V. Semiconductor Detectors

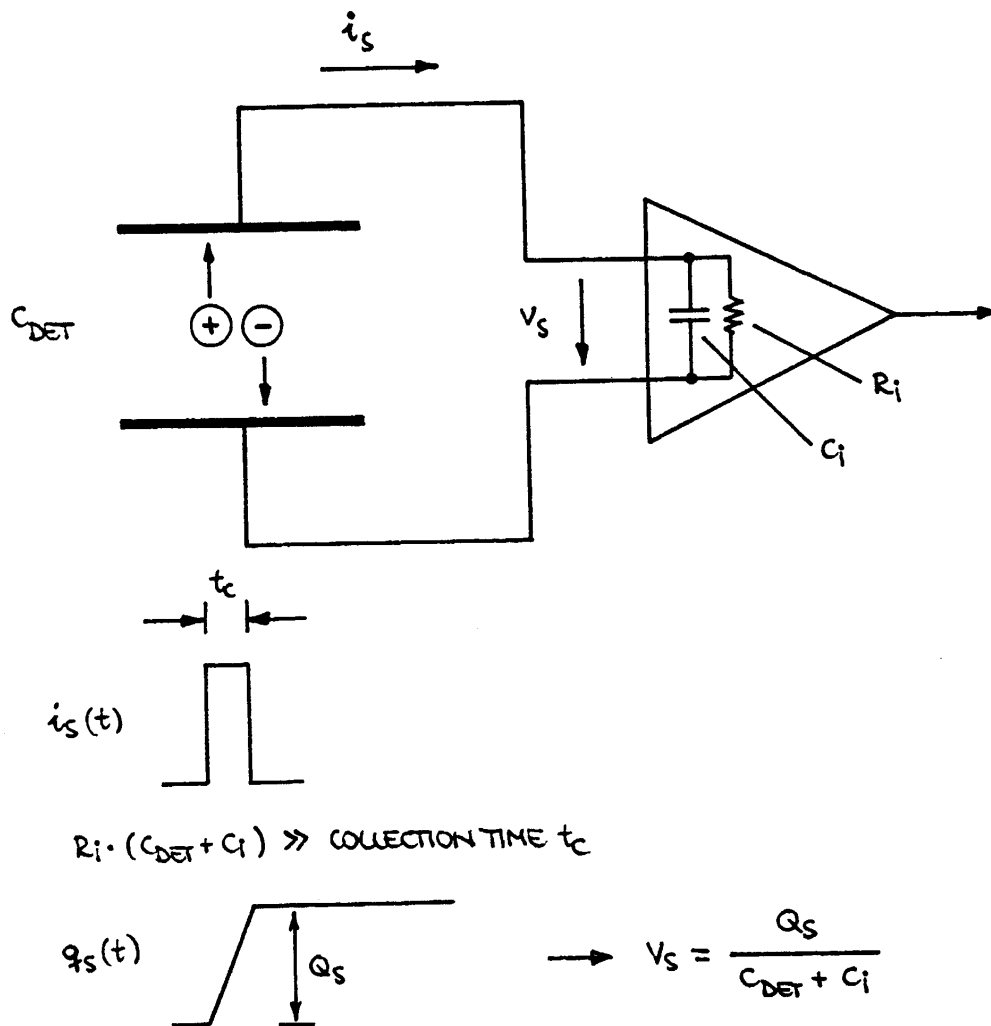
V.1. Principles

Semiconductor Detectors are Ionization Chambers

Detection volume with electric field

Energy deposited $\rightarrow$ positive and negative charge pairs

Charges move in field $\rightarrow$ current in external circuit (continuity equation)



Ionization chambers can be made with any medium that allows charge collection to a pair of electrodes.

Medium can be gas
liquid
solid

Crude comparison of relevant properties

	gas	**liquid**	**solid**
density	low	moderate	high
atomic number Z	low	moderate	moderate
ionization energy ε_i	moderate	moderate	low
signal speed	moderate	moderate	fast

Desirable properties:

low ionization energy ⇒ 1. increased charge yield dq/dE

2. superior resolution

$$\frac{\Delta E}{E} \propto \frac{1}{\sqrt{N}} \propto \frac{1}{\sqrt{E/\varepsilon_i}} \propto \sqrt{\varepsilon_i}$$

high field in detection volume

⇒ 1. fast response

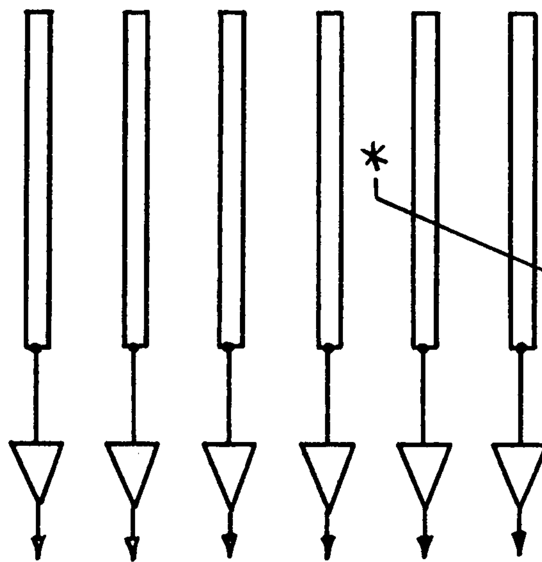
2. improved charge collection efficiency (reduced trapping)

In addition to energy measurements, semiconductor detectors allow precision position sensing.

Resolution determined by precision of micron scale patterning of the detector electrodes (e.g. strips on 50 μm pitch).

Two options:

Binary Readout



to discriminators

Position resolution determined directly by pitch

$$\sigma_x = pitch/\sqrt{12}$$

Analog Readout

Interpolation yields resolution < pitch

Relies on transverse diffusion

$$\sigma_x \propto \sqrt{t_{coll}}$$

e.g. in Si

$t_{coll} \approx$ 10 ns

$\Rightarrow$ σ_x = 5 μm

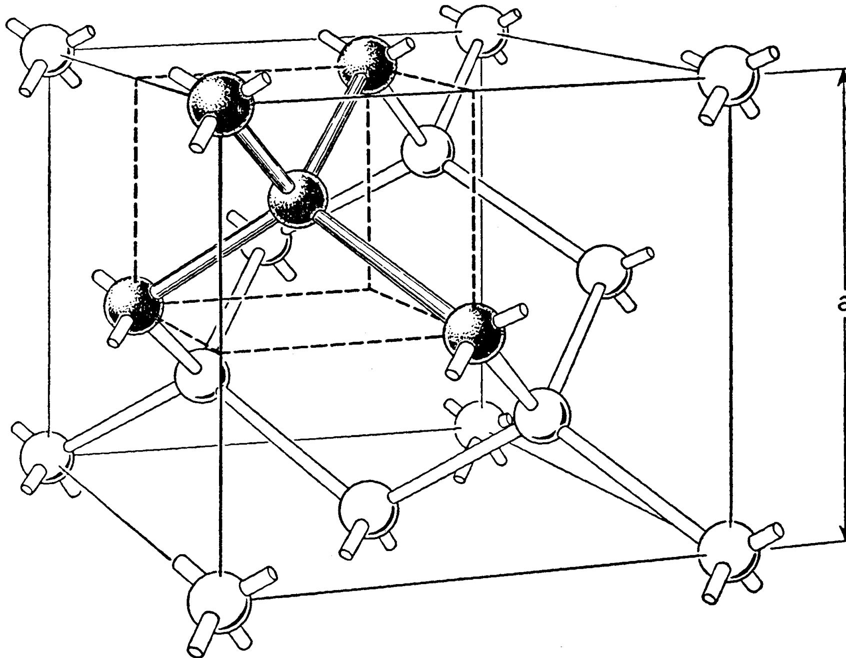
Interpolation precision depends on S/N and p

p = 25 μm and S/N=50

$\Rightarrow$ 3 – 4 μm resolution

Semiconductor Crystals

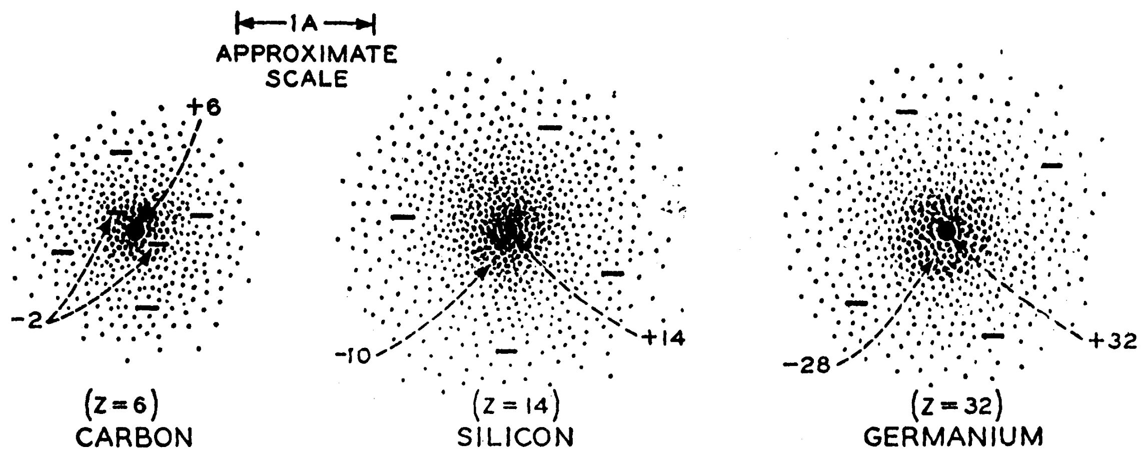
Lattice structure of diamond, Si, Ge ("diamond lattice")



(from Schockley)

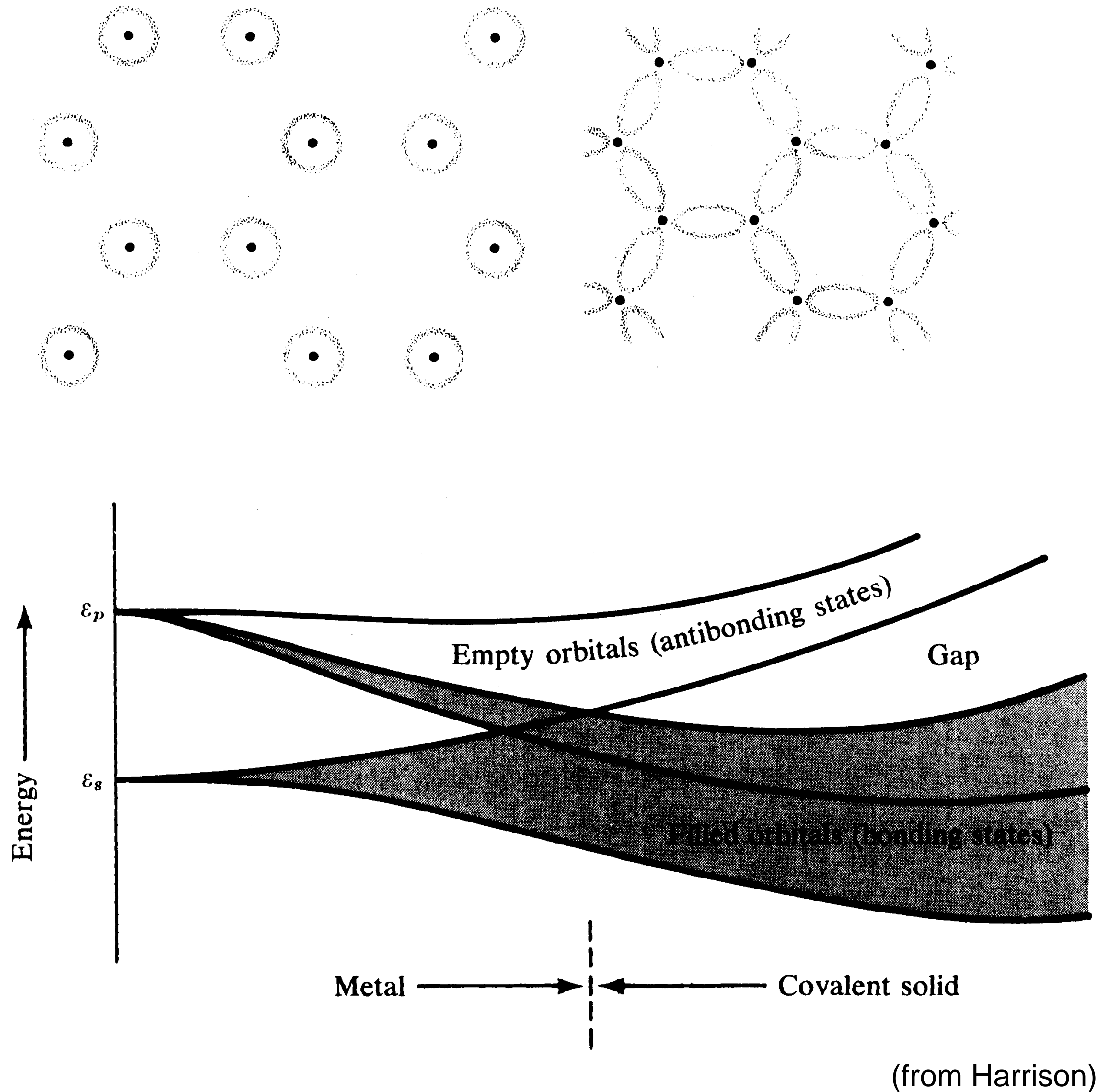
dimension a:	lattice constant	Diamond:	3.56 Å
		Ge:	5.65 Å
		Si:	5.43 Å

Extent of wavefunctions of typical constituent atoms:



(from Schockley)

When isolated atoms are brought together to form a lattice, the discrete atomic states shift to form energy bands:



(from Harrison)

Filled band formed by bonding states: $\Psi = \Psi_a + \Psi_a$

(Ψ_a = wavefunction of individual atom)

Empty band formed by anti-bonding states: $\Psi = \Psi_a - \Psi_a$

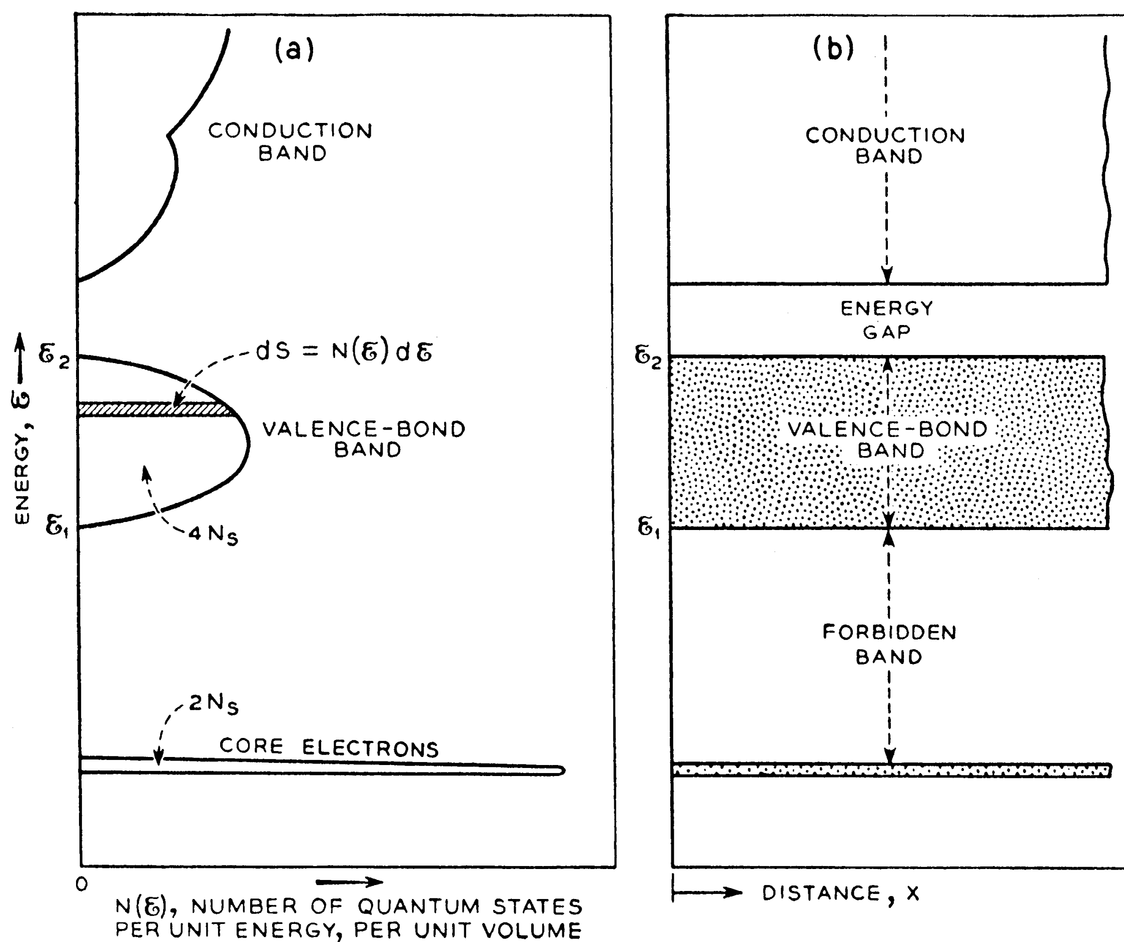
(vanishing occupancy at mid-point between atoms)

Each atom in the lattice contributes its quantum states to each band:

> The number of quantum states in the band is equal to the number of states from which the band was formed.

The bands are extended states, i.e. the state contributed by an individual atom extends throughout the crystal.

Energy band structure

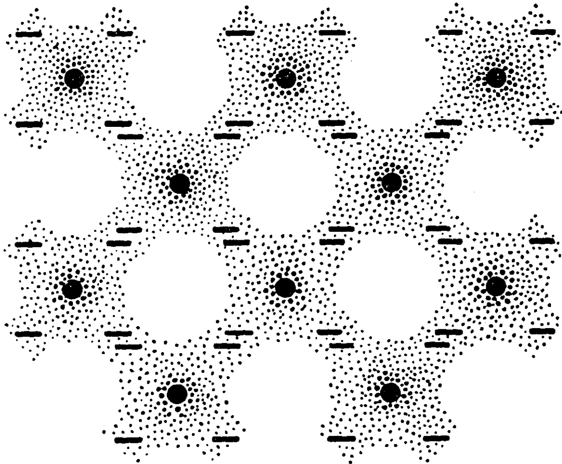


(from Schockley)

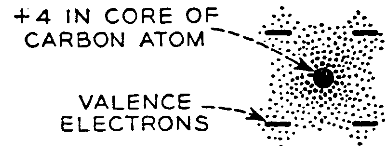
Typical band gaps (valence – conduction band)

Ge	0.7 eV	GaAs	1.4 eV
Si	1.1 eV	Diamond	5.5 eV

At 0K all electrons occupy bonding states, completely filling the valence band.

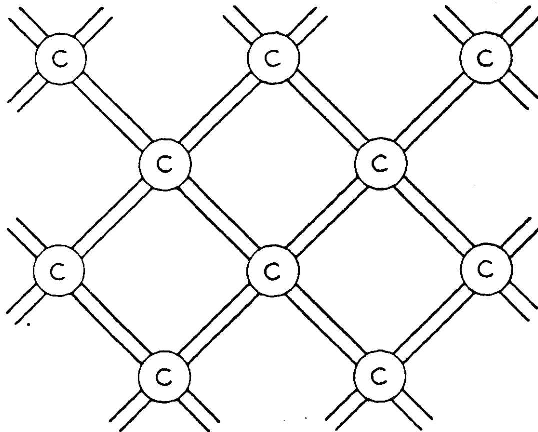


(a) ELECTRON PAIR BONDS

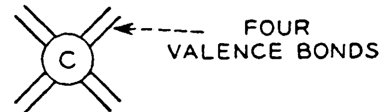


EACH ATOM, WITH THE CHARGE OF ITS SHARE OF VALENCE-BOND ELECTRONS, IS ELECTRICALLY NEUTRAL.

(b)



(C) PLANE DIAGRAM OF DIAMOND LATTICE WITH BONDS REPRESENTED BY LINES



(from Schockley)

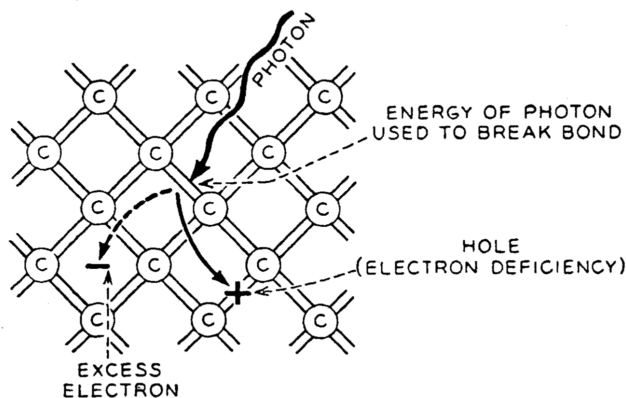
If an electric field is applied to the crystal, no current can flow, as this requires that the electrons acquire energy, which they can't, as no higher energy states are available in the valence band.

If energy is imparted to a bond by incident radiation, for example a photon, the bond can be broken,

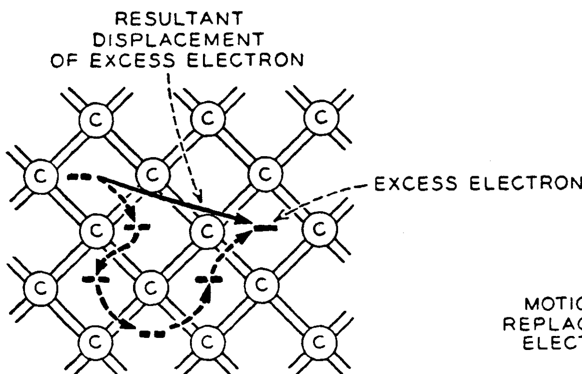
- exciting an electron into the conduction band and
- leaving back a vacant state in the valence band, a “hole”.

The electron can move freely in its extended state.

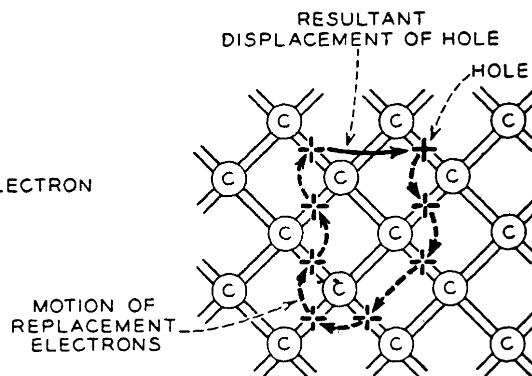
The hole can be filled by an electron from a nearby atom, thereby moving to another position.



(a) PRODUCTION OF A HOLE-ELECTRON PAIR BY A PHOTON



(C) RANDOM MOTION OF AN EXCESS ELECTRON



(d) RANDOM MOTION OF A HOLE

The motion of the electron and hole can be directed by an electric field.
Holes can be treated as positive charge carriers just like the electrons, although they tend to move more slowly as hole transport involves sequential transition probabilities (the wavefunction overlap of the hole and its replacement electron).

Energy required for creation of an electron-hole pair

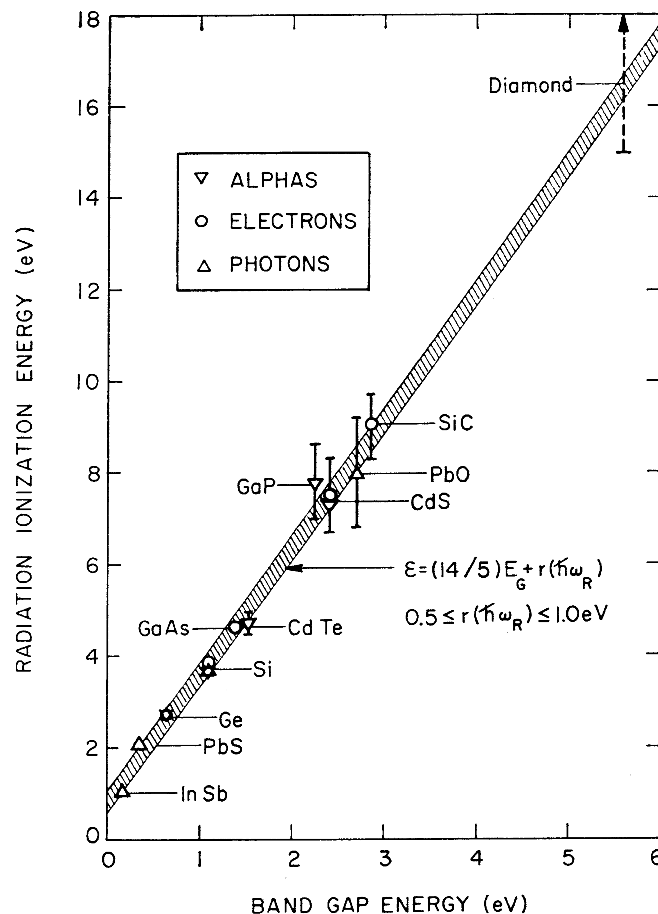
Ionization Energy > Band Gap

Formation of e-h pair requires both …

1. Conservation of energy
2. Conservation of momentum

⇒ additional energy excites phonons

Remarkably, $\varepsilon_i / E_g \approx$ const for all materials, types of radiation.



C.A. Klein, J. Applied Physics **39** (1968) 2029

Ionization energy in solids is proportional to the band gap

small band gap	⇒	~ conductor
		electric field small
		DC current >> signal current
large band gap	⇒	insulator
		high electric field
		small signal charge + small DC current
		example: diamond
moderate band gap	⇒	semiconductor
		high electric field
		"large" signal charge small DC current, but "*pn*-junction" required.
		examples: Si, Ge, GaAs

Formation of a High-Field Region

The conduction band is only empty at 0K.

As the temperature is increased, thermal excitation can promote electrons across the band gap into the conduction band.

Pure Si: carrier concentration ~ 10^{10} cm^{-3} at 300K
(resistivity ≈ 400 kΩ·cm)

Since the Si lattice comprises $5 \cdot 10^{22}$ atoms/cm^3, many states are available in the conduction band to allow carrier motion.

In reality, crystal imperfections and minute impurity concentrations limit Si carrier concentrations to ~10^{11} cm^{-3} at 300K.

This is too high for use in a simple crystal detector.

> A crystal detector is feasible with diamond, but the charge yield is smaller due to the larger band gap.

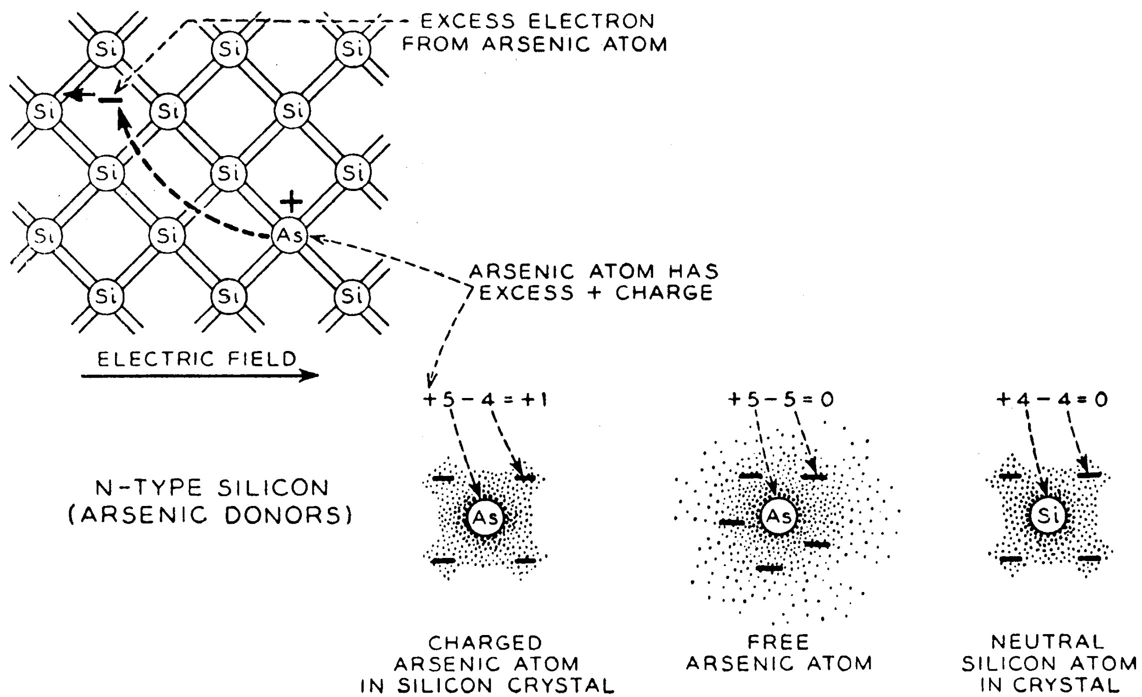
High-field region with low DC current in semiconductors
is most easily achieved utilizing a p-n junction.

⇒ Introduction of impurities to control conductivity.

The conductivity of semiconductors can be controlled by introducing special impurities.

required concentrations: $\sim 10^{12} - 10^{18}$ cm^{-3}

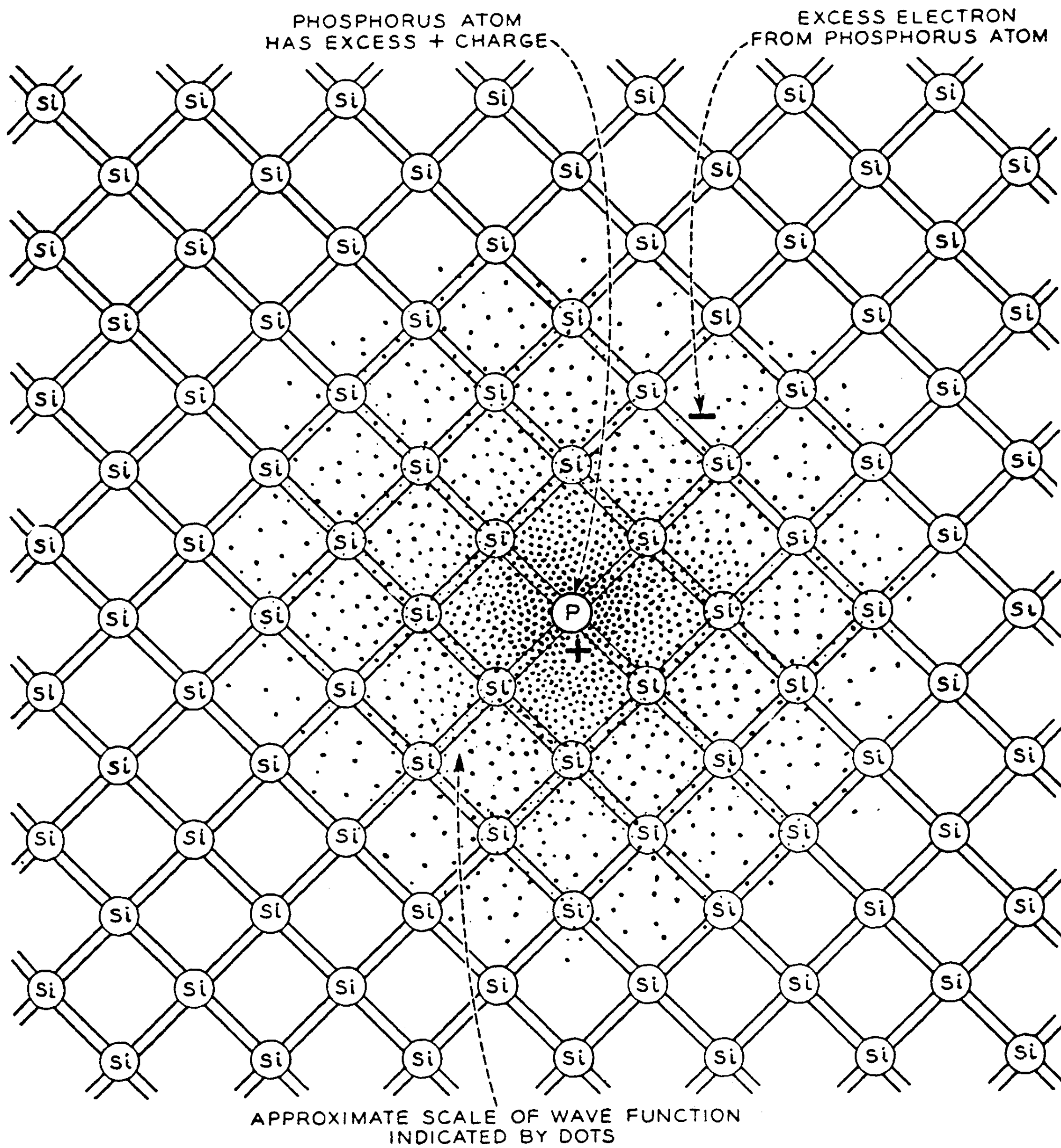
Replacing a silicon atom (group 4 in periodic table, i.e. 4 valence electrons) by an atom with 5 valence electrons, e.g. P, As, Sb, leaves one valence electron without a partner.



(from Schockley)

Since the impurity contributes an excess electron to the lattice, it is called a donor.

The wavefunction of the dopant atom extends over many neighbors.

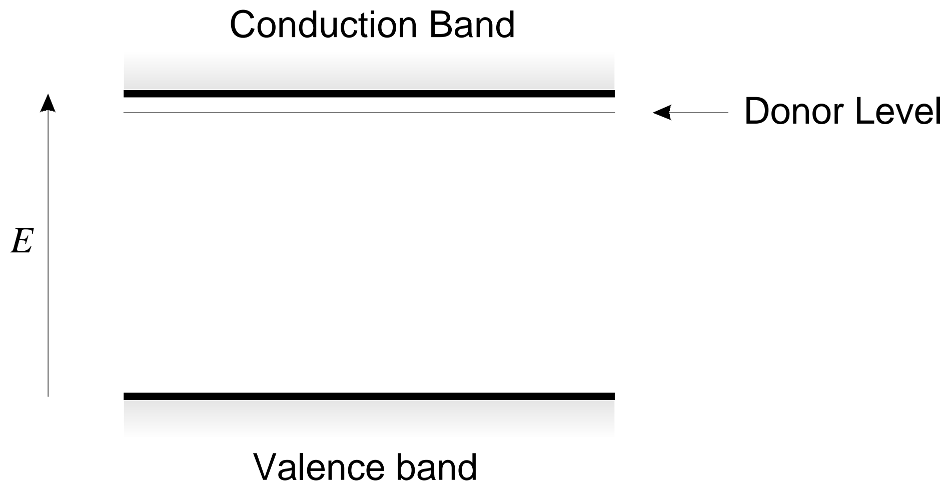


(from Schockley)

The excess electron is only loosely bound, as the coulomb force is reduced by the dielectric constant ε of the medium (ε =12 in Si).

$$E_i(lattice) \propto \frac{E_i(atom)}{\varepsilon^2}$$

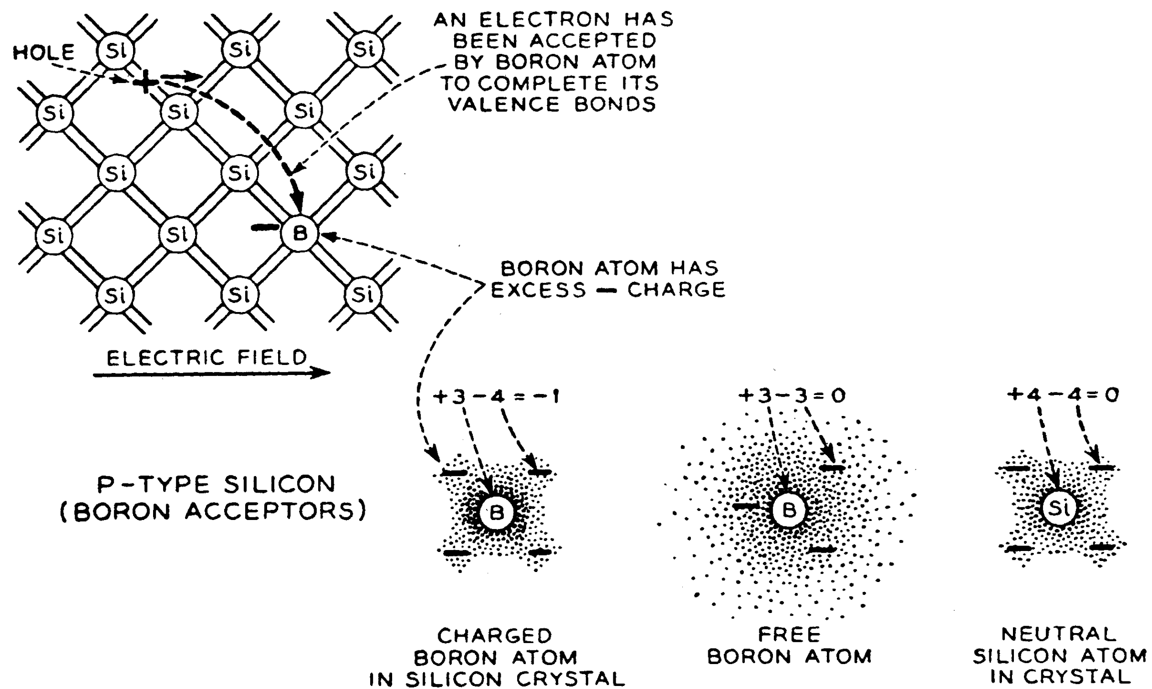
The bound level of this unpaired electron is of order 0.01 eV below the conduction band (e.g. for P: E_c - 0.045 eV).



⇒ substantial ionization probability at room temperature (E= 0.026 eV) – “donor”

⇒ electrons in conduction band

Conversely, introducing a group 3 atom (B, Al, Ga, In) leaves a Si valence electron without a partner.



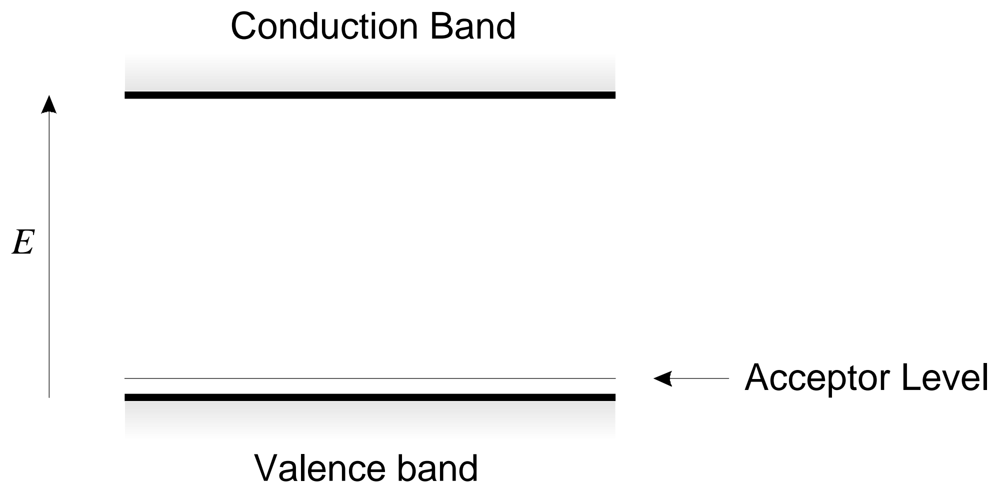
(from Schockley)

To close its shell the B atom “borrows” an electron from a lattice atom in the vicinity.

This type of dopant is called an “acceptor”.

The “borrowed” electron is bound, but somewhat less than other valence electrons since the B nucleus only has charge 3.

This introduces a bound state close to the valence band, also of order 0.01 eV from the band edge.



For example, a B atom in Si forms a state at E_v + 0.045 eV.

Again, as this energy is comparable to kT at room temperature, electrons from the valence band can be excited to fill a substantial fraction of these states.

The electrons missing from the valence band form mobile charge states called “holes”, which behave similarly to an electron in the conduction band, i.e. they can move freely throughout the crystal.

Since the charge carriers in the donor region are electrons, i.e. negative, it is called “n-type”.

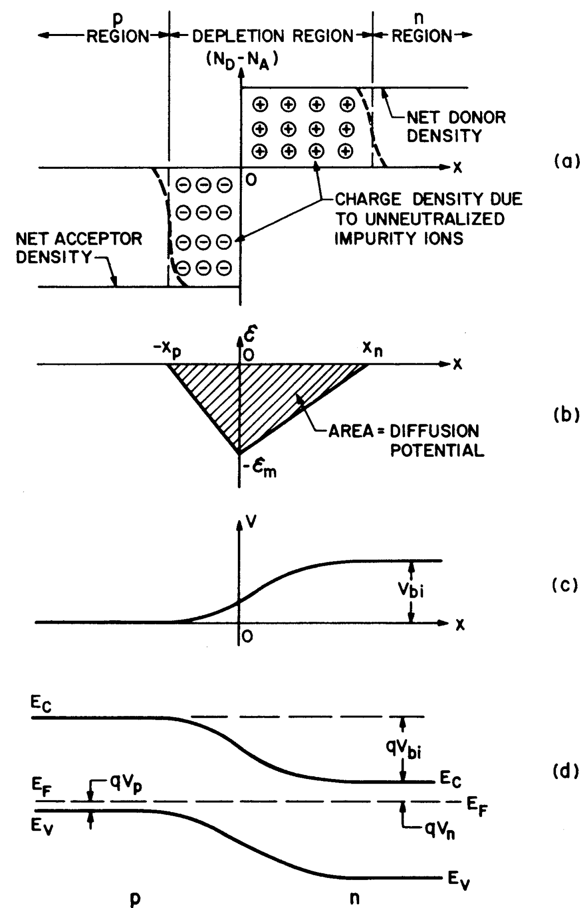
Conversely, as the charge carriers in the acceptor region are holes, i.e. positive, it is called “p-type”.

Consider a crystal suitably doped that a donor region and an acceptor adjoin each other, a "*p-n* junction".

Thermal diffusion will drive holes and electrons across the junction. Although the *p* and *n* regions were originally electrically neutral, as electrons diffuse from the *n* to the *p* region, they uncover their respective donor atoms, leaving a net positive charge in the *n* region.

This positive space charge exerts a restraining force on the electrons that diffused into the *p* region, i.e. diffusion of electrons into the *p* region builds up a potential. The diffusion depth is limited when the space charge potential exceeds the available energy for thermal diffusion.

The corresponding process also limits the diffusion of holes into the *n*-region.



(from Sze, *Physics of Semiconductor Devices*)

The diffusion of holes and electrons across the junction leads to a region free of mobile carriers – the “depletion region”, bounded by conductive regions, which are n- and p-doped, respectively.

Strictly speaking, the depletion region is not completely devoid of mobile carriers, as the diffusion profile is a gradual transition.

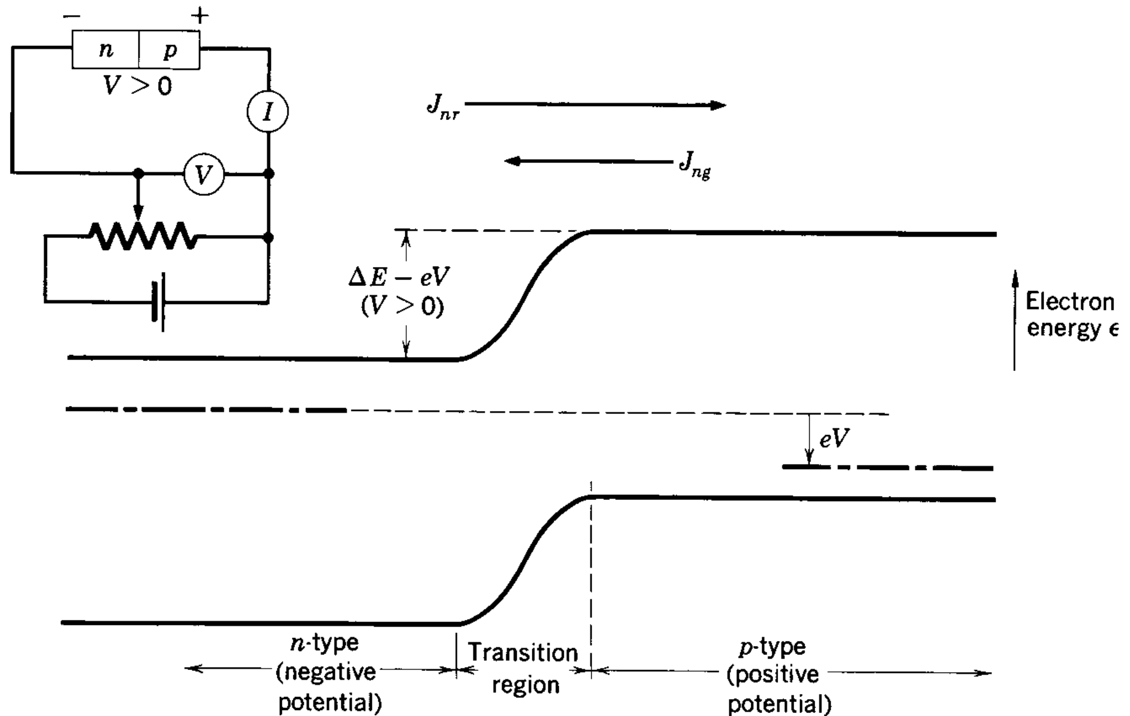
Nevertheless, since the carrier concentration is substantially reduced, it is convenient to treat the depletion zone as an abrupt transition between bulk and 0 carrier concentration.

Furthermore, the formation of the two adjacent space charge regions builds up a potential barrier between the n and p regions, which impedes the further flow of charge.

The magnitude of this potential barrier is typically 50 – 90% of the band-gap, depending on relative doping levels.

This represents the situation in thermal equilibrium. By application of an external potential, two distinctly different non-equilibrium modes can be established.

a) positive potential applied to the p region
negative potential applied to the n region



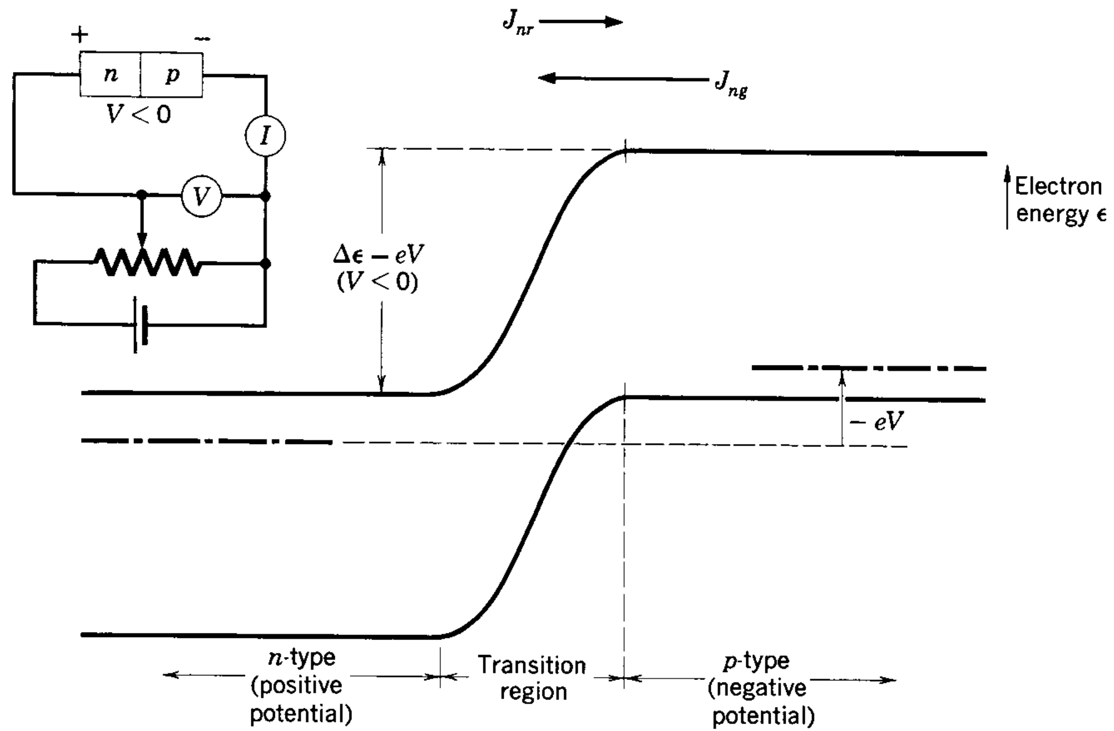
(from Kittel, *Introduction to Solid State Physics*)

The externally applied voltage reduces the potential barrier, allowing increased charge transfer across the junction.

$\Rightarrow$ "forward bias"

Electrons flowing from the n-region across the junction are replenished from the external voltage supply and large current flow is possible.

b) negative potential applied to the p region
positive potential applied to the n region



(from Kittel, *Introduction to Solid State Physics*)

This arrangement increases the potential barrier across the junction, impeding the flow of current.

$\Rightarrow$ "reverse bias"

The p-n junction is asymmetric with respect to current flow (diode).

a) forward bias

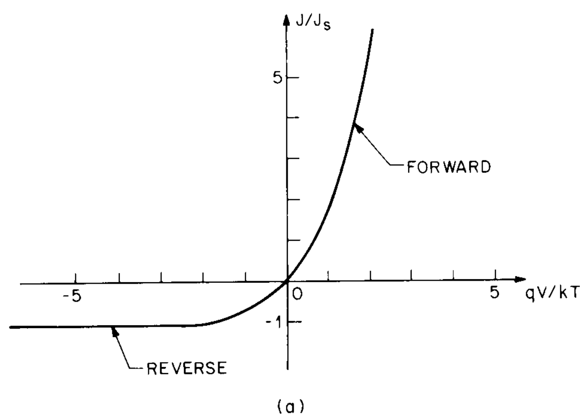
positive supply connection $\rightarrow$ p contact
negative supply connection $\rightarrow$ n contact

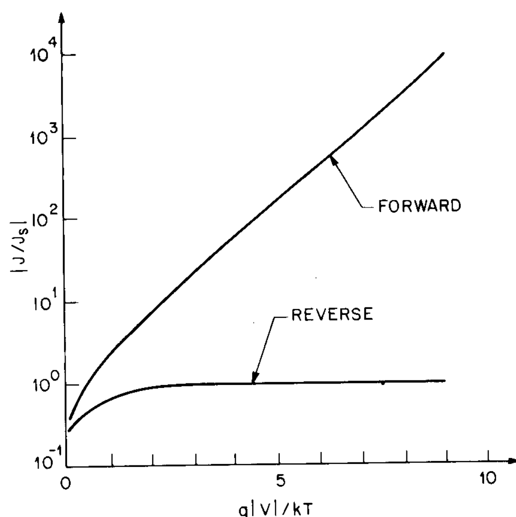
$\Rightarrow$ large current flow

b) reverse bias

positive supply connection $\rightarrow$ n contact
negative supply connection $\rightarrow$ p contact

$\Rightarrow$ small current flow





(from Sze, *Physics of Semiconductor Devices*)

Since the depletion region is a volume with an electric field, it by itself could be used as a radiation detector.

- The width of the depletion region is increased by reverse bias.

1. Depletion width and electric field in p-n junction

Assume a reverse bias voltage V_b and that the potential changes only in the direction perpendicular to the n-p interface. Poisson's equation is then

$$\frac{d^2V}{dx^2} + \frac{Nq_e}{\varepsilon} = 0 \tag{1}$$

where N is the dopant concentration and q_e the electron charge.

Consider an abrupt junction where charge densities on the n and p sides are $N_d\, q_e$ and $N_a\, q_e$, respectively.

If the limits of the depletion region are x_n on the n-side and x_p on the p-side, after two successive integrations one obtains on the n-side

$$\frac{dV}{dx} = -\frac{q_e N_d}{\varepsilon}(x - x_n) \tag{2}$$

and

$$V = -\frac{q_e N_d}{\varepsilon}\frac{x^2}{2} + \frac{q_e N_d x x_n}{\varepsilon} + V_j \tag{3}$$

where V_j is the potential at the metallurgical junction. For $x = x_n$

$$V(x_n) = V_b = \frac{q_e N_d x_n^2}{2\varepsilon} + V_j \tag{4}$$

and the contribution of the n-region to the total reverse bias potential becomes

$$V_b - V_j = \frac{q_e N_d x_n^2}{2\varepsilon} \ . \tag{5a}$$

Correspondingly, in the p-region

$$V_j = \frac{q_e N_a x_p^2}{2\varepsilon} \tag{5b}$$

and the total potential becomes

$$V_b = \frac{q_e}{2\varepsilon}(N_d x_n^2 + N_a x_p^2) \ . \tag{6}$$

Due to overall charge neutrality

$$N_d x_n = N_a x_p \tag{7}$$

and

$$V_b = \frac{q_e}{2\varepsilon}\left(1+\frac{N_a}{N_d}\right) N_a x_p^2 = \frac{q_e}{2\varepsilon}\left(1+\frac{N_d}{N_a}\right) N_d x_n^2 \ . \tag{8}$$

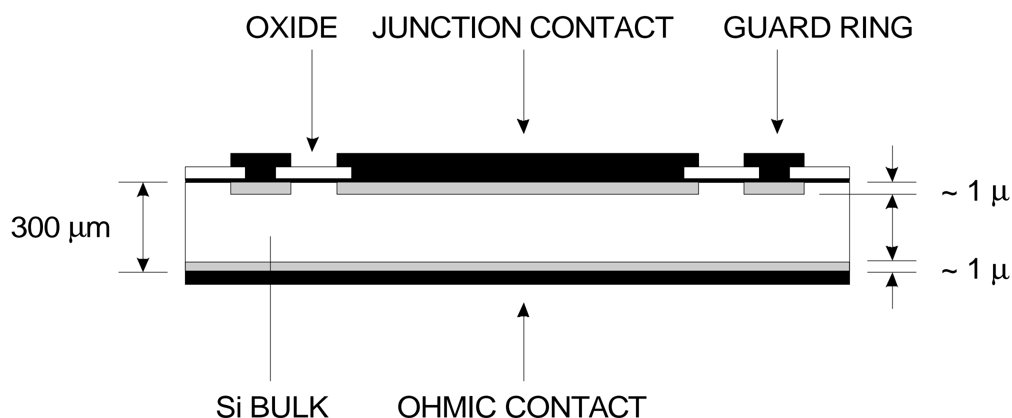
The depletion widths on the n- and p-side of the junction are

$$x_n = \sqrt{\frac{2\varepsilon V_b}{q_e N_d (1+N_d / N_a)}} \quad ; \quad x_p = \sqrt{\frac{2\varepsilon V_b}{q_e N_a (1+N_a / N_d)}} \tag{9}$$

and the total depletion width becomes

$$W = x_n + x_p = \sqrt{\frac{2\varepsilon V_b}{q_e} \frac{N_a + N_d}{N_a N_d}} \tag{10}$$

Detector diodes are usually asymmetrically doped. The starting material (bulk) is lightly doped and the junction is formed by diffusing or ion-implanting a highly doped layer.



The external connection to the lightly doped bulk is made by an additional highly doped layer of the same type (non-rectifying, “ohmic” contact).

- The depletion region then extends predominantly into the lightly doped bulk.

Other details:

The guard ring isolates the wafer edge (saw cut) from the active region.

In the gap between the detector electrode and the guard ring it is critical to provide a neutral interface at the silicon surface to prevent formation of a conductive path.

This is best accomplished by oxide passivation (SiO_2).

If, for example, $N_a >> N_d$, the depletion region extends predominantly into the n-side and the total depletion width is

$$W \approx x_n = \sqrt{\frac{2\varepsilon V_b}{q_e N_d}} \quad . \tag{11}$$

The doping concentration is commonly expressed in terms of resistivity

$$\rho = (\mu q_e N)^{-1},$$

because this is a readily measurable quantity. The parameter μ describes the relationship between the applied field and carrier velocity (to be discussed later).

Using resistivity the depletion width becomes

$$W = \sqrt{2\varepsilon \mu_n \rho_n V_b} \quad . \tag{12}$$

Note that this introduces an artificial distinction between the n- and p-regions, because the mobilities μ for electrons and holes are different.

Since the mobility of holes is approximately 1/3 that of electrons, p-type material of a given doping concentration will have 3 times the resistivity of n-type material of the same concentration.

As discussed earlier, even in the absence of an external voltage electrons and holes to diffuse across the junction, establishing a "built-in" reverse bias voltage V_{bi}. If we take this inherent bias voltage into account, $V_b \equiv V_b + V_{bi}$ and one obtains for the one-sided junction

$$W \approx x_1 = \sqrt{\frac{2\varepsilon(V_b + V_{bi})}{q_e N_d}} = \sqrt{2\varepsilon\mu_n \rho_n (V_b + V_{bi})}.$$

For example, in n-type silicon (V_b in volts and ρ in Ω·cm)

$$W = 0.5\,\mu m \text{ x } \sqrt{\rho(V_b + V_{bi})}$$

and in p-type material

$$W = 0.3\,\mu m \text{ x } \sqrt{\rho(V_b + V_{bi})}$$

The depleted junction volume is free of mobile charge and thus forms a capacitor, bounded by the conducting p- and n-type semiconductor on each side.

The capacitance is

$$C = \varepsilon \frac{A}{W} = A \sqrt{\frac{\varepsilon\, q_e N}{2(V_b + V_{bi})}}$$

For bias voltages $V_b >> V_{bi}$

$$C \propto \frac{1}{\sqrt{V_b}}$$

In technical units

$$\frac{C}{A} = \frac{\varepsilon}{W} \approx 1\,[\text{pF/cm}]\frac{1}{W}$$

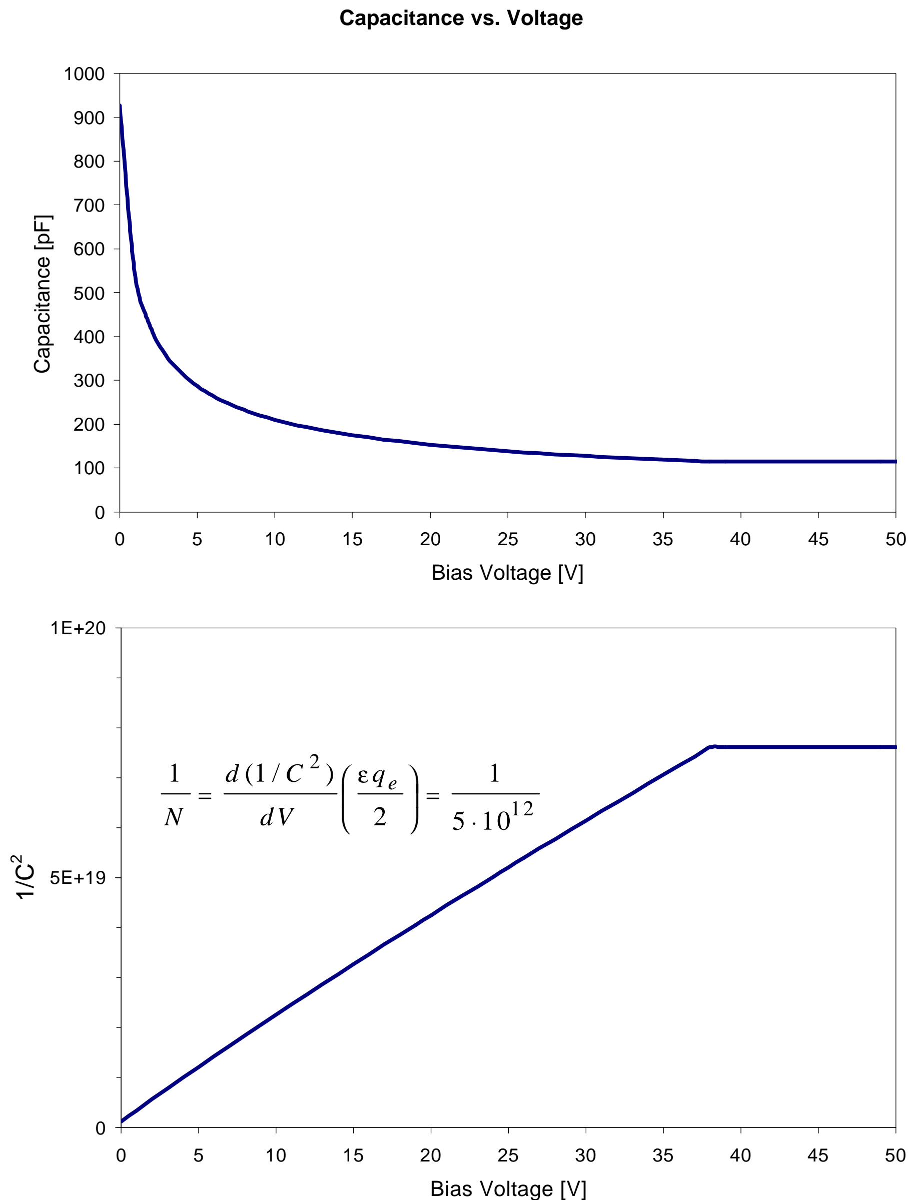
A diode with 100 μm thickness has about 1 pF/mm^2.

The capacitance vs. voltage characteristic of a diode can be used to determine the doping concentration of the detector material.

$$\frac{C}{A} = \sqrt{\frac{\varepsilon\, q_e N}{2(V_b + V_{bi})}}$$

In a plot of $(A/C)^2$ vs. the detector bias voltage V_b the slope of the voltage dependent portion yields the doping concentration N.

Example: Si pad detector, A= 1 cm^2, 100 µm thick



2. Charge Collection

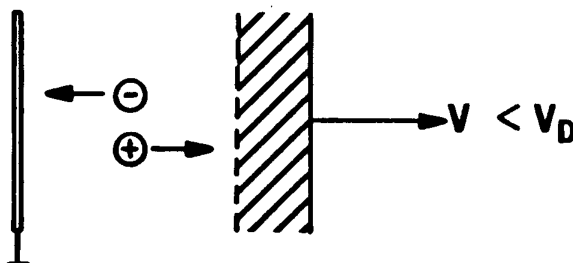
Mobile electrons and holes formed by radiation move under the influence of the electric field in the junction.

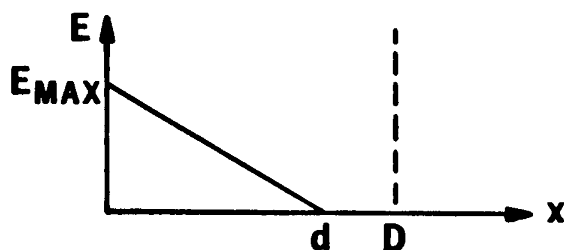
Although electrons and holes move in opposite directions, their contribution to the signal current is of the same polarity.

The time required for a charge carrier to traverse the sensitive volume is called the collection time.

Using the depletion width eq. 13 one can rewrite eq. 2 for the electric field

$$E(x) = \frac{2(V_b + V_{bi})}{W}\left(\frac{x}{W} - 1\right) \qquad (14)$$





The detector bulk is completely depleted of mobile charge when $W = d$, the thickness of the substrate. This occurs at the externally applied depletion voltage

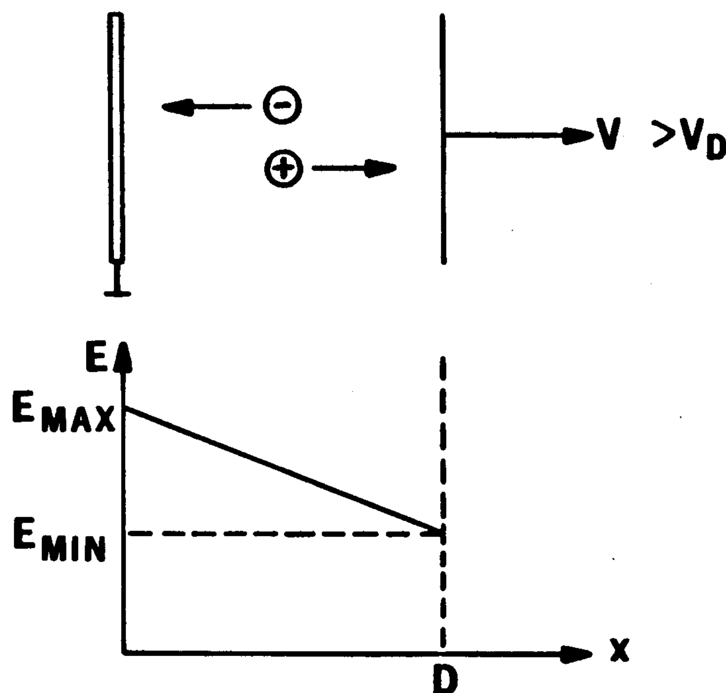
$$V_d = \frac{q_e N_d W^2}{2\varepsilon} - V_{bi} \;. \tag{15}$$

The field drops linearly from its maximum value at the junction to zero at the opposite contact.

Increasing the bias voltage beyond this value adds a uniform field due to the voltage beyond depletion, yielding a distribution

$$E(x) = \frac{2V_{di}}{W}\left(1 - \frac{x}{W}\right) + \frac{V_b - V_{di}}{W} \tag{16}$$

where $V_{di} \equiv V_d + V_{bi}$ has been defined as the internal depletion voltage.



First consider a detector operated at partial depletion $V_b < V_d$. The field

$$E(x) = -\frac{q_e N_d}{\varepsilon}(W - x) \equiv E_0 (W - x) \qquad (17)$$

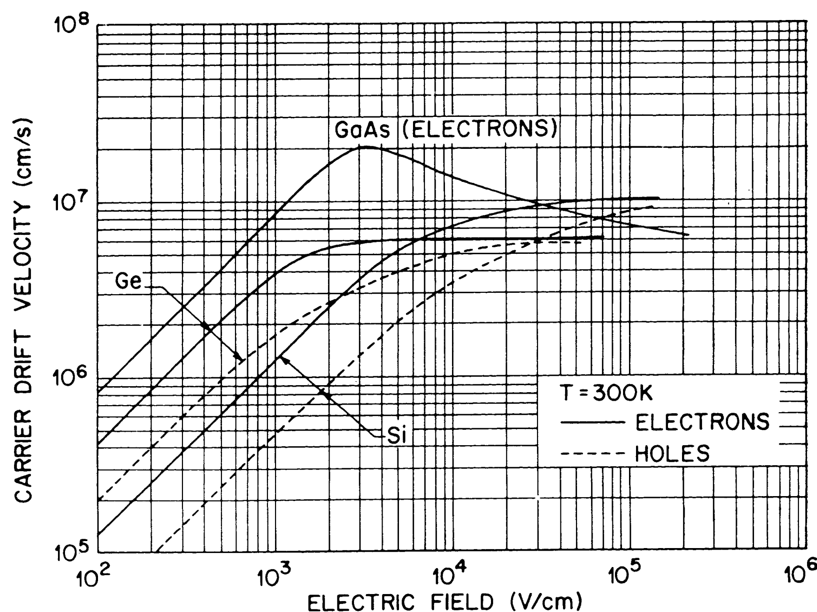
The local velocity of a charge carrier

$$v(x) = \mu E(x) = \mu E_0 (W - x) \qquad (18)$$

Note that the velocity does not depend on the time during which the charge carrier is accelerated, as in normal ballistic motion, since the charge carrier also interacts with the crystal lattice, exciting lattice vibrations (phonons). Since the characteristic times for phonon excitation are much smaller than the transport times, the carrier is always in equilibrium with the lattice, so the velocity is only a function of the electric field, at every position in the depletion region.

In Si at 300K the mobility at low fields is 1350 cm^2/ Vs for electrons and 480 cm^2/ Vs for holes.

The mobility is constant up to about 10^4 V/cm, but then increased phonon emission reduces the energy going into electron motion, so the mobility decreases. At high fields $E > 10^5$ V/cm the mobility $\mu \propto 1/E$ and carriers attain a constant drift velocity of 10^7 cm/s.



(from Sze)

The time required for a charge originating at x_0 to reach a point x is

$$t(x) = \int_{x_0}^{x} \frac{1}{v(x)} dx = \frac{1}{\mu E_0} \int_{x_0}^{x} \frac{1}{W-x} dx = -\frac{1}{\mu E_0} \left[\ln(W-x) \right]_{x_0}^{x}$$

$$t(x) = -\frac{1}{\mu E_0} \ln \frac{W-x}{W-x_0} = \frac{\varepsilon}{\mu q_e N_d} \ln \frac{W-x}{W-x_0} \qquad (19)$$

Consider a hole drifting toward the high-field region and collected at the p-electrode $x = 0$. Using the hole mobility μ_p eq. 19 yields

$$t(x_0) = -\frac{1}{\mu_p E_0} \ln \frac{W}{W-x_0} = \frac{\varepsilon}{\mu_p q_e N_d} \ln \frac{W}{W-x_0} \qquad (20)$$

If we define a characteristic collection time

$$\tau_p \equiv \frac{\varepsilon}{\mu_p q_e N_d} \ ,$$

then

$$t(x_0) = \tau_p \ln \frac{W}{W-x_0} \qquad (20a)$$

For example, $t(x_0 = 0.5W) = 0.7\tau_p$ and $t(x_0 = 0.95W) = 3.0\ \tau_p$.

For the electrons drifting toward the low-field electrode $x = W$, eq. 19 does not yield a solution. However, it can be rewritten to yield the position as a function of time

$$x(t) = W - (W - x_0) e^{-t/\tau_n} \qquad (21)$$

where τ_n has been defined analogously to τ_p. For a charge originating at the metallurgical junction $x_0 = 0$ and drifting toward $x = W$

$$x(t) = W(1 - e^{-t/\tau_n}) \ . \qquad (22)$$

In this simple picture, a charge drifting toward the low field region is never collected (in reality this is accomplished by diffusion), although after a time $t = 3\tau_n$ the carrier will have traversed 95% of the detector. Note that in a partially depleted detector the collection time constants τ_n and τ_p are independent of the applied bias voltage (and depletion thickness), but determined only by the doping concentration of the bulk material and the carrier mobility. τ_n is numerically equal to the dielectric relaxation time of the n-type bulk

$$\tau = \rho\varepsilon = \varepsilon_{Si}\, \varepsilon_0\, \rho = 1.05 \left[\frac{ns}{k\Omega \cdot cm} \right] \rho \qquad (23)$$

In n-type silicon of 10 kΩ·cm resistivity τ_n= 10.5 ns and τ_p= 31.5 ns, and typical collection times in partially depleted detectors are about 30 and 90 ns, respectively.

The collection time can be reduced by operating the detector at bias voltages exceeding the depletion voltage. The field distribution was given in eq. 16, which can be rewritten as

$$E(x) = E_0 \left(1 - \frac{x}{W} \right) + E_1 \qquad (24)$$

This yields a collection time

$$t(x) = \int_{x_0}^{x} \frac{1}{v(x)}\, dx = \frac{1}{\mu} \int_{x_0}^{x} \frac{1}{E_0 \left(1 - \frac{x}{W} \right) + E_1}\, dx$$

$$t(x) = -\frac{W}{\mu E_0} \left[\ln(E_0 + E_1 - E_0 \frac{x}{W} \right]_{x_0}^{x}$$

$$t(x) = \frac{W}{\mu E_0} \ln \frac{E_0 + E_1 - E_0 \dfrac{x}{W}}{E_0 + E_1 - E_0 \dfrac{x_0}{W}} . \tag{25}$$

For holes originating at $x_0 = W$ and drifting to the p-electrode $x = 0$

$$t_{cp} = \frac{W}{\mu_p E_0} \ln \left(1 + \frac{E_0}{E_1} \right) . \tag{26a}$$

The corresponding result obtains for electrons originating at $x_0 = 0$ and drifting to the n-electrode $x = W$

$$t_{cn} = \frac{W}{\mu_n E_0} \ln \left(1 + \frac{E_0}{E_1} \right) . \tag{26b}$$

For large overbias $E_1 >> E_0$,

$$\ln \left(1 + \frac{E_0}{E_1} \right) \approx \frac{E_0}{E_1}$$

and
$$t_{cp} = \frac{W}{\mu_p E_1}$$

as expected for a uniform field.

Rewritten in terms of voltages, eqs. 26a and 26b become

$$t_{cp} = \frac{W^2}{2\mu_p V_{di}} \ln\left(\frac{V_b + V_{di}}{V_b - V_{di}}\right)$$

and

$$t_{cn} = \frac{W^2}{2\mu_n V_{di}} \ln\left(\frac{V_b + V_{di}}{V_b - V_{di}}\right) \qquad (27)$$

For n-type silicon of 10 kΩ·cm resistivity,

a detector thickness of 300 µm, and

a reverse bias voltage V_b= 60V= 2V_d

(i.e. E_0 =2·10^3 and E_1 =10^3 V/cm)

the collection times for electrons and holes are 12 and 36 ns, substantially less than in the partially depleted device.