No.02 APP NOTE
April 1982

Silicon Photodiodes
Physics and Technology

Silicon photodiodes are semiconductor devices used for the detection of light in ultra-violet, visible and infrared spectral regions. Because of their small size, low noise, high speed and good spectral response, silicon photodiodes are being used for both civilian and defense related applications. Depending on the requirement of a particular application, photodiodes can be made in any desired geometry, and provided in a special package with a filter for any special application.

Basic considerations in the design of silicon photodiodes for desired performance characteristics will be discussed. Various parameters used to characterize photodiodes will be introduced. The differences between photoconductive and photovoltaic modes of operation, and their relative merits and limitations will be pointed out to enable users of silicon photodiodes to make an intelligent choice. The use of parameters like thickness, surface finish and anti-reflection coatings to control responsivity and speed in UV, visible and near IR spectral regions for fiber optic and other applications will be discussed.

Photodiode – operational amplifier combinations used for the detection of very low light intensities will be introduced.

General Discussion

Silicon photodiodes are solid state semiconductor devices, sensitive to light in the wide spectral range of 200 – 1200nm, which extends from deep ultraviolet through the visible to the near infrared. The silicon photodiodes can therefore see whatever the human eye can and much more. They can be used to detect the presence or absence of minute light intensities and can be calibrated to measure the intensity of light extremely accurately from very minute light intensities of below 10^{-13} watts/ cm^2 to high intensities above $10mW/cm^2$. In the photovoltaic mode they work as "solar cells" and convert solar and light energy into electrical energy.

It is because of these characteristics that silicon photodiodes have found use in such diverse applications as photography, pollution monitoring, distance and speed measurement, laser ranging, optical communications, energy conversion, computer peripherals and many others.

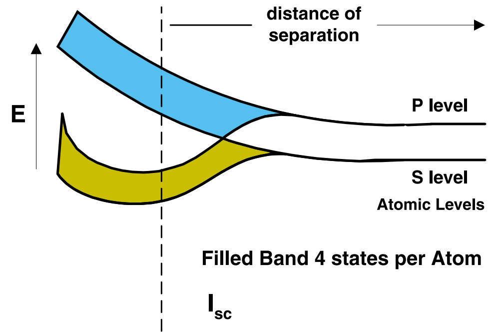
Physics of Semiconductor Photodiodes

Silicon photodiodes are solid state semiconductor devices. To understand how and why a photodiode works as a detector of light wavelengths in a certain spectral region one must first understand what a semiconductor is.

All matter in nature exists as solids, liquids, gases or in plasma form. As we know, all matter is made up of atoms, which essentially consist of a positively charged nucleus and negatively charged electrons rotating around it in fixed orbits. The electrons are bound to the nucleus with a force of attraction because of opposite charge and this binding energy decreases as the electron is farther away from the nucleus. The electrons in the outermost orbit are called valence electrons and determine most of the chemical properties of these elements. Silicon and germanium have four electrons in their outermost orbits and are therefore in the fourth (IV) column of the periodic table of elements. When an electron is removed from its orbit and separated from the atom, the atom is said to be ionized and the result is a free electron and a positively charged ion. A plasma is a collection of charged particles. An atom can be ionized by heat, pressure, high energy particle bombardment or exposure to light of appropriate wavelength such that the photons have more energy than the binding energy of the electrons. This is called the ionization energy.

As the atoms are brought closer by cooling or applying pressure the matter changes physical form from gas to liquid to solid. The reason for this change of form lies in an equilibrium being reached at every temperature between the thermal energy of the atoms, which is a random oscillatory motion and the attractive force because of the overlapping of outermost electron orbits. At each temperature there is an equilibrium distance between atoms, which results in the changes of phases from gas to liquid to solid.

As the atoms are brought closer together their electron orbits overlap and hence the discreet energy levels of the free atoms turn into energy bands in the solid phase. The uppermost band, completely empty or partially filled, is called the conduction band. The next lower band is called the valence band. Depending on the relative position of these bands all solids are divided into three categories; (1) metals, (2) semi-



Lattice Spacing in solids

Fig. 1 As the atoms are brought closer to form a solid, their discrete energy levels split into energy bands of the solid.

conductors, and (3) insulators. In metals the conduction band is partially filled or the conduction and valence bands overlap. In semiconductors there is a small energy gap between the conduction and valence bands (0.1 – 0.5 eV, Silicon Eg = 1.119 eV; and Ge Eg = 0.67 eV). In insulators there is a large energy gap between the conduction band and the valence band (Eg > 6 eV) like $Si0_2$, Si_3 N_4 etc. Table I shows energy gap values for semiconductors at room temperature.

CRYSTAL	E_g (eV)	Crystal	E_g (eV)
Diamond	5.33	PbS	0.34-0.37
Si	1.14	PbSe	0.27
Ge	0.67	PbTe	0.30
InSb	0.23	CdS	2.42
InAs	0.33	CdSe	1.74
InP	1.25	CdTe	1.45
GaAs	1.4	ZnO	3.2
AlSb	1.6-1.7	ZnS	3.6
GaP	2.25	ZnSe	2.60
SiC	3	AgCl	3.2
Te	0.33	AgI	2.8
ZnSb	0.56	Cu_20	2.1
GaSb	0.78	$Ti0_2$	3

TABLE I
Values of the Energy Gap between the valence and conduction bands in semiconductors, at room temperature.

These bands also have discrete energy levels but they are so close together that the energy of an electron in a partially filled conduction band can be increased almost continuously. The metals have high electrical conductivity, because all the electrons in the conduction band behave like free electrons and are available to conduct electricity. Conductivity of a solid = neμ where n is the number of free charge carriers, e their charge and μ their electrical mobility.

$$\text{Re}\,sistivity = \frac{1}{Conductivi\ ty} = \frac{1}{ne\mu}$$

On the other hand, in a semiconductor such as silicon, there are no free electrons since as shown in Figure 2, the valence band is completely filled and the conduction band, which is empty, is separated from the valence band by an energy gap. In order to conduct electricity, electrons have to be excited from the valence band into the conduction band. This can be done by supplying enough energy to the electrons to excite them from the valence band into the conduction band either thermally, by heating, or by exposure to high energy particles or photons with energy hν ≥ Eg. The number of electrons excited thermally from the valence band into the conduction band in silicon is given by:

$$n_i = 2\left[\frac{2\pi\, m_e\ kT}{h^2}\right]^{3/2} \exp\left[\frac{E_g}{2kT}\right]$$

To excite electrons from the valence band into the conduction band by photons, their energy should be ≥ 1.119 eV for silicon at 300K, which corresponds to a light wavelength of λ = 1500nm. This is the reason why silicon photodiodes at room temperature cannot be used for wavelengths λ > 1200nm. However, as the temperature is increased it's bandgap decreases and hence the light of longer wavelengths can also be detected.

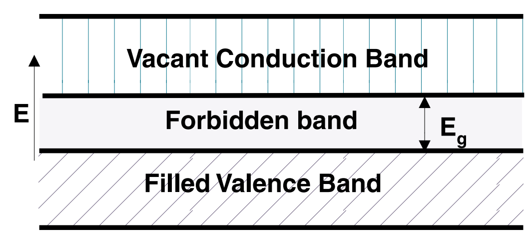


Fig. 2 Band scheme for intrinsic conductivity in a semiconductor. At 0°K the conductivity is zero because all states in the conduction band are vacant. As the temperature is increased, electrons are thermally excited from the valence band to the conduction band, where they become mobile.

The total number of electrons in a solid can be distributed over all the allowed energy levels. The highest energy level, up to which all of the energy levels are occupied at absolute zero temperature, is called the Fermi level. For metals the Fermi level (denoted by E_f) lies in the conduction band. In intrinsic semiconductors and insulators as shown in Figure 3, the Fermi level lies at the center of the energy gap.

$$E_f = \frac{1}{2}(E_c + E_v)$$

Where E_c denotes the bottom of the conduction band and E_v the top of the valence band.

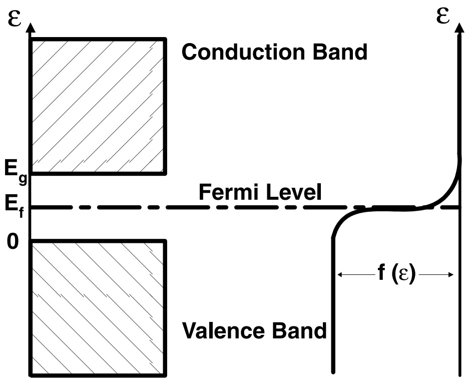
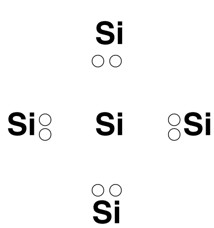


Fig. 3 Energy scale for statistical calculations. The Fermi distribution function is shown on the same scale, for a temperature $K_BT < E_g$; the Fermi level E_f for an intrinsic semiconductor lies at the center of the band gap.

When any two solids are brought into contact with each other the Fermi level in one of the two moves up or down to be at the same level as in the other. The Fermi level acts as a chemical potential and must be the same throughout the solid.

This concept of bandgap can be understood in terms of an atomic picture as follows:

Silicon has four electrons in the outermost orbit and has a diamond crystal structure, i.e., all atoms are bonded tetrahedrally to each other.



Each silicon atom Shares its four electrons with four other silicon atoms such that each has eight electrons in the outermost orbit thus creating a stable configuration. Since all the electrons are bound to one atom or another, there are no free electrons to conduct electricity. The bandgap energy is the energy required to break these tetrahedral bonds, such that the electron is free to move around and conduct electricity.

In a pure semiconductor, when an electron is excited from the valence band to the conduction band it leaves behind a hole in the valence band. So, for an intrinsic semiconductor in thermal equilibrium the number of electrons and holes are equal. For silicon at 22ºC the number of electrons excited from valence band into the conduction band is given by:

$$n_i = 3.873\times10^{16}\ T^{3/2}\ \exp\left(-\frac{70\ \ 14}{T}\right) = 9.3\times10^{9}\ cm$$

	P	As	Sb
Si	0.045	0.049	0.039
Ge	0.0120	0.0127	0.0096

TABLE 2
Donor ionization energies E_d of pentavalent impurities in germanium and silicon, in eV

	B	Al	Ga	In
Si	0.045	0.057	0.065	0.16
Ge	0.0104	0.0102	0.0108	0.0112

TABLE 3
Acceptor ionization energies Ea of trivalent impurities in germanium and silicon, in eV

However, if we put in an atom with five electrons in the outermost orbit in place of a silicon atom, there will be one extra unbonded electron very loosely bound to the parent atom. Such an atom thus acts as a donor and the energy required to knock this electron free from the parent atom is called the ionization energy of the donor. Phosphorous and arsenic are the most commonly used donor impurities for silicon. Since semiconductors with excess electrons are called N-Type, the donor impurities are called N-Type dopants. In terms of band picture the donor impurities result in donor levels just below the condition band edge by the donor ionization energy. Table 2 shows the donor ionization energies for the most commonly used donors in silicon and germanium. The Fermi level for N-Type silicon lies closer to the conduction band edge.

Similarly, elements of group III in the periodic table like Boron, Aluminum, Gallium, etc., act as acceptors since they have one electron less than silicon. The absence of an electron from its normal site is called a "hole". The acceptor impurities result in excess holes and a semiconductor with excess holes is called P-Type, so the acceptor impurities are called P-Type dopants. As shown in Figure 5, the acceptor levels lie just above the top of the valence band, and the Fermi level is closer to the valence band edge. Table 3 shows the ionization energies of the acceptor impurities in silicon and germanium.

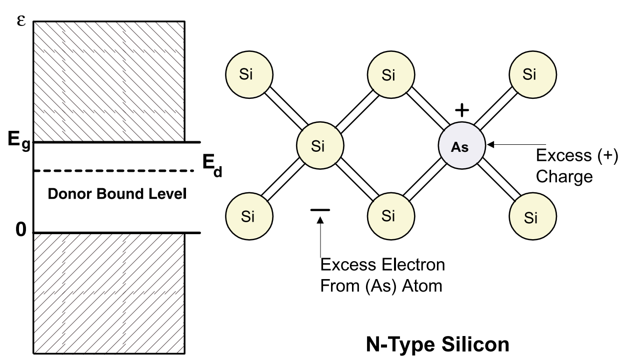


Fig. 4 Charges associated with impurity atom in silicon; with arsenic impurity an electron is available for conduction. The arsenic atom is called a donor because when ionized it gives up an electron to the conduction band.

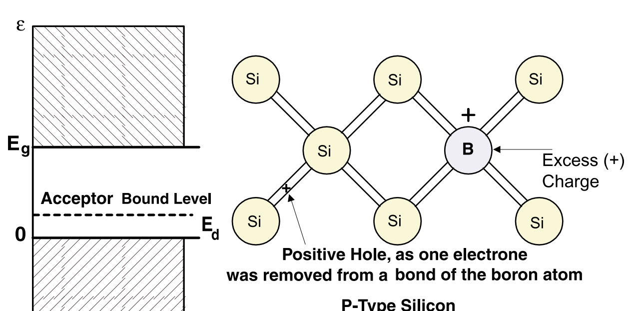


Fig. 5

Fig-5 With boron impurity a positive hole is available for ionization and conduction.The boron atom is called an acceptor atom because when ionized it takes up an electron from the valence band. (the ionization of the hole associated with the acceptor corresponds to the addition of an electron to the acceptor, the hole moving to the former stste of the valence electron.)

The electrical conductivity is given by:

$$\sigma = ne\mu_n + pe\mu_p$$

Where n and p are the concentrations of electrons and holes and u_n and u_p their respective mobility's. For an intrinsic semiconductor

$$n = p = n_i$$

$$\sigma = n_i e \quad (\mu_n + \mu_p)$$

For an N-Type semiconductor, n >> p

$$\sigma = ne\mu_n$$

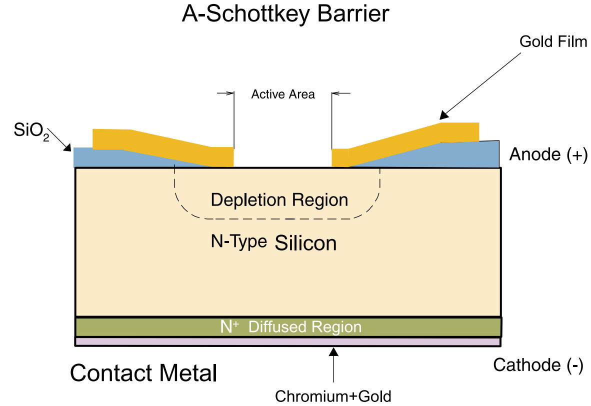
For a P-Type semiconductor, p >> n

$$\sigma = pe\mu_p$$

When a metal comes in contact with a semiconductor, a schottky barrier results. When N-Type silicon comes in contact with P-Type silicon a P-N junction is formed.

Schottky barrier type silicon photodiodes are made by diffusing an N+ layer on the back of a high resistivity N-Type substrate for ohmic contact, and by evaporating a thin gold metal layer (= 150 Å thick) on a specially prepared surface on the front side. These Schottky barrier photodiodes behave just like the P^+N junction type photodiodes.

A-Schottkey Barrier



B-Planar Diffused

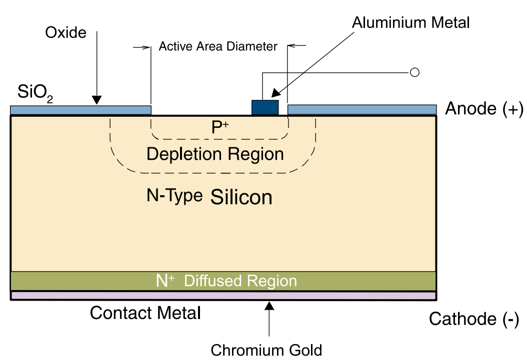


Figure 6 (a) shows the construction of a Schottky barrier photodiode.

The planar diffused $P^+N\ N^+$ photodiodes are made by diffusing an N^+ layer on the back for Ohmic contact, and a P^+ layer in the active area on the front, defined by an oxide mask, to make the P^+N junction. The bulk region between the junction and the N^+ back layer serves as the absorption region. The back metallization is chromium and gold and the front metallization is usually aluminum.

Figure 6 (b) shows the construction of a planar diffused $P^+N\ N^+$ silicon photodiode.

The photodiodes are always reverse biased to collect the charge carriers generated by the incident light. The N-Type region behaves like a cathode and is always made positive. The P-Type region is the anode and is made negative. The electrons flow from P-Type region to the N-Type region inside the photodiode and the photocurrent flows from P-Type region to the N-Type region in the outside circuit.

III Characteristic Parameters of Photodiodes

A. Built In Voltage

When a P-Type dopant is diffused into an N-Type substrate a P-N junction is formed. Since the N-Type region has an excess of electrons, the concentration gradiant causes electrons to diffuse from the N-Type region to the P-Type region and the holes from the P-Type region to the N-Type region. As shown in Figure 7 an equilibrium is reached when the electric field generated because of these carriers, is equal to the concentration gradient. The electric field is such, that it exerts a force on electrons to drive them back into the N-Type region to balance the force due to concentration gradients. The electric field integrated over the junction region, is the built-in voltage, Vb_i. The magnitude of this voltage is equal to the difference between the Fermi levels in N and P-Type regions.

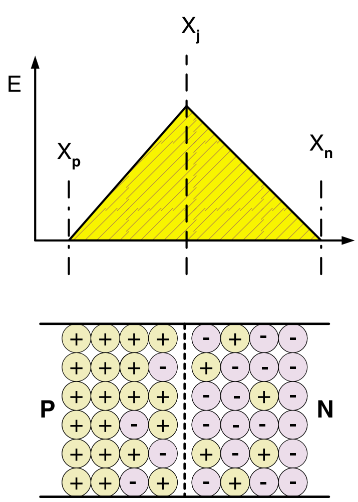
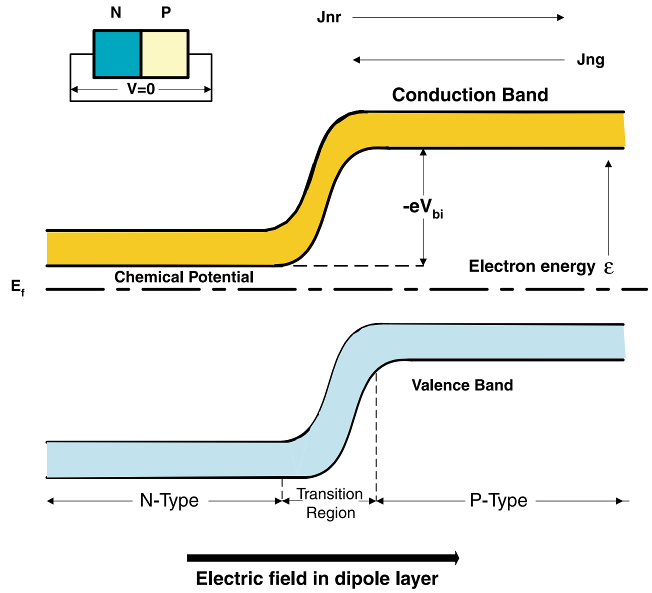


Figure 7



The region over which the electric field extends is called the depletion region.

The electric field $$E = e\int_{x=x_n}^{x=x_j} n(x)dx$$

Built in voltage

$$Vbi = \int_{X_p}^{X_n} E(X)dx = (E_{Fn} - E_{FP})\frac{1}{e} = \frac{KT}{e}\ln\left[\frac{N_A N_D}{ni^2}\right]$$

The boundaries of the depletion region are determined by equating the integral of the electric field to the difference between Fermi levels on the two boundary lines in the N and P-Type regions. N_A and N_D are the acceptor and donor concentrations at the boundaries of the depletion region.

B. Dark Current

There are four major contributors to the dark current

1. Diffusion Current
2. Generation – Recombination Current
3. Surface Current
4. Avalanche Current

1. Diffusion Current

The diffusion current arises from the regions within a diffusion length of the minority carriers next to the junction

$$\left(L_p = \sqrt{D_p \varsigma_p}\right) \qquad \textit{Diffusion current density:}$$

$$J = J_o \left[\exp . \left(\frac{ev}{kT} \right) - 1 \right]$$

$$\text{Where } J_o = \frac{e n_i^2}{N_B} \sqrt{\frac{D_p}{\varsigma_p}} \text{ for a } p^+ N \text{ junction}$$

n_i = instrinsic carrier concentration
N_B = substrate carrier concentration
D_p = Hole (minority carrier) diffusion coefficient
ς_p = Hole (minority carrier) lifetime

This is the standard diode equation, with J_o the reverse saturation current:

$$\text{Also } J_o \, \alpha \, \frac{1}{\sqrt{\varsigma p}}$$

$$J_o \alpha \frac{1}{NB} \quad \text{or } J_o \, \alpha \, p \text{ the substrate resistivity}$$

The temperature dependence of diffusion current comes through the intrinsic carrier concentration:

$$J_o \, \alpha \, n_i^2 \quad \text{or } J_o \alpha \quad T^3 \exp . \left(\frac{E_g}{kT} \right)$$

Thus the diffusion part of the dark current is larger for higher substrate resistivity and lower minority carrier lifetime and increases with temperature as $T^3 \exp(E_G/kT)$.

$$I_{DB} = 1/2 \, e \left(\frac{n_i}{\varsigma} \right) \quad A_j . \; W_d .$$

1. Generation Recombination Current

All impurity atoms and defects in the silicon crystal lattice result in trap levels in the forbidden energy gap, which act as generation-recombination centers. This current is generated in the bulk depletion region:

$$\text{Where: } A_j = \text{diffused junction area.}$$

$$W_d = \text{depletion region width.}$$

$$W_d = \sqrt{\frac{2 K_s \varepsilon_o V}{e \, N_B}} = \sqrt{2 K_s \varepsilon_o p \mu \; (V + V_{bi})}$$

Thus the bulk contribution to dark current is larger for all diffused area, shorter minority carrier lifetime and larger substrate resistivity. The dark current increases as when V is the applied reverse bias. This dependence comes through increased depletion region width. After the diode is fully depleted, the bulk generated dark current no longer increases with applied bias.

3. Surface Current

If the depletion region extends to silicon surface, all the impurities and defects at the surface contribute to dark current. At the silicon-silicon dioxide interface there are defects and trapped impurities which act as generation recombination centers. If the surface is not passivated by an oxide or nitride layer, impurities can diffuse in, over a period of time, from outside and lead to increased dark current.

$$I_{DS} = 1/2 \, e s_o \;\; ni \;\; A_{SD}$$

Where S_o is the surface recombination velocity, and A_{SD} is the depleted surface area. The surface generated dark current becomes particularly important for surface inversion layer devices, like UV diodes, and also for Schottky diodes, because in these devices the depleted surface area A_{SD} is larger than Those in diffused junction devices.

4. Avalanche Current

Avalanche multiplication occurs when the electric field in some part of the photodiode reaches 30 V/µm. Such high fields could occur at sharp edges in the junction region or near an impurity or defect.

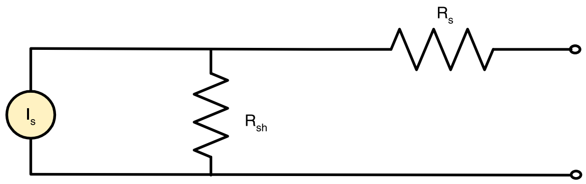
Usually at very low voltages (i.e. a few millivolts) the diffusion current makes up the majority of the dark leakage current. As the voltage is increased and the depletion region extends to the surface, the surface generated dark current dominates. As the voltage is increased further the bulk generated dark current dominates.

At still higher voltages, when electric field at some point in the junction region(usually near sharp corners) becomes > 30V per µm, the avalanche multiplication current dominates.
The temperature dependence of dark current comes through ni, the intrinsic carrier concentration. Diffusion current is proportional to n_i^2 whereas the bulk and surface generated dark currents are proportional to n_i.

C.Shunt (or Source) Resistance. R_{sh}

A photodiode can be represented as a current source with a shunt (or source) resistance, connected in parallel to it, and a series (or forward) resistance R_S in series. An ideal diode has:

$$R_{sh} = \infty \quad \text{and} \quad R_s = O$$



Shunt resistance is the slope of the I-V curve at the origin (V=0). Experimentally R_{sh} is obtained by applying ± 10mV and calculating the effective resistance of the diode.

D.Series Resistance

Series resistance of a photodiode is the resistance of the contacts and the undepleted bulk of the substrate.

$$R_s = \left[\frac{W_o - W_d}{A_j}\rho\right] + \textit{contact resistance}$$

$$\textit{where } W_o = \textit{substrate thickness and}$$

$$W_d = \textit{depletion region width.}$$

For photodiodes which are fully depleted, the series resistance is just the resistance of the contacts.

E. Capacitance

The diode acts as a capacitor with the boundaries of the depletion region as the two plates of a parallel plate capacitor. Usually the diodes are made as step junctions with heavy doping in the active area region to get an Ohmic contact. The depletion region widths in the p+ region of a P on N diodes is small compared with that in the comparatively higher resistivity N-Type substrate. The zero bias capacitance is therefore inversely proportional to the substrate resistivity being smaller for higher resistivity substrates. At any applied reverse bias voltage V the capacitance is given by:

$$C = \frac{K_s \varepsilon_o}{W_d} A_j = K_s \varepsilon_o A_j \left[2K_s \varepsilon_o P\mu \left(V + V_{bi}\right)\right]^{-1/2}$$

Where W is the depletion region width, A_j the diffused area, K_s the dielectric constact, ε_o the permitivity of free space, p the resistivity of the substrate and μ the mobility of the majority carriers in the substrate.

F. Breakdown Voltage

Breakdown voltage is usually taken as the voltage when dark current I_d is ≥ 10 μA. For planar diffused junction diodes the electric field is maximum at the curved edges of the diffused region. Avalanche breakdown occurs when the electric field is such, that the carriers are accelerated to high enough energies to generate more carriers by avalanche multiplication. Maximum electric field at the junction in the plane region.

$$E_m = e \int_{x_j}^{X_n} n(x)\, d_x$$

Where n (x) is the net carrier concentration at any position x, from the junction down into the substrate, and x_n is the boundary of the depletion region in the substrate. As the applied reverse bias is increased the depletion region extends deeper into the substrate and therefore E_m increases. $X_n \propto \sqrt{\rho V}$

Applied voltage V is the integral of the electric field over the depletion region. For higher resistivity material, the carrier concentration is smaller and the depletion region width X_n is larger, since .

Therefore the maximum field E_m is smaller for the same applied bias V. For the same resistivity material, since N-Type silicon has a smaller carrier concentration than P-Type, for the same applied bias voltage, depletion region width in N-Type silicon will be larger than in P-Type silicon and hence for the same substrate resistivity a P on N diode has a higher breakdown voltage than an N on P.

Since thicker substrates absorb more light at longer wavelengths and larger depletion region widths for higher resistivity material helps better collections of the carriers generated by light, the responsivity at longer wavelengths is greater for higher resistivity materials operated at higher bias.

However, as the depletion region-width increases, for the same minority carrier lifetime in the bulk, the dark current increases. Hence the smaller capacitance, higher responsivity for longer wavelengths, and faster risetimes are accompanied by higher operating voltages and larger dark currents.

G. Responsivity and Quantum Efficiency

The responsivity of a photodiode is a measure of its sensitivity to light and is defined as the ratio of the current produced by the photodiode (amps), to the amount of light falling on it (watts).

$$R = \frac{I\,(amps)}{L\,(watts)} = \frac{\lambda\,(\mu m)}{1.24}\, T(I - e^{-\alpha x})$$

Where T is the fraction of incident light transmitted into silicon, the absorption coefficient of the light in silicon and x is the silicon substrate thickness.

The absorption coefficient of light in silicon depends on its wavelength λ 2 cm^{-1} for λ = 1100nm, corresponding to photon energy hν = 1.13 eV. For λ 400nm (corrensponding to photon energy hν = 3.10 eV) the absorption coefficient is = 5.4 x 10^6 cm^{-1}.

The responsivity of silicon photodiodes peaks between 850 and 950nm and is about 0.6 A/W at the peak. At longer wavelengths the responsivity decreases because the absorption coefficient is too small, and therefore the silicon thickness required to absorb 99.9% of λ = 11nm light, for example, is about 3.45 cm, which is much longer than the thickness of the silicon substrate wafers normally used (12-14 mils or 0.031 – 0.036cm). At shorter wavelengths, although the absorption coefficient is large and most of the light is absorbed, but only one electron holes pair is generated in silicon corresponding to each photon. So the part of photon energy hν in excess of the energy gap E_g, is wasted in producing heat. Therefore the current produced by the photodiode per watt of light falling on the detector is smaller. At shorter wavelengths < 400nm, the silicon thickness required to absorb 99.9% of the light is small < 1.28μm, and therefore, a large fraction of carrier's generated are lost in the heavily doped region close to the surface of the planar diffused diodes, because of small lifetime and high surface recombination velocity.

The quantum efficiency of a photodiode is the percentage of the number of photons incident on the photodiode which contribute to the photocurrent

$$QE = \frac{R_{Observed}}{R\,(100)} \times 100 = T(1-e^{-\alpha x}) \times 100\%$$

$$= \frac{124 R_\lambda}{\lambda(nm)}\%$$

The fraction of light transmitted T, can be increased by putting AR coating on the photodiode. Figure-9 shows calculated responsivity and quantum efficiency for a 120 μm thick silicon photodiode without any AR coating. Figure-10 shows the modified spectral response with AR coating. By putting a suitable AR coating the response of a photodode at a particular wavelength can be decreased or increasd by 20 – 30%. The thickness of the AR coating is usually a mutiple of where λ is the wavelength of light in the anti-reflection coating.

H.Noise N.E.P. and detectivity D*

The shot noise of a photodiode is related to the dark current as < in ^{2}> = 2e I_d. F, where f is the noise bandwidth. When a load R_L is connected across the diode, thermal noise voltage across the load is: < $v_{th}{}^2$ > = 4kT f R_L.

$$< Vnoise >= \left(2\, e\, I_d\, \Delta f\, R_L{}^2 + 4kT\, \Delta f\, R_L\right)^{1/2}$$

Total noise voltage across the load R_L is the sum of the shot noise and thermal noise.

$$< Inoise >= \left(2\, e\, I_d\, \Delta f + \frac{4kT\, \Delta f}{R_L}\right)^{1/2}$$

Noise equivalent power, is the amount of light falling on a photodiode which produces a signal equal to the noise generated internally by the photodiode.

$$N.E.P. = \frac{< i_n >}{Responsivity} = \left(2\, e\, I_d \quad \Delta f \quad + \quad \frac{4kT\, \Delta f}{R_L}\right)^{1/2}$$

$$When\ R_L >> \left(\frac{2kT}{eI_o}\right),\ shot\ noise >> thermal\ noise$$

$$N.E.P. = \sqrt{\frac{2eI_d}{R}} \qquad \frac{watts}{\sqrt{H_z}}$$

The detectivity D* is a measure of detecting ability of the photodiode.

$$Detectivity\ D^* = \sqrt{\frac{Area\ x\ bandwidth}{N.E.P.}} = \sqrt{\frac{A\ x\ \Delta f}{N.E.P.}}$$

$$= R\left(\frac{A}{2eI_d}\right)^{1/2} \qquad \frac{cm.\sqrt{H_z}}{watt}$$

$$D^*\ (850, 900, 5) = 5.16x10^{12}\ cm - \sqrt{\frac{H_z}{watt}}$$

for 850nm light chopped at 900 c/sec in a noise bandwidth of 5 Hz.

1/f noise appears at frequencies below 1000 cps. This is a surface related noise. In general Schottky photodiodes have lower 1/f noise than planar diffused diodes.

I Response Time

The rise time of a photodiode is defined as the time for response between 10% and 90% of the final value of the signal:

$$\varsigma \frac{90}{10} = \left(\varsigma_{cc}^2 + \varsigma_{Diff}^2 + \varsigma_{RC}^2\right)^{1/2}$$

Where ς_{cc} is the time for charge collection from the depleted region, ς_{Diff} the time for the carriers generated by light in the undepleted bulk to diffuse to depleted regions and ς_{Rc} is the RC time constant. The time for the collection of charges generated in the depleted region is given by:

$$\varsigma_{cc} = \frac{W_d}{2v_d} \sim \frac{W_d^{\,2}}{2\mu V} \sim K_s \varepsilon_o \; \rho \sim 3\rho \; x\, 10^{-12} \text{ sec}$$

for P on N devices where W_d is the width of the depletion region and ν_d is the average drift velocity of the carriers. The ime for diffusion of carriers from the undepleted bulk to the depleted region is given by:

$$\varsigma_{Diff} = \left(\frac{W_o - W_d}{D_p} \right)^2$$

where W_o is the substrate thickness and D_p is the diffusion constant.

The RC time constant of the photodiode with a load resistance R_L is given by:

$$\varsigma_{RC} = 2.2 \; (R_s + R_L) C_j$$

Where R_s is the series resistance of the photodiode and C_j is the capacitance of the photodiode at applied bias V.

For a P+N diode on 10 ohm cm N-Type substrate, assumming a built-in voltage ~ 0.5V the zero bias depletion region width W_d ~ 1.2µm. for a diffused area A_J ~ 4.67 mm² and substrate thickness W_o ~ 12 mils = 300µm, series resistance R_s ~ 10 ohms. Zero bias junction capacitance C_J ~ 400 pF. For a load resistance R_L ~ 50

ohm. RC time constant ς_{RC} ~ 2.2 x 60 x 400 x 10^{-12} ~53 n.sec.
Charge collection time ςcc ~ 30 P. sec.

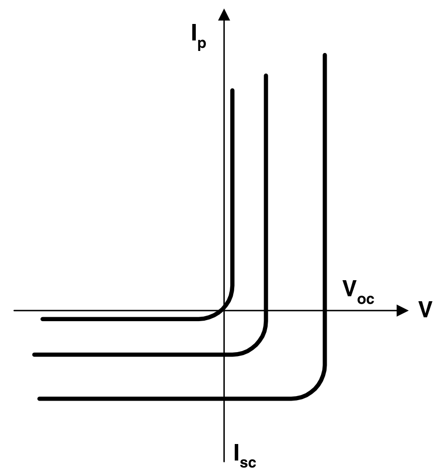
$$\textit{Diffusion time } \varsigma_{Diff} \sim \frac{(300x10^{-4})^2}{12.3} \sim 73\; \mu \text{sec}$$

Thus for unbiased undepleted photodiodes the response time is dominated by the diffusion time of the carriers generated in the undepleted bulk region for longer wavelengths. For short wavelengths when the light is completely absorbed in the depleted region, the charge collection time ς_{cc} and the RC time constant ς_{RC} may become comparable.

IV Photoconductive versus Photovoltaic Operation

In photoconductive mode of operation no external bias is applied across the photodiode. The photodiode is used as a current source. The photocurrent increases linearly with intensity in the range from 10^{-13} watts/cm² to 10^{-3} watts/cm². The photocurrent I_p, is given by:

$$I_p = I_g - I_O \left(Exp \; \left(\frac{ev}{kT} \right) \; -1 \right)$$



Where I_g is the light generated current. The open circuit voltage V_{oc} varies as log of the short circuit current I_{sc}. The region of linearity of short circuit current with light intensity breaksdown when the open circuit voltage V_{oc} becomes close to the built-in voltage V_{bi}. Also the region of linearity decreases as the load resistance R_L increases. The region of linearity can be increased by applying a reverse bias across the diode, which helps collection of all the carriers generated by light. In photovoltaic mode the capacitnce is the zero bias capacitance, which is larger for lower resistivity substrates because the depletion region width due to the built-in voltage is smaller for lower resistivity material. The responsivity for longer wavelengths is also smaller because of the smaller depletion width from which all the carriers generated by light are collected. The carriers from the rest of the undepleted material in which the light gets absorbed are collected by diffusion, and hence the response time for devices operated in the photovoltaic mode is longer (typically for 1'µsec. To 100 µsec.). The zero bias depletion region width due to the built-in voltage is larger for higher resistivity substrates, and for the same resistivity, the depletion region width is longer for N-Type compared to P-Type substrates because of smaller carrier concentration. Therefore, the capacitance is small and long wavelength responsivity is larger for higher resistivity substrates, but the series resistance may be larger and the shunt resistance may be smaller. For shorter wavelengths, if most of the light is absorbed in the depleted region, the response time will be shorter for higher resistivity substrates.

In the photovoltaic mode, in the absence of applied bias, there is no dark current and hence the only source of noise is thermal noise, therefore very low light level intensities can measured.

In the photoconductive mode, a reverse bias is applied across the photodiode which results in a wider depletion region, smaller capacitance, smaller series resistance, shorter rise times and linear photo response over a wider range of light intensities. As the reverse bias is increased, dark current increases and hence the noise is larger.

V. Design considerations for desired performance

A.UV Enhanced

The UV radiation in the spectral region 200-400nm gets absorbed completely in the top 2µm of the silicon surface layer. To increase UV responsivity it is therefore essential to avoid dead layer formation on the surface due to heavy dopant diffusion required for ohmic contacts. The inversion layer diodes have no dead layer and hence their UV response is much better than p-n junction type photodiodes. Application of a small reverse bias depletes the inversion layer region and results in considerably improved responsivity by better collection of charge carriers generated by light. For photovoltaic operation built-in voltage generated by a lightly doped shallow diffused junction will result in better charge collection and improved responsivity.

UV responsivity at a particular wavelength λ can be improved by growing an SiO_2 layer on the silicon surface with a thickness equal to λ/4, or any odd multiple of it, where λ is the wavelength of light in SiO_2. Since SiO_2 does not absorb UV light, it is a good AR coating as well as a passivation layer.

B. IR Enhanced

The absorption coefficient of infra-red radiation in silicon is rather small $^{-2}$ cm $^{-1}$ at 1100nm and therefore the silicon thickness required to absorb the incident radiation is large, ~ 3.5cm for 1100nm. For maximum responsivity the silicon substrate thickness should be greater than or equal to the absorption length of the light in silicon. But for thicker substrate wafers the charge carriers have to travel longer distances to get collected and therefore the response time becomes longer. In order to increase IR responsivity at longer wavelengths for smaller substrate thicknesses a mirror coating of chromium or silver is applied to the back surface to reflect the light, thus increasing the light path by multiple reflections. Still better responsivity is obtained by using textured, or as lapped back surface with a mirror coating so that light from the back surface is reflected in all directions thus increasing the light path in silicon resulting in greater absorption.

Anti-reflection coating of SiO_2, SiO or Si_3 N4 are applied to silicon surface to increase the transmission of incident light into silicon.

IR enhanced photodiodes are made with deep diffused junctions so that the breakdown voltage is high and the diodes are fully depleted at the operating voltage.

A.High breakdown voltage

When electric field at some point close to the junction reaches ≥ 30 V/µm the charge carriers are accelerated to high enough energies to cause avalanche multiplication. The electric field in a planar diffused junction diode is highest at the junction edges because of the junction curvature. To increase the breakdown voltage the diffused area should have no sharp edges and the junction should be deep. Deeper junctions result in higher breakdown voltages. When the junction in the active area has to be made shallow to maintain high UV response, a deep ring is diffused around the junction which determines the breakdown voltage.

The breakdown voltage can also be increased by putting a metal electrode overlapping the junction, such that when reverse bias is applied, an inversion layer is produced which results in a junction with a large effective radius of curvature and hence higher breakdown voltage. In high resistivity p-type substrates since an inversion layer is formed on the surface during oxidation, it is comparatively easy to make photodiodes with higher breakdown voltages.

D.Fast Response

For devices operated in the photovoltaic mode, the response time is dominated by the transit time of the carriers in the undepleted bulk. For shorter wavelengths since the light is absorbed in a short distance from the surface, the substrate resistivity is chosen such that majority of the carriers are generated in the depleted region and are collected by the built-in electric field near the junction. For longer wavelengths when the absorption length is much greater than the depletion region width the response time can be decreased by deep back diffusion resulting in a back surface field because of the large concentration gradient of the majority carriers. If the lifetime of the minority carriers in the bulk of the drift region is smaller than the transit time of the carriers then the response time of the photodiode will be limited by the minority carrier lifetime. Therefore it may be possible to get faster rise times by reducing minority carrier lifetime in the undepleted bulk so that the carriers generated in the undepleted bulk are lost by recombination which will result in lower responsivity but faster rise time.

When a reverse bias is applied the depletion region width increases and the charge collection time in the drift region becomes longer because of longer distances the carriers have to travel. For fastest possible response of fully depleted devices the substrate thickness, diffused area, the series resistance and the load resistance have to be such that

$$\varsigma_{Rc} = \varsigma_{cc}$$

$$or \quad 2\cdot 2 \quad (R_L + R_S)\ C_J = \frac{0.5W_o}{v_d}$$

As shown in figure the drift velocity of the carriers reaches the saturation value of ~ 10^7 cm/sec for electrons and $4.5x10^6$ cm/sec for holes at an electric field of 2 volts/µm. By designing photodiodes using these criteria, it is possible to get rise times of less than a nanosecond. For a device with 20 mils diameter area (Area ~ 0.2mm²)

Load resistance = 50Ω
Series resistance = 10Ω
Device thickness = 200 µm (8 mils)

Assumming a depletion voltage of 50 volts at an applied reverse bias of 300 volts the average electric field is ~ 1.5 V/µm and the average drift velocity of the carriers

$$v_d \sim 5\ x\ 10^6 \quad cm/\sec.$$

$$\cong 2x10^{-9}\ \sec.$$

$$\varsigma_{cc} = \frac{0.5W}{v_d} \sim \frac{0.5\ x\ 200x10^{-4}}{5x10^{-6}}$$

$$\varsigma_{Rc} \sim (50+10)\ x\ 9.998x10^{-14} = 6x10^{-12}\ \sec$$

The rise time of the photodiode is dominated by the transist time of the carriers and can be reduced by decreasing the substrate thickness.

VI. Photodiode Op-Amp combinations for low light level detection

To detect very low light levels, for such applications as short haul optical-fiber links, hybrid preamplifiers or operational amplifiers are used to boost the signal from the photodiodes. The operational amplifier acts as a current to voltage converter with the output signal V_o given by

$$V_o = I_s\ R$$

$$= (P_{in}\ x\ R)\,R_f$$

Where P_{in} is the signal power incident on the photodiode. The effective detector load resistance R_L = where A_o is the open loop gain of the op-amp. This is valid only when

$$R_{sh} >> R_f\ and\ A_o >> 1 + \frac{R_f}{R_{sh}}$$

Zero signal output error, superimposed on signal is

$$V_{OD} = -R_f\,(I_B + I_D + V_{OS}\,(1 + \frac{R_F}{R_{SH}})$$

Where I_B = amplifier biase Current

I_D =Detector Dark Current at the operating Voltage

(I_D = 0 at V = 0)

V_{OS} =amplifier input offset Voltage

R_{SH} = Detector Shunt Resistance (When V=0) = $\frac{V}{I_D}$

if the detector is biased.
for Photodiode op-amp Combination since the effective

Load resistance is $R_{L=}\ \frac{R_f}{A}$ the Photocurrent increases

Linearity with light Intensity Over a wider range of Incident

for application where little space is available or reduced component count is desired, UDT provides a product line of detector-op-amp combinations called Photops ™
In some cases feedback elements can be provided to give specified gain, volts/watts. The standard line of Photops ™ consists of 5.1 mm diameter photodiode and and an op-amp.
Specifications are provided to enable the user to calculate system performance exactly as would be done with discreat elements.