collected. Practical detectors are almost always operated with a bias voltage that is very large compared with the contact potential, so that the applied voltage completely dominates the magnitude of the potential difference across the junction.

In the analysis that follows, we first assume that the semiconductor wafer in which the junction is formed is sufficiently thick so that the depletion region does not reach either surface and is contained within the interior volume of the wafer. This condition holds for *partially depleted* detectors in which some portion of the wafer thickness remains undepleted. Many semiconductor detectors are operated with sufficient reverse bias voltage so that the depletion region extends through the full wafer thickness, creating a *fully depleted* (or *totally depleted*) detector. These configurations share many of the properties derived below, except that the depletion region is obviously limited by the physical thickness of the wafer.

Some properties of the reverse bias junction can be derived if we represent the charge distribution sketched in Fig. 11.8 by the idealized distribution shown below:

$$\rho(x) = \begin{cases} eN_D & (-a < x \leq 0) \\ -eN_A & (0 < x \leq b) \end{cases}$$



Here the electron diffusion is assumed to result in a uniform positive space charge (the ionized donor sites) over the region $-a < x \leq 0$ on the n side of the junction. A corresponding negative space charge (the filled acceptor sites) resulting from hole diffusion is assumed to extend over the region $0 < x \leq b$ on the p side. Because the net charge must be zero, $N_D a = N_A b$.

Equation (11.15) applied to this case takes the form

$$\frac{d^2\varphi}{dx^2} = \begin{cases} -\dfrac{eN_D}{\epsilon} & (-a < x \leq 0) \\[2ex] +\dfrac{eN_A}{\epsilon} & (0 < x \leq b) \end{cases}$$

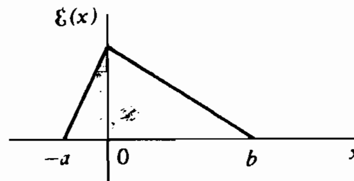
We now carry out an integration and apply the boundary conditions that the electric field $\mathscr{E} = -d\varphi/dx$ must vanish at both edges of the charge distribution:

$$\frac{d\varphi}{dx}(-a) = 0 \quad \text{and} \quad \frac{d\varphi}{dx}(b) = 0$$

The result is then

$$\frac{d\varphi}{dx} = \begin{cases} -\dfrac{eN_D}{\epsilon}(x + a) & (-a < x \leq 0) \\ +\dfrac{eN_A}{\epsilon}(x - b) & (0 < x \leq b) \end{cases}$$

The corresponding shape of the electric field $\mathscr{E} = -d\varphi/dx$ is sketched below:

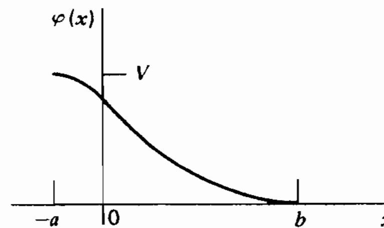


Another integration will now yield the electric potential $\varphi(x)$. The difference in potential from the n side to the p side of the junction, if we neglect the relatively small contact potential, is just the value of the applied reverse bias V. We can therefore apply the boundary conditions

$$\varphi(-a) = V \quad \text{and} \quad \varphi(b) = 0$$

The solution then takes the form

$$\varphi(x) = \begin{cases} -\dfrac{eN_D}{2\epsilon}(x + a)^2 + V & (-a < x \leq 0) \\ +\dfrac{eN_A}{2\epsilon}(x - b)^2 & (0 < x \leq b) \end{cases}$$



Since the solutions for either side of the junction must match at $x = 0$, we can write

$$V - \frac{eN_D a^2}{2\epsilon} = \frac{eN_A b^2}{2\epsilon}$$

or

$$N_A b^2 + N_D a^2 = \frac{2\epsilon V}{e}$$

Now since $N_D a = N_A b$, the expression above can be rewritten:

$$(a + b)b = \frac{2\epsilon V}{eN_A}$$

The total width of the depletion region d is the entire distance over which the space charge extends, or $d = a + b$.

For purposes of the present example, we have assumed that the n-side doping level is much higher than on the p side, so that $N_D >> N_A$. Because $N_D a = N_A b$, it follows that $b >> a$, and therefore the space charge extends much farther into the p side than the n side. Then $d \cong b$ and we can write

$$d \cong \left(\frac{2\epsilon V}{eN_A}\right)^{1/2}$$

If we had started from the opposite assumption that the p-side doping level was predominant, a similar result would be obtained except that N_A in the above expression would be replaced by N_D. A generalized solution for the thickness of the depletion region is therefore

$$d \cong \left(\frac{2\epsilon V}{eN}\right)^{1/2} \tag{11.18}$$

In this expression, N now represents the dopant concentration (either donors or acceptors) on the side of the junction that has the lower dopant level. (For surface barriers described later in this chapter, N is the dopant concentration in the bulk of the crystal.)

The resistivity ρ_d of the doped semiconductor [see Eq. (11.11)] is given by $1/e\mu N$, where μ is the mobility of the majority carrier. Equation (11.18) may thus be written

$$d \cong (2\epsilon V \mu \rho_d)^{1/2} \tag{11.19}$$

Because one often would like the largest depletion width possible for a given applied voltage, it is advantageous to have the resistivity as high as possible. This resistivity is limited by the purity of the semiconductor material before the doping process, because enough dopant must be added to override the nonuniform effects of the residual impurities. A premium is therefore placed on obtaining detectors fabricated from the highest purity material possible.

Because of the fixed charges that are built up on either side of the junction, the depletion region exhibits some properties of a charged capacitor. If the reverse bias is increased, the depletion region grows thicker and the capacitance represented by the separated charges therefore decreases. The value of the capacitance per unit area is

$$C = \frac{\epsilon}{d} \cong \left(\frac{e\epsilon N}{2V}\right)^{1/2} \tag{11.20}$$

Good energy resolution under conditions in which electronic noise is dominant depends on achieving a small detector capacitance and is thus promoted by using the largest possible applied voltage, up to the point that the detector becomes fully depleted.

The maximum electric field will occur at the point of transition between the n- and p-type material. Its magnitude is given by

$$\mathscr{E}_{max} \cong \frac{2V}{d} = \left(\frac{2VNe}{\epsilon}\right)^{1/2} \tag{11.21}$$

and can easily reach $10^6 - 10^7$ V/m under typical conditions. For partially depleted junctions, the depletion layer thickness d is proportional to $\sqrt{V}$ so that the value of $\mathscr{E}_{max}$ increases with applied voltage as $\sqrt{V}$.

The interrelation between these parameters is illustrated in the nomogram for silicon detectors given in Fig. 11.10. Also shown are scales corresponding to the ranges of various charged particles to allow selection of conditions required to produce a depletion depth that exceeds the range.

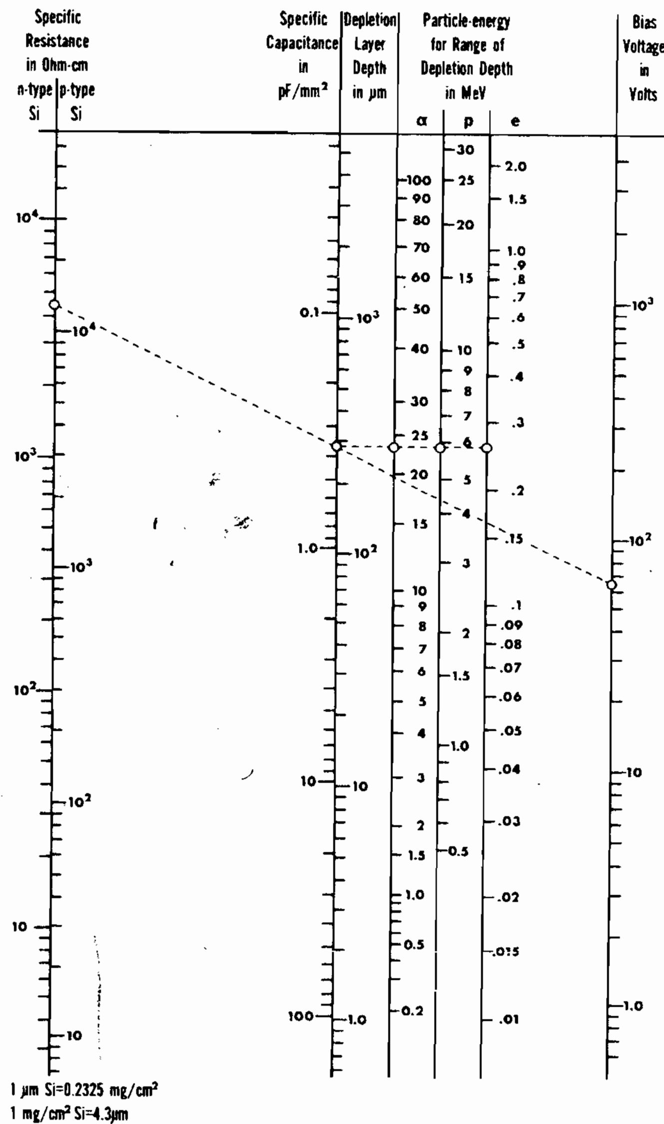


Figure 11.10 Nomogram illustrating interrelation between parameters for silicon junction detectors. (Similar to nomogram originally published by Blankenship.[26])

The maximum operating voltage for any diode detector must be kept below the breakdown voltage to avoid a catastrophic deterioration of detector properties. Commercially manufactured detectors are supplied with a maximum voltage rating that should always be strictly observed. Additional protection can be provided by monitoring the leakage current during application of the voltage (see the discussion later in this chapter).

To summarize, the reverse biased *p-n* junction makes an attractive radiation detector because charge carriers created within the depletion region can be quickly and efficiently collected. The width of the depletion region represents the active volume of the detector and is changed in partially depleted detectors by varying the reverse bias. The variable active volume of semiconductor junctions is unique among radiation detectors and sometimes is used to good advantage. The capacitance of a partially depleted detector also varies with applied voltage, and stable operation therefore requires the use of charge sensitive preamplifiers (see Chapter 17).

IV. SEMICONDUCTOR DETECTOR CONFIGURATIONS

A. Diffused Junction Detectors

One common fabrication method for semiconductor diode detectors starts with a homogeneous crystal of p-type material. One surface is treated by exposing it to a vapor of n-type impurity (typically phosphorus), which then converts a region of the crystal near the surface from p-type to n-type material. A junction is therefore formed some distance from the surface at the point at which the n- and p-type impurities reverse their relative concentration. Typical depths of the diffused n-type layer range from 0.1 to 2.0 μm. Because the n-type surface layer is heavily doped compared with the p-type original crystal, the depletion region extends primarily into the p side of the junction. Therefore, much of the surface layer remains outside the depletion region and represents a *dead layer* or window through which the incident radiation must pass before reaching the depletion region. In charged particle spectroscopy, this dead layer can be a real disadvantage because a portion of the particle energy will be lost before the active region of the detector is reached. Methods for experimentally determining its thickness are given later in this chapter.

To avoid the disadvantages of the dead layer, diffused junction detectors have been replaced in many applications by other configurations described in the following sections. Diffused junction detectors are still commercially manufactured, however, and offer some advantage over surface barrier detectors. They are somewhat more rugged and less prone to the problems that can arise due to the accumulation of oil or other foreign matter on the surface of the detector.

B. Surface Barrier Detectors

The role of the p-type material in forming the junction can be assumed by a high density of electron traps formed at the surface of an n-type crystal. The resulting depletion region behaves in much the same way as discussed earlier for a diffused junction detector. Formation of the surface states is carried out using recipes that have evolved somewhat empirically. One such set of typical procedures is described in Ref. 27. The usual treatment is etching of the surface, followed by evaporation of a thin gold layer for electrical contact. Best results are obtained if the evaporation is carried out under conditions that promote slight oxidation of the surface; the resulting oxide layer between the gold and silicon apparently plays an important role in the resulting properties of the surface barrier. Surface barriers can also be produced by starting with a p-type crystal and evaporating aluminum to form an equivalent n-type contact. The very thin dead layers that characterize surface barrier detectors are further discussed later in this chapter.

One potential disadvantage of surface barriers is their sensitivity to light. The thin entrance windows are optically transparent, and photons striking the detector surface can reach the active volume. The energy of visible light photons of 2–4 eV is greater than the bandgap energy of most semiconductors, and electron–hole pairs can therefore be produced by photon interactions. A very high noise level is produced by normal room lighting, but the vacuum enclosure required for most charged particle applications usually reduces light-induced noise to insignificant levels. The thin entrance window also makes the detector sensitive to damage from exposure to vapors, and the front surface must never be directly handled.

A cross-sectional diagram of a typical mounting arrangement for a surface barrier detector is shown in Fig. 11.11*a*. The outer housing and front surface are normally grounded, and an electrical lead from the back surface of the semiconductor wafer attaches to the center electrode of the coaxial connector at the rear. Because normal surface barriers are usually created on n-type crystals, a positive polarity voltage is required to reverse bias the junction.

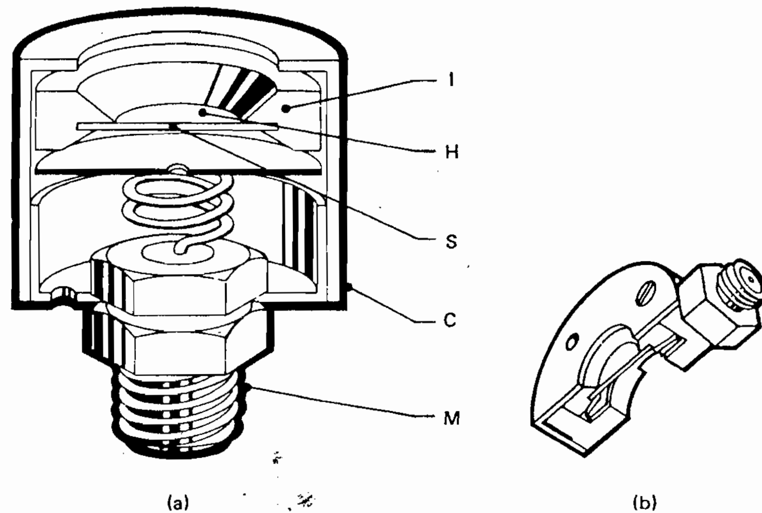


Figure 11.11 Construction and mounting of silicon junction detectors shown in cross-sectional view. (*a*) Surface barrier mount with coaxial connector (M) at rear. The silicon wafer (S) is mounted in a ceramic ring (I) with electrical contact made between either side of the junction and opposite metalized surfaces of the ring. The front surface is connected to the outer case (C) and grounded, whereas the back surface is connected to the center conductor of the coaxial connector. (*b*) Cutaway view of a transmission mount, in which both surfaces of the silicon wafer are accessible. The coaxial connector is placed at the edge of the ceramic ring. (Courtesy of EG & G ORTEC, Oak Ridge, TN.)

C. Ion Implanted Layers

An alternative method of introducing doping impurities at the surface of the semiconductor is to expose that surface to a beam of ions produced by an accelerator. This method is known as *ion implantation* and can be used to form n^+ or p^+ layers by accelerating, for example, either phosphorus or boron ions, respectively. At a fixed accelerator voltage (typically about 10 kV) monoenergetic ions are produced that have a well-defined range in the semiconductor material. By changing the energy of the incident ions, the concentration profile of the added impurity can be closely controlled. Following exposure to the ion beam, an annealing step is normally carried out to reduce the effects of radiation damage caused by the incident ions. One of the advantages of ion implantation is that the annealing temperature required (less than 500°C) is considerably lower than that needed for the thermal diffusion of dopants to form a diffused junction. The structure of the crystal is therefore less disturbed, and carrier lifetimes are not unnecessarily reduced. Compared with surface barriers, ion-implanted detectors tend to be more stable and less subject to ambient conditions. Also, they can be formed with entrance windows as thin as 34 nm silicon equivalent[28,29] and they are available commercially. A review of the use of ion implantation to form radiation detectors can be found in Ref. 30.

D. Fully Depleted Detectors

As shown by Eq. (11.18), the width of the depletion region associated with a *p-n* junction increases as the reverse bias voltage is increased. If the voltage can be increased far enough, the depletion region eventually extends across virtually the entire thickness of the silicon wafer, resulting in a fully depleted (or totally depleted) detector. Because of the several advantages this configuration presents over partially depleted detectors, the fully depleted configuration is the preferred type in most applications.

In the usual case, one side of the junction is made up of a heavily doped n^+ or p^+ layer or, alternatively, a surface barrier. The opposite side of the junction generally consists of high-purity semiconductor material that is only mildly n or p type. (Such material is often designated ν or π respectively.) The reason that high-purity material is important is reflected in Eq. (11.18). For a given applied voltage, the depletion depth is maximized by minimizing the concentration of doping impurities on the higher-purity side of the junction. Thick depletion regions can therefore only be obtained by starting from semiconductor material with the lowest possible impurity concentration. Also, with a large difference in the doping levels, the depletion layer essentially extends only into the high-purity side of the junction. The heavily doped layer can then be very thin, providing an *entrance window* for weakly penetrating radiations.

In Fig. 11.12, we assume that we have such a junction formed between a heavily doped p^+ surface layer and a high-purity n-type silicon wafer. As the reverse bias voltage applied to the detector is raised from zero, the depletion region extends further from the p^+ surface into the bulk of the wafer. For low values of the voltage, the wafer is only partially depleted and the electric field goes to zero at the far edge of the depletion region. Between this point and the back surface of the wafer, a region of undepleted silicon exists in which there is no electric field. This region then represents a very thick dead layer from which charge carriers are not collected. For all practical purposes, partially depleted detectors are therefore only sensitive to charged particles incident on the front surface.

If the applied voltage is increased further, the depletion region may be made to extend all the way to the back surface of the wafer. The voltage required to achieve this condition is sometimes called the *depletion voltage*. Its value is found by setting the depletion depth d in Eq. (11.18) equal to the wafer thickness T:

$$V_d = \frac{eNT^2}{2\epsilon}$$

When this stage is reached, a finite electric field exists all the way through the wafer, and the back dead layer thickness is reduced to that of the surface electrical contact that is employed. This condition is represented by the middle plot in Fig. 11.12. Once the wafer is

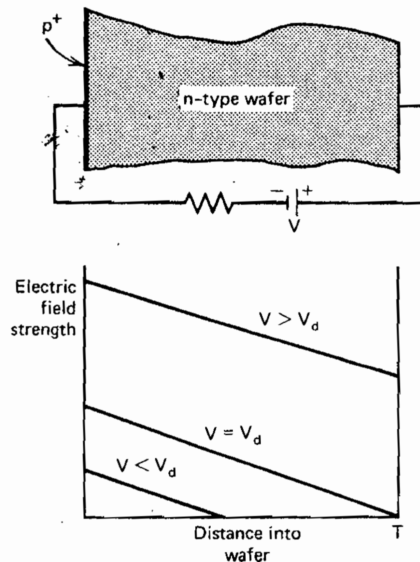


Figure 11.12 The electric field shape in a reverse bias semiconductor detector. Three plots are shown for bias voltages that are below, equal to, and above the depletion voltage V_d.

fully depleted, raising the applied voltage further simply results in a constant increase in the electric field everywhere in the wafer. At voltages much larger than the depletion voltage, the electric field profile therefore tends to become more nearly uniform across the entire wafer thickness. Under these conditions, the detector is sometimes said to be *overdepleted.* Because of the advantages of having a high electric field everywhere within the detector active volume, virtually all totally depleted detectors are operated at sufficient voltage to achieve this condition.

Figure 11.13 shows several configurations that are typical of fully depleted detectors, together with the corresponding electric field profile through the wafer. In order to deplete the wafer fully at as low a voltage as possible, one normally starts with material with the highest available purity, either n or p type. The junction is then formed by providing a heavily doped surface layer of the opposite type. This is often called the *rectifying contact.* Because of its high doping level, it also serves as an excellent blocking contact in which the minority carrier concentration is very low. In the nearly pure bulk of the wafer, however, the minority carriers are not highly suppressed and an additional blocking contact is normally provided at the opposite face of the wafer. If the high-purity silicon is mildly n type, then a thin n^+ layer is provided at this back surface. Since both materials are n type, no semiconductor junction exists at this surface. Instead, the n^+ layer provides the noninjecting conditions necessary to suppress leakage current due to minority carrier motion across the junction.

As shown on the right in Fig. 11.13, the roles of the n^+ and p^+ surface layers are reversed if one starts with high-purity material that is mildly p type. In both cases, the electric field is a maximum at the rectifying contact and decreases linearly to a minimum at the blocking contact. Shown in the center in Fig. 11.13 is the case in which intrinsic or perfectly compensated material is used for the wafer. In this case, the distinction between the two contacts disappears and the electric field is uniform throughout the entire wafer. The detector is fully depleted even for very low values of applied voltage. This *p-i-n* configuration is discussed in connection with lithium-drifted detectors in the next chapter.

Fully depleted silicon detectors are very useful as *transmission detectors* for incident particles that have sufficient energy to pass completely through the wafer. The pulse amplitude then indicates the energy lost by the incident radiation during its transit through the device. Totally depleted silicon detectors are commercially available in thicknesses from about 50 to 2000 μm. Transmission mounts of the type shown in Fig. 11.11*b* allow access to both surfaces of the wafer.

Several properties of totally depleted detectors are of primary importance. The dead layers must be as small as possible at both the front and rear surfaces of the detector if the

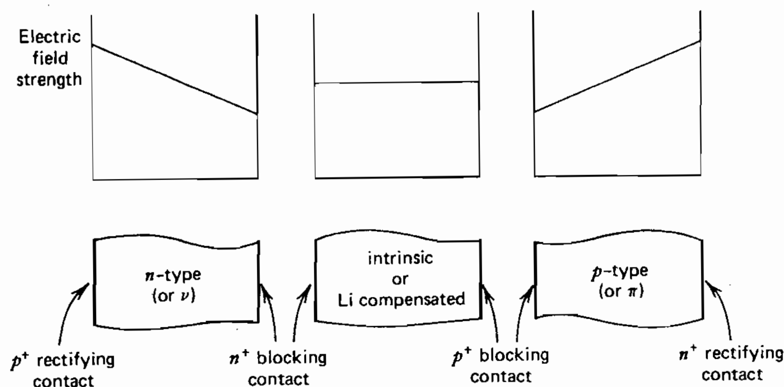


Figure 11.13 The electric field shapes for fully depleted planar semiconductor detectors of different configurations.

pulse is to indicate accurately the energy loss of the particle during its transit. An empirical test is often carried out to determine the minimum bias voltage at which these detectors are totally depleted. The pulse height from a monoenergetic source of charged particles is recorded for the particles incident on both the front and back face of the detector. When the detector is totally depleted, the pulse height should be approximately the same for either orientation. In interpreting such measurements, allowance must be made for the fact that the inherent window thicknesses of the front and back contacts of these detectors are often not the same.

In partially depleted detectors, the thickness uniformity of the crystal from which the detector is fabricated is not critical because the active volume of the detector is determined by the limited depletion depth. In fully depleted detectors, however, the wafer thickness must be kept quite uniform to avoid energy loss variations across the surface of the detector. Consequently, considerable effort is taken to provide uniform crystal wafers when totally depleted configurations are produced.

Fully depleted detectors have other advantages over partially depleted configurations in which there is an undepleted back dead layer. The finite electrical resistance of this dead layer is a source of Johnson noise that can contribute to the degradation of energy resolution. It is eliminated in a fully depleted detector by extending the depletion region all the way to the back contact. Timing properties also tend to be superior in fully depleted configurations. In a partially depleted detector, the electric field drops to zero at the edge of the depletion region. Charge carrier velocities therefore become very low in these low-field regions, slowing the rise of the signal pulse. In a totally depleted detector, the electric field can be maintained at a high value everywhere within the detector volume. Finally, some added stability results from the fact that the active volume and capacitance of a fully depleted detector are no longer functions of the applied voltage as they are in a partially depleted configuration.

The thickness of wafer that can be fully depleted using voltages short of catastrophic breakdown depends on the purity of the semiconductor. In this respect, there is a significant difference between silicon and germanium. Using ultrapure germanium described in the following chapter, a depletion thickness of several centimeters can be achieved. The impurity levels in currently available silicon are somewhat higher, and depletion thicknesses are generally limited to no more than several millimeters. Greater thicknesses in silicon are currently possible only through the use of material compensated by the lithium-drifting process described in Chapter 13.

E. Passivated Planar Detectors

The newest method of fabricating silicon junction detectors combines the techniques of ion implantation and photolithography to produce detectors with very low leakage currents and excellent operational characteristics.[31–33] Methods that were first developed in the semiconductor industry to produce integrated circuits have now been adapted successfully[34,35] to the fabrication of detectors. The techniques described below lend themselves to the batch production of multiple detectors simultaneously starting with a large-area silicon wafer, thus providing potential cost savings. These techniques can also accommodate the type of complex electrode geometry required, for example, in the silicon microstrip detectors described in Chapter 13.

The planar fabrication process generally begins with high-purity silicon that is mildly n type due to residual donor impurities. The steps in the fabrication process are shown in Fig. 11.14. After the wafer has been polished and cleaned, the surface is "passified" through the creation of an oxide layer at elevated temperature. Next, the techniques of photolithography are used to remove selectively areas of the oxide where the entrance windows of the finished detectors are to be located. The junction is then formed by converting a very

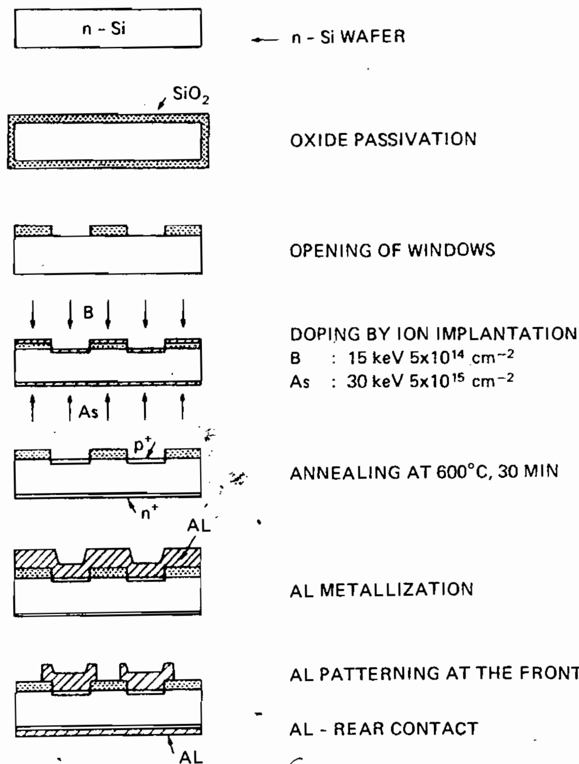


Figure 11.14 Steps in the fabrication of passivated planar silicon diode detectors. (From Kemmer.[34])

thin layer of silicon within the windows into p-type material through the implanting of acceptor ions (boron) using an accelerator. To serve as a blocking electrical contact, the rear surface of the wafer is converted into n^+ material through implantation of donor (As) ions. The radiation damage in the implanted layers is next removed through annealing at elevated temperature. Finally, aluminum is evaporated and patterned by photolithography to provide thin ohmic electrical contacts at the front and rear surfaces. The individual detectors are then separated and encapsulated.

One advantage of this planar fabrication process is that the junction edges are defined by the ion implantation pattern and can be kept within the bulk of the wafer. The oxide-passivated surface keeps leakage currents much lower than in surface barrier detectors, where the junction edge extends all the way to the edge of the wafer. Much of the leakage current in these designs then occurs where these edges are encapsulated in epoxy or similar material. Formation of the p^+ layer through ion implantation also provides planar detectors with a very thin and uniform entrance window or dead layer, an important consideration in preserving good energy resolution for the detector. The aluminized front surface is more rugged and less subject to damage compared with the gold front surface used in surface barrier fabrication.

V. OPERATIONAL CHARACTERISTICS

A. Leakage Current

When voltage is applied to a junction detector in the normal fashion, that is, to reverse bias the junction, a small current of the order of a fraction of a microampere is normally observed. The origins of this *leakage current* are related both to the bulk volume and

surface of the detector. Bulk leakage currents arising internally within the volume of the detector can be caused by either of two mechanisms.

The direction of the electric field across the depletion region is such that any majority carriers that diffuse from the normal p and n regions of the detector to the edges of the depletion region will be repelled away from the junction. However, the minority carriers in either case are attracted and will therefore be conducted across the junction. Because the minority carriers are generated continuously on both sides of the junction and are free to diffuse, a steady-state current will result that will be roughly proportional to the area of the junction. In most cases, the minority carrier current is small and is seldom an important leakage source.

A second source of bulk leakage is the thermal generation of electron–hole pairs within the depletion region. This rate will obviously increase with the volume of the depletion region and can be reduced only by cooling the material. Silicon detectors of usual dimensions have a sufficiently low thermally generated current to allow their use at room temperature, but germanium detectors, because of the lower gap energy, must always be operated at reduced temperatures.

Surface leakage effects take place at the edges of the junction where relatively large voltage gradients must be supported over small distances. The amount of surface leakage can vary greatly, depending on such factors as the type of detector encapsulation used, humidity, and any contamination of the detector surface by fingerprints, vacuum pump oil, or other condensable vapors. Guard rings analogous to those described in Chapter 5 are sometimes incorporated into the design of semiconductor diode detectors[36,37] to reduce the surface leakage, but in commercial detectors the normal approach is to rely on clean encapsulation techniques to keep the surface leakage within tolerable levels. The introduction of the planar fabrication process has allowed the production of detectors in which the junction edges are buried within the silicon wafer. As a result, leakage current is reduced to a small fraction of that typically observed in either diffused junction or surface barrier devices.

In addition to the effects on energy resolution discussed in the following section, the leakage current has another practical influence on detector operation. The bias voltage to the detector is always supplied through a large-value series resistor (R_L in Fig. 17.5) for signal isolation purposes. Therefore, the true bias voltage applied to the junction is reduced from that of the voltage source by the product of the leakage current and the series resistance. If the leakage current is large enough, the drop across the resistor can appreciably diminish the actual voltage applied to the detector, and the supply voltage must then be raised to compensate for this loss. It is therefore a fairly common practice to monitor the leakage current with an ammeter in series with the voltage supply.

Monitoring the leakage current can also detect the onset of abnormal detector behavior. During steady operation, the leakage should normally maintain a steady value, and any abrupt change or increase in the leakage current can indicate a change in detector performance, which may degrade the energy resolution. Also, it is useful to monitor the leakage current as the bias voltage is first applied to the detector. Normally, the leakage current will increase as the bias voltage is raised. However, any sudden increases can signal the approach of the breakdown of the diode, and the voltage should therefore be reduced to a lower value. Finally, the long-term behavior of the leakage current is often a useful monitor on the degree of radiation damage suffered by a given detector when used under conditions in which such damage is significant.

B. Detector Noise and Energy Resolution

Sources of electronic noise in spectroscopic measurements fall into two main categories: *series noise* and *parallel noise*. A discussion of the distinctions between these categories and their variations with operating parameters such as the choice of electronic shaping time is

postponed until Chapter 17. For the type of silicon diode detectors used for charged particle measurements, three main contributors to the electronic noise are most significant:

1. Fluctuations in the bulk generated leakage current, a component of parallel noise.
2. Fluctuations in the surface leakage current, another component of parallel noise.
3. Noise associated with series resistance or poor electrical contacts to the detector, a contributor to series noise.

The relative importance of these sources will depend on the magnitude of the leakage currents for the specific detector, the capacitance of the detector, and whether the diode is partially or fully depleted. The overall peak broadening due to electronic noise is often measured experimentally by injecting the output of a stable pulse generator into the preamplifier input while the detector remains connected. If the pulse generator has negligible spread in the amplitude of its output pulses, then the width of the corresponding "pulser peak" in the recorded spectrum is a direct measure of the electronic noise contribution.

This noise width combines in quadrature with other sources of peak broadening, such as the contributions of charge carrier statistics and fluctuations in particle energy loss in dead layers to determine the peak widths actually observed. Examples of the energy resolution attainable with semiconductor diode detectors are given in the section on alpha particle spectroscopy later in this chapter.

If trapping effects become significant in the detector, it is usually evidenced by the appearance of low-energy tails on the peaks observed from monoenergetic sources of radiation. These tails correspond to pulses in which less than the total amount of charge generated by the radiation has been collected. Because the amount of trapping varies according to the distance traveled by the carriers before reaching the collection electrodes, the amount of energy loss is variable and the full-energy peaks are spread only to the low-energy side.

C. Changes with Detector Bias Voltage

When the bias voltage and electric field are low, the pulse height from radiations that are fully stopped within the depletion layer continues to rise with applied voltage. This variation is caused by the incomplete collection of charge carriers because of trapping or recombination along the track of the incident particle. The fraction that escape collection will decrease as the electric field is increased. Similar losses to recombination are observed in a gas-filled ion chamber at low values of the electric field. Once the electric field is sufficiently high, charge collection becomes complete and the pulse height no longer changes with further increases in the detector bias voltage. This region of operation is called the *saturation region* and corresponds to the region of ion saturation in a gas-filled ion chamber.

If radiations of a single energy and type are involved, it is sometimes possible to operate the detector at a bias voltage that is short of true saturation without significant deterioration in the energy resolution, because the fraction of charge lost for each event is likely to be nearly constant. When measuring radiations of diverse energy and specific ionization, however, it is quite important to ensure that the detector is operating in the region of true saturation to avoid significant deterioration in the energy resolution. To reach the saturation region, somewhat higher electric fields are generally required as the detector undergoes radiation damage.[38]

If the electric field is made sufficiently high, multiplication effects can be induced in a semiconductor detector that are analogous to gas multiplication in proportional counters or Geiger–Mueller tubes. The multiplication arises when electrons liberated in the initial radiation interaction gain sufficient energy from the field to create further electron–hole pairs as they drift toward the collecting electrodes. These multiplication effects are discussed in some detail in Ref. 39 and are the basis of operation for silicon avalanche detectors described in Chapter 13.

D. Pulse Rise Time

Semiconductor diode detectors are generally among the fastest of all commonly used radiation detectors. A general review of theoretical and experimental work on their timing properties is given in Ref. 40. Under normal conditions, the observed pulse rise time is of the order of 10 ns or less. The detector contribution to this rise time is composed of the *charge transit time* and the *plasma time.*

The charge transit time corresponds to the migration of the electrons and holes formed by the incident radiation across the region of high electric field in the depletion region. The rise time of the output pulse is therefore limited by the time required for complete migration of these charges from their point of formation to the opposite extremes of the depletion region. These times are minimized in detectors with high electric fields and small depletion widths. In totally depleted detectors, the depletion width is fixed by the physical thickness of the silicon wafer, and therefore the transit time is decreased as the bias voltage is increased. In partially depleted detectors, however, the depletion width increases with increasing bias [see Eq. (11.18)], and therefore the effect of a larger bias voltage is to increase both the electric field and the distance over which charges must be collected. Furthermore, because the electric field is not uniform, the drift velocity of electrons and holes will vary as they move across the depletion region. The dependence of the charge transit time on bias voltage in these detectors is therefore somewhat more complicated, but it can be shown to be independent of the voltage if certain simplifying assumptions are made.[41] A derivation is given in Chapter 12 (see p. 421) of the time profile of the signal pulse attributable to charge migration in solid-state detectors in which the electric field is uniform.

For the case of a particle range that is much less than the width of the depletion region, all the charge carriers are created near one boundary. The collection time of one type of carrier corresponds to its migration across the entire depletion region and is therefore much longer than that for the other carrier. For a surface barrier on an n-type crystal, it is thus the electron collection time that dominates the time response for weakly penetrating particles.

A second component called the *plasma time* is observed when heavy charged particles, such as alpha particles or fission fragments, comprise the incident radiation. For these radiations, the density of electron–hole pairs along the track of the particle is sufficiently high to form a plasma-like cloud of charge that shields the interior from the influence of the electric field. Only those charge carriers at the outer edge of the cloud are subject to the influence of the field, and they begin to migrate immediately. The outer regions are gradually eroded away until the charges at the interior are finally subject to the applied field and also begin to drift. The plasma time is roughly defined as the time required for the charge cloud to disperse to the point where normal charge collection proceeds.

A number of theoretical models have been developed to describe the plasma erosion process;[42–47] it is predicted[42] that the plasma time should vary inversely with the electric field strength at the position of the track and increase as the cube root of the linear carrier density along the track. The effects of the plasma formation are observed to be a fixed delay of several nanoseconds between the time of track formation and the onset of the rise of the output pulse together with a slowing of the rise time of the output pulse. Measurements of the delay time with silicon surface barrier detectors,[48–53] give typical values of 1–3 ns for alpha particles, and 2–5 ns for heavy ions and fission fragments.

The actual rise time observed from a detector–preamplifier combination may also be influenced by the preamplifier properties. The time constant of the equivalent input circuit must be short if the rise time is to be held to that determined by the detector charge collection and plasma time properties only. One contributor to the input time constant is the series resistance of the undepleted region in partially depleted detectors. Therefore, fully depleted detectors in which the series resistance is largely eliminated are often favored in fast-timing situations.

E. Entrance Window or Dead Layer

When heavy charged particles or other weakly penetrating radiations are involved, the energy loss that may take place before the particle reaches the active volume of the detector can be significant. Because the thickness of the dead layer includes not only the metallic electrode but also an indeterminate thickness of silicon immediately beneath the electrode in which charge collection is inefficient, the dead layer can be a function of the applied voltage. Its effective thickness must often be measured directly by the user if accurate compensation is to be made.

The simplest and most frequently used technique is to vary the angle of incidence of a monoenergetic charged particle radiation. When the angle of incidence is zero (i.e., perpendicular to the detector surface), the energy loss in the dead layer is given by

$$\Delta E_0 = \frac{dE_0}{dx} t \tag{11.22}$$

where t is the thickness of the dead layer. The energy loss for an angle of incidence of θ is given by

$$\Delta E(\theta) = \frac{\Delta E_0}{\cos\theta} \tag{11.23}$$

Therefore, the difference between the measured pulse height for angles of incidence of zero and θ is given by

$$E' = (E_0 - \Delta E_0) - (E_0 - \Delta E(\theta))$$

$$E' = \Delta E_0 \left(\frac{1}{\cos\theta} - 1 \right) \tag{11.24}$$

If a series of measurements are made as the angle of incidence is varied, a plot of E' as a function of $(1/\cos\theta - 1)$ should be a straight line whose slope is equal to ΔE_0. Using tabular data for dE_0/dx for the incident radiation, we can calculate the dead layer thickness from Eq. (11.22).

One possible flaw in this method involves the assumption that the energy loss through the dead layer depends only on the total path length traversed and not on the relative orientation of the particle path with respect to the detector axis. There is some evidence that recombination should be more severe for particle paths parallel to the direction of the electric field in the detector compared with paths perpendicular to the field. This recombination would tend to cause a lower than expected response for paths near normal incidence and should be evidenced by a curvature in the plot described above.

The thinnest dead layers are produced in semiconductor detectors of the ion implanted or surface barrier types. Typical values of 100 nm of silicon equivalent correspond to an energy loss of 4 keV for 1 MeV protons, 14 keV for 5 MeV alpha particles, and several hundred keV for fission fragments. Because variations in this energy loss due to straggling or variable angle of incidence will potentially detract from energy resolution, thin dead layers are quite important in high-resolution charged particle spectroscopy. Using special techniques, dead layers of less than 30 nm have been successfully fabricated.[54]

F. Channeling

In crystalline materials, the rate of energy loss of a charged particle can depend on the orientation of its path with respect to the crystal axes.[55] Particles that travel parallel to crystal planes can, on the average, show a rate of energy loss that is lower than that for particles directed in some arbitrary direction. Therefore, these "channeled" particles can penetrate

significantly farther through the crystal. The effects are particularly significant for thin totally depleted detectors because the amount of energy deposited is then dependent on the orientation of the crystal planes with respect to the particle direction. To minimize the tendency for incident particles to channel, detectors are normally fabricated from silicon cut so that the ⟨111⟩ crystal orientation is perpendicular to the wafer surface.

Channeling can affect the recorded pulse height even in situations in which the particle is fully stopped within the active volume. Nuclear collisions are less probable for channeled particles and therefore the pulse height defect for heavy ions (discussed later in this chapter) may be reduced.

G. Radiation Damage

The proper operation of any semiconductor detector depends on the near perfection of the crystalline lattice to prevent defects that can trap charge carriers and lead to incomplete charge collection. Any extensive use of these detectors, however, ensures that some damage to the lattice will take place because of the disruptive effects of the radiation being measured as it passes through the crystal. While the energy that goes into the creation of electron–hole pairs leads to fully reversible processes that leave no damage, nonionizing energy transfers to the atoms of the crystal lead to irreversible changes. The effects tend to be relatively minor for lightly ionizing radiations (beta particles or gamma rays) but can become quite significant under typical conditions of use for heavy charged particles. For example, prolonged exposure of silicon surface barrier detectors to heavy ions or fission fragments will lead to a measurable increase in the leakage current and a significant loss in energy resolution of the detector. With extreme radiation damage, multiple peaks may appear in the pulse height spectrum recorded for monoenergetic particles. Furthermore, the time characteristics of the detector may be degraded even at doses that are too low to show measurable spectral effects.[56] In silicon microstrip detectors of the type discussed in Ch. 13, radiation damage can lead to a decrease in the inter-strip resistance and a loss of spatial resolution.[57]

The radiation-induced damage can be classified into the two categories of bulk and surface effects.[58] The most fundamental type of bulk radiation damage is the *Frenkel defect*, produced by the displacement of an atom of the semiconductor material from its normal lattice site. The vacancy left behind, together with the original atom now at an interstitial position, constitutes a trapping site for normal charge carriers. These are sometimes called *point defects* to distinguish them from more complex "clusters" of crystalline damage that are formed along the track of a primary "knock-on" atom if sufficient energy is transferred. Gamma rays and electrons with energy of a few MeV or less create only point defects, whereas heavy charged particles of equivalent energy are generally more damaging because they also form clusters. The number of Frenkel defects produced by a fission fragment is estimated to be about 100–1000 times greater than that produced by an alpha particle.[59] At the other extreme, an incident electron or beta particle requires a minimum of about 145 keV to produce a defect, and very little damage is observed for electrons whose energy is much below 250 keV.[60] The severity of damage to be anticipated is therefore a strong function of the nature of the radiation involved. When enough of these defects have been formed, carrier lifetime and charge collection efficiency are reduced and the energy resolution of the detector is degraded due to fluctuations in the amount of charge lost. Some minor annealing of the radiation damage can occur[61] over long periods of time (see Fig. 11.17), but for all intents and purposes, the damage is permanent.

The increase in leakage current appears to be more directly related to surface effects[59] and also contributes to a loss of detector resolution from the corresponding increase in leakage current fluctuation. In devices such as silicon microstrip detectors that include

oxide passivation layers, the surface effects are closely related to the ionization created within the oxide and its trapping at interfaces. For penetrating radiations including gamma rays or neutrons, the damage is generally distributed throughout the detector and the direction of incidence of the radiation has little effect. For electrons or charged particles, however, the orientation with respect to the detector is important. Irradiation of the front (or gold) surface of totally depleted detectors requires exposures that are several orders of magnitude less than those needed to produce the same effects by irradiation of the back (or aluminum) contact.[62]

For silicon surface barriers irradiated on the gold or front surface, various data have been published on the integrated flux of charged particles required to produce a significant deterioration in detector performance. Although subject to a great deal of variability, depending on the specifics of each experiment, serious changes appear to take place for irradiations of about 10^{14} fast electrons/cm^2 (Ref. 60), 10^{12} to 10^{13} protons/cm^2 (Refs. 62 and 63), 10^{11} alpha particles/cm^2 (Ref. 64), and about 3×10^8 fission fragments/cm^2 (Ref. 59). Exposure to fast neutron fluxes of about 3×10^{11} neutrons/cm^2 (Ref. 65) and gamma-ray doses of about 10^6 R (Ref. 66) are also sufficient to lead to significant performance degradation. In general, fully depleted detectors are less sensitive than partially depleted devices because the average electric field throughout the detector is somewhat higher. The effects of radiation-induced charge trapping are minimized when the highest possible electric field is present in the active volume.

An effect known as *type inversion*[67,68] has been observed to occur in high resistivity n-type silicon after prolonged exposure to fast neutrons or high-energy particles with integrated fluence of about 10^{13}/cm^2. The effective concentration of donors gradually decreases with exposure, until a transition to p-type behavior is observed. Some models for this change[67,69] postulate the radiation-induced formation of deep acceptor levels, close to the center of the bandgap, that become electrically active when voltage is applied to the detector. In detectors with a p^+-ν-n^+ structure (see Fig. 11.13), an interesting consequence of the type inversion is the effective shift of the rectifying contact from one side of the detector to the other. The device may continue to function after the inversion with the same polarity of applied voltage, but eventually the growing concentration of acceptors raises the voltage level required for full depletion to levels beyond those causing breakdown.

H. Energy Calibration

When applied to the measurement of fast electrons or light ions such as protons or alpha particles, semiconductor diode radiation detectors respond very linearly, and the energy calibration obtained for one particle type is very close to that obtained using a different radiation type. Some observations show that there is a small difference in the pulse height observed for protons and alpha particles of the same energy,[70,71] but such differences are usually on the order of 1% or less. It is not clear whether these differences reflect a fundamental difference in the ionization energy ϵ for various particles or whether other factors related to the modes of energy loss may be responsible. For any application in which an absolute energy calibration of about 1% or less is required, it is always best to calibrate the detector using the same type of particle that is involved in the measurement itself.

The most common calibration source is the alpha-emitting isotope ^{241}Am. This isotope emits alpha particles of 5.486 MeV (85%) and 5.443 MeV (13%), and a representative pulse height spectrum is illustrated in Fig. 11.15. An accurate calibration of the energy scale requires that account be taken of the energy loss of these alpha particles in the source itself, in any intervening material between the source and detector, and in the window or dead layer of the detector.

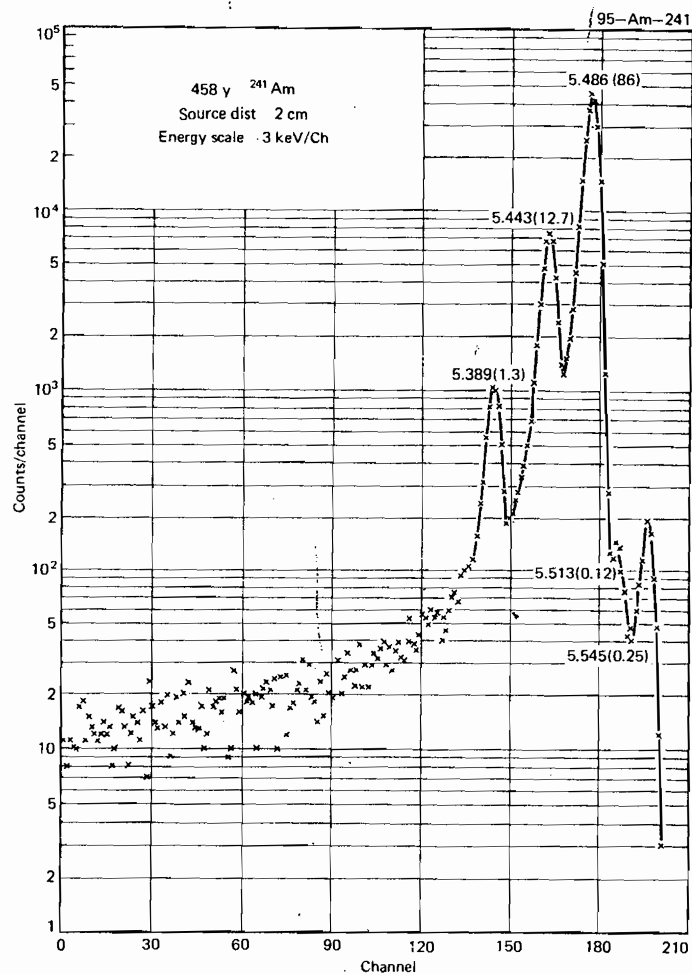


Figure 11.15 Upper portion of the ^{241}Am alpha spectrum as recorded by a high-resolution surface barrier detector. (From Chanda and Deal.[72])

I. Pulse Height Defect

The response of semiconductor detectors to very heavy ions such as fission fragments is less straightforward. There is abundant evidence that the pulse height observed for heavy ions is substantially less than that observed for a light ion of the same energy. The *pulse height defect* is defined in units of energy as the difference between the true energy of the heavy ion and its apparent energy, as determined from an energy calibration of the detector obtained using alpha particles.

A plot of the effects of the pulse height defect is given in Fig. 11.16. Measurements of the pulse height defect for surface barriers when irradiated by fission fragments[74] show that a value as large as 15 MeV for the defect is possible, compared with an average energy

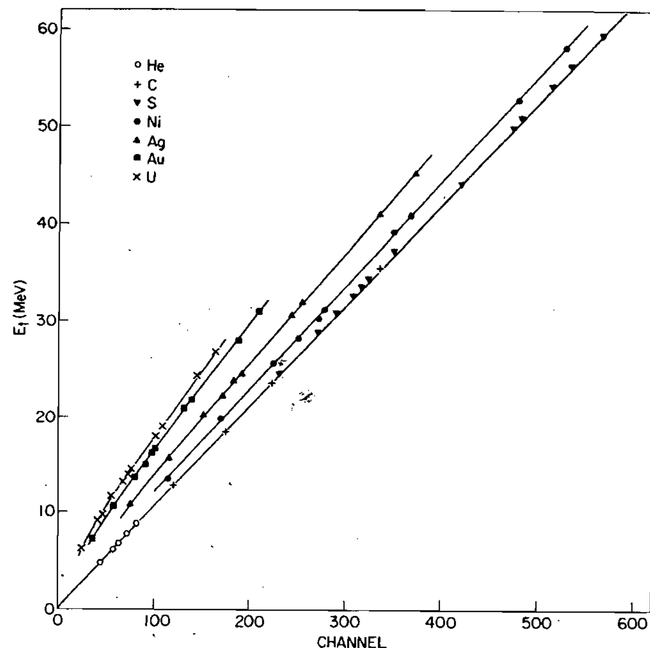


Figure 11.16 The true energy of various ions versus the pulse height channel number from a silicon surface barrier detector. (From Wilkins et al.[73])

of about 80 MeV. Other measurements[75–77] using fragments that were separated in mass and energy obtained smaller defect values of 3–7 MeV.

Analysis has shown[73,78–80] that three separate phenomena contribute to the observed pulse height defect. The first and simplest contributor is the energy loss of the ion in the entrance window and dead layer of the detector. The magnitude of the contribution to the pulse height defect can be calculated from the stopping power of the ion and measurements of the dead layer thickness using methods discussed earlier. Because heavy ions such as fission fragments show maximum dE/dx at the start of their range, whereas light ions such as alpha particles show the reverse behavior, the fractional energy loss in the dead layer will be much more significant for the heavy ions.

A second contributor to the pulse height defect involves the tendency for heavy ions to lose energy by means other than simple electronic collisions. As the velocity of the ion decreases, nuclear collisions become important, and recoil nuclei are the direct result of such interactions. Because of the low velocity of these recoil nuclei, the probability of electronic interactions is reduced and a net decrease in the efficiency with which electron–hole pairs are produced is observed. The contribution of nuclear collisions increases with the effective charge on the ion and is therefore most significant for heavy ions. A calculation of the magnitude of this effect has been carried out by Haines and Whitehead[81] from basic theories of ion slowing down in solids. An experimental study of the effect in surface barriers is described in Ref. 82.

A third factor in the pulse height defect involves the high rate of electron-hole recombination expected in the dense plasma created along the ion track, particularly near its end.[74,83,84] The magnitude of the recombination would be expected to decrease with increasing bias voltage and may also depend on the relative orientation of the particle path

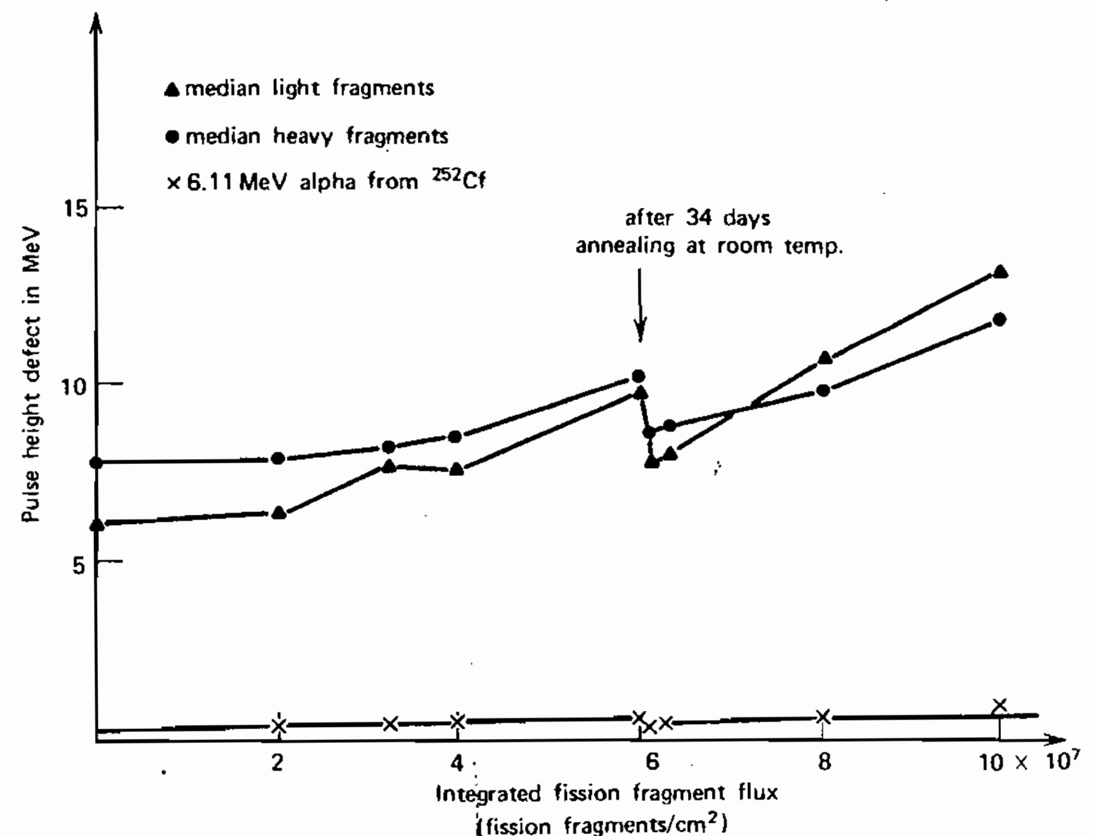


Figure 11.17 Dependence of measured pulse height defect versus fission fragment exposure for a heavy ion silicon detector. (From Bozorgmanesh.[85])

with respect to the electric field within the detector. A practical method of reducing the pulse height defect is to minimize the effect of recombination by creating the largest possible electric field within the detector.

Because the effects of trapping and recombination are influenced by radiation damage within the detector, it should be anticipated that the pulse height defect may increase with normal use. Figure 11.17 shows the measured pulse height defect for a silicon heavy ion detector as a function of integrated fission fragment flux. The general trend toward increasing pulse height defect is evident, together with the effects of partial annealing of the radiation damage over several weeks following removal of the detector from the flux.

VI. APPLICATIONS OF SILICON DIODE DETECTORS

A. General Charged Particle Spectroscopy

Since their development as practical detectors in the early 1960s, silicon diodes have become the detectors of choice for the majority of applications in which heavy charged particles are involved. Some of the more common applications involving the spectroscopy of alpha particles and fission fragments and the measurement of energy loss of charged particles in transmission detectors are discussed in the following section. Most of the conclusions carry over into other charged particle spectroscopy involving protons, deuterons, or other heavy ions.

Silicon diode detectors are sometimes used for the measurement of beta particles and fast electrons, particularly as thin totally depleted transmission detectors. However, because it is more common to use ion-drifted detectors for this purpose, a discussion of the response of semiconductor detectors to fast electrons will be postponed until Chapter 13.

Compared with competing techniques, the use of semiconductor detectors often provides advantages in a number of key areas. These include exceptionally good energy

resolution, good stability and freedom from drift, excellent timing characteristics, very thin entrance windows, and simplicity of operation. The relatively small size can be an advantage in some situations but is also a limitation in those applications in which a detector with a large surface area is required. Silicon diodes are commercially available with surface area up to 20 cm^2, but the corresponding large capacitance results in a poorer energy resolution than is attainable with smaller detectors. More usual sizes range from 1 to 5 cm^2. Depletion depths up to 5 mm can be obtained commercially in some special configurations, but the more common detectors are limited to a depletion depth of 1 mm or less.

In the event that the detector depletion depth is greater than the range of the incident heavy ions, the response of the detector is very simple. For monoenergetic incident particles, only a single full-energy peak is observed because there are no competing processes that can significantly scatter out the ion or otherwise lead to partial energy deposition.† For fully depleted detectors, the depletion depth is given simply by the thickness of the silicon wafer. In partially depleted detectors, the depletion depth increases with applied bias and is therefore limited by the maximum detector bias that can be applied without risking detector breakdown. The maximum bias and the corresponding depletion depth are normally provided as a specification by the detector manufacturer. If a measurement of the radiation energy is not required, simple counting of charged particle radiation can be carried out with semiconductor detectors whose depletion depth is less than the range, provided the energy deposited within the depletion region is sufficiently high to generate a pulse that lies above the noise level of the instrumentation system.

B. Alpha Particle Spectroscopy

Silicon diodes operated at room temperature are near-ideal detectors for alpha particles and other light ions. Because of the wide availability of convenient monoenergetic sources of alpha particles, the performance of semiconductor detectors conventionally is tested by recording the pulse height spectrum from such sources. The most common of these is ^{241}Am, and the corresponding alpha spectrum is widely used for comparison of the energy resolution of solid-state detectors. A representative spectrum taken with a detector of good resolution is shown in Fig. 11.15.

With alpha particles in this energy range (5.486 MeV), the noise contribution of the preamplifier and other electronic components can be much smaller than the inherent energy resolution of the detector itself. In this case, one might expect that the statistics of charge carrier formation would limit the energy resolution that is achievable. This limiting resolution can be calculated beginning with Eq. (4.15) in terms of the Fano factor F and the average number of charge carriers per pulse N, or the energy of the alpha particle E and the ionization energy ϵ

$$R_{\text{lim}} = 2.35\sqrt{\frac{F}{N}} = 2.35\sqrt{\frac{F\epsilon}{E}} \tag{11.25}$$

In this equation, R is expressed as a fractional (or percentage) energy resolution. The convention with silicon semiconductor detectors is to instead quote the FWHM in units of energy, so we multiply both sides of the above equation by the alpha particle energy E

$$\text{FWHM}]_{\text{lim}} = 2.35\sqrt{FE\epsilon} \tag{11.26}$$

†For incident ions with relatively high energies, some complications in the response can be observed due to nuclear reactions induced by the incident particle in the material of the detector. One example occurs in silicon detectors due to the inelastic scattering of light ions from ^{28}Si nuclei in the detector that are followed by the emission of a 1.78 MeV gamma ray. Since the probability is high that this gamma ray escapes from the detector, a satellite peak located 1.78 MeV below the full-energy peak can sometimes be observed.[86] In the case of incident protons, the probability for this process remains lower than 0.1% per incident particle for energies below 10 MeV.

Putting in values for silicon of $F = 0.11$, $\epsilon = 3.62$ eV, and evaluating at $E = 5.486$ MeV, we obtain

$$\mathrm{FWHM}]_{\mathrm{lim}} = 3.47 \text{ keV}$$

for the predicted statistical limit. Under very carefully controlled conditions, an energy resolution for alpha particles of 8 keV has been demonstrated[87] for passivated planar silicon detectors. In commercially available small-size detectors, however, the energy resolution that is achievable in practice tends to be no better than about 10 keV. The discrepancy from the predicted statistical limit represents the additional contributions to peak broadening that are still significant in these detector systems.

One of these contributions is due to the fact that a small portion of the alpha particle energy is transferred to recoil nuclei rather than to electrons. These low-energy recoil nuclei lose their energy in quasielastic collisions with surrounding atoms and form almost no additional electron–hole pairs. If the fraction of energy lost in this manner were constant for each alpha particle, there would be no effect on energy resolution. However, this energy loss is subject to large fluctuation since it is influenced by a few relatively large events. It has been estimated[88] that the FWHM contribution due to these fluctuations in silicon amount to about 3.5 keV for 6 MeV alpha particles.

Other significant contributions to peak broadening are the effects of incomplete charge collection and variations in the energy lost by the particle in dead layers at the detector surface.[89] For lower charged particle energies and detectors of large capacitance, the electronic noise can also be a significant contribution. The noise level from the detector–preamplifier–amplifier combination is dominated by fluctuations in the detector leakage current, the inherent preamplifier noise, and the characteristics of the FET used in the input stage of the preamplifier.

Since all these sources of peak broadening are normally independent, the square of each FWHM that would theoretically be observed for each source alone can be summed together to give the square of the overall FWHM.

C. Heavy Ion and Fission Fragment Spectroscopy

The energy measurement of fission fragments or other ions of large mass involves several special concerns. Most stem from the high density of charge carriers that are created along the track of these ions. Recombination of electron–hole pairs is accentuated, and the detector may require a higher bias voltage than that required to saturate the signal from alpha particles. The pulse height defect discussed previously is also accentuated by this high carrier density, complicating the energy calibration procedures. Also, the prolonged exposure of detectors to heavy ions or fission fragments creates rapid performance deterioration due to radiation damage in the detector.

Detector manufacturers generally offer some silicon detectors especially tailored for heavy ion spectroscopy. They are designed to minimize the problems of slow rise time and pulse height defect caused by the high specific energy loss along the particle track. The most effective step is to ensure that the electric field is as high as possible. One approach is to use thin slices of low-resistivity silicon to prepare totally depleted detectors that require a large value of the reverse bias to become fully depleted. Special methods of preparing contacts can also be used[90] to allow extreme overbiasing (beyond the point of full depletion), which can raise the field magnitude by as much as a factor of 20 over its value at the depletion voltage.[91] The silicon used for heavy ion detectors should also have large carrier lifetimes to help reduce the pulse height defect due to recombination.

Typical measurements[92–94] of the response function of silicon diodes to monoenergetic heavy ions show an asymmetric peak with significant tailing toward the low energy side. The cause of the tailing is probably related to fluctuations in the energy loss due to the