Chapter 5

Photodetectors and Solar Cells

3.1 Photodetectors

Photodetectors come in two basic flavors:

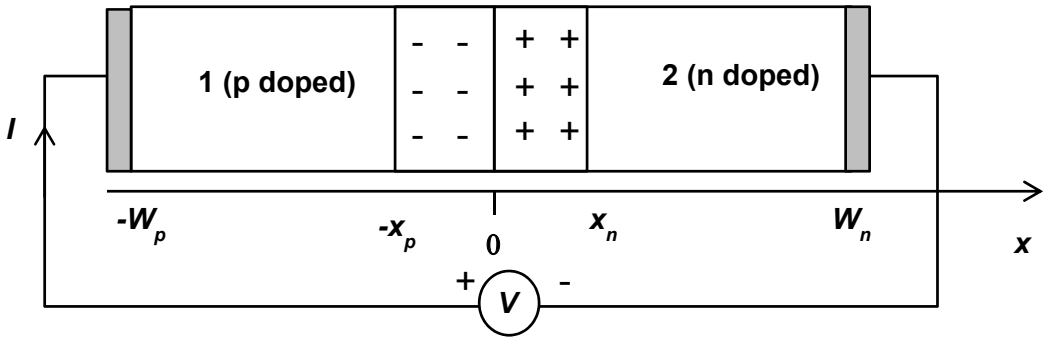
i) Photoconductors
ii) Photovoltaics

A photoconductor is a device whose resistance (or conductivity) changes in the presence of light. A photovoltaic device produces a current or a voltage at its output in the presence of light. In this Chapter, we discuss photodiodes which are by far the most common type of photovoltaic devices. Photoconductors will be the subject of a homework problem.

3.2 Photodiodes

A pn diode can be used to realize a photodetector of the photovoltaic type. Consider the pn diode structure shown in the figure below. Assume that the current-voltage relation of the pn diode, in the absence of light, is given as,

$$I = I_o\left(e^{\frac{qV}{KT}} - 1\right)$$



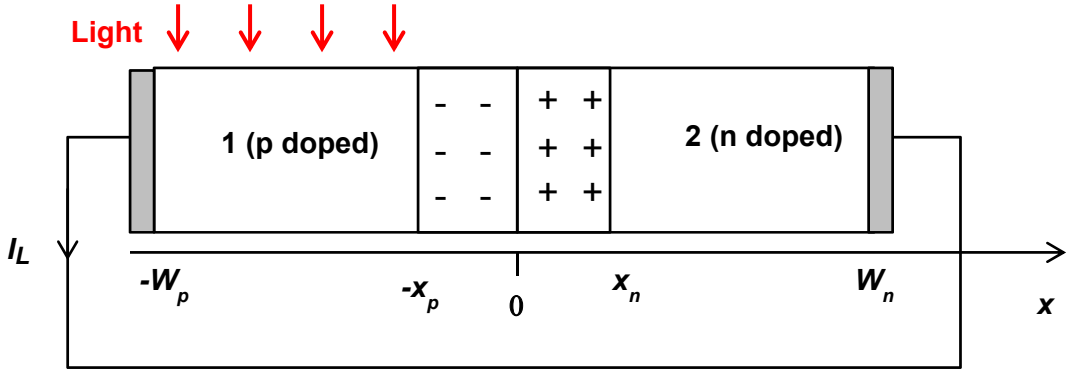
3.2.1 Case I: Photogeneration in the Quasineutral Region:

Now suppose that the applied bias is zero and light is shining on the p-side of the diode, as shown in the Figure below. We need to figure out how much current is coming out of the device. We start from the continuity equation for the minority carriers (i.e. electrons) on the p-side,

$$\frac{\partial n}{\partial t} - \frac{1}{q}\frac{\partial}{\partial x}J_e(x) = G_e(x) - R_e(x)$$

As a result of photon absorption electron and hole pairs are created on the p-side. Let the generation rate of electron-hole pairs due to light be $G_L(x)$. We can add $G_L(x)$ to the right hand side of the above equation to get,

$$\frac{\partial n}{\partial t} - \frac{1}{q}\frac{\partial}{\partial x} J_e(x) = G_e(x) - R_e(x) + G_L(x)$$



Assuming,

$$\frac{\partial n}{\partial t} = 0 \qquad \{\text{steady state}$$

$$J_e(x) = qD_{e1}\frac{\partial n(x)}{\partial x} \qquad \{\text{minority carrier current by diffusion only}$$

$$G_e(x) - R_e(x) = -\frac{\left(n(x) - n_{po}\right)}{\tau_{e1}}$$

$$G_L(x) = \text{constant} = G_L$$

One obtains the following equation for the minority carrier density on the p-side,

$$\frac{\partial^2 n(x)}{\partial x^2} = \frac{\left(n(x) - n_{po}\right)}{L_{e1}^2} - \frac{G_L}{D_{e1}}$$

This is a second order differential equation and needs two boundary conditions. The first boundary condition is that the electron density at the left metal contact must be the equilibrium electron density n_{po}. The second boundary condition is that the electron density at the edge of the depletion region (at $x = -x_p$), given by the expression,

$$n\left(x = -x_p\right) = n_{po} e^{qV/KT}$$

must equal n_{po} since there is no bias across the pn junction. The solution, subject to these two boundary conditions is,

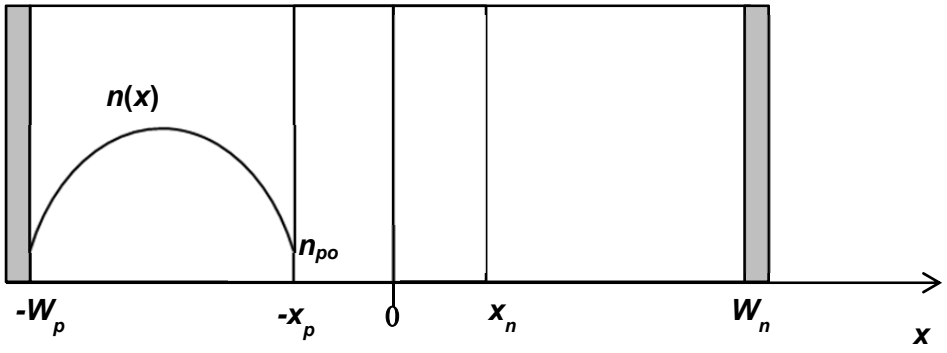
$$n(x) = n_{po} + G_L\tau_{e1} - G_L\tau_{e1}\left[\frac{\sinh\left(\frac{W_p + x}{L_{e1}}\right) - \sinh\left(\frac{x + x_p}{L_{e1}}\right)}{\sinh\left(\frac{W_p - x_p}{L_{e1}}\right)}\right]$$

A particularly interesting and relevant case is when the minority carrier diffusion length is much longer than the length of the p-side,

$$L_{e1} >> \left(W_p - x_p\right)$$

This condition, called the short base limit, implies that that the minority carriers can diffuse through the entire length of the device without recombining and is equivalent to assuming no electron-hole recombination takes place on the p-side. In this case,

$$n(x) = n_{po} + \frac{G_L}{2D_{e1}}\left(\frac{W_p - x_p}{2}\right)^2 - \frac{G_L}{2D_{e1}}\left(x + \frac{W_p + x_p}{2}\right)^2$$



3.2.2 Electron Current:

The electron current flows by diffusion,

$$J_e(x) = q\,D_{e1}\frac{\partial n}{\partial x} = -qG_L L_{e1}\left[\frac{\cosh\left(\frac{W_p + x}{L_{e1}}\right) - \cosh\left(\frac{x + x_p}{L_{e1}}\right)}{\sinh\left(\frac{W_p - x_p}{L_{e1}}\right)}\right]$$

$$= -qG_L\left(x + \frac{W_p + x_p}{2}\right) \qquad \{\text{Short base limit}$$

3.2.3 Total Current:

The electrons that recombine with the holes on the p-side are lost. The electron that diffuse and reach the left metal contact recombine with the holes at the contact and are also lost. However, the electrons that diffuse to the right and reach the edge of the depletion region are quickly transported across the depletion region by the electric field present inside the depletion region. These electrons contribute to the current in the external circuit. The total current J_T is due to the electrons that are able to cross the depletion region, and equals $J_e(x_n)$.

$$J_T = J_e(x_n)$$

Since recombination and generation are assumed to be zero inside the depletion region,

$$J_e(x_n) = J_e(-x_p)$$

$$J_T = J_e(-x_p) = -qG_L L_{e1}\left[\frac{\cosh\left(\frac{W_p - x_p}{L_{e1}}\right) - 1}{\sinh\left(\frac{W_p - x_p}{L_{e1}}\right)}\right]$$

$$= -qG_L\left(\frac{W_p - x_p}{2}\right) \qquad \{\text{Short base limit}$$

3.2.4 Hole Current:

Since $J_T = J_e(x) + J_h(x)$, we have,

$$J_h(x) = J_T - J_e(x)$$

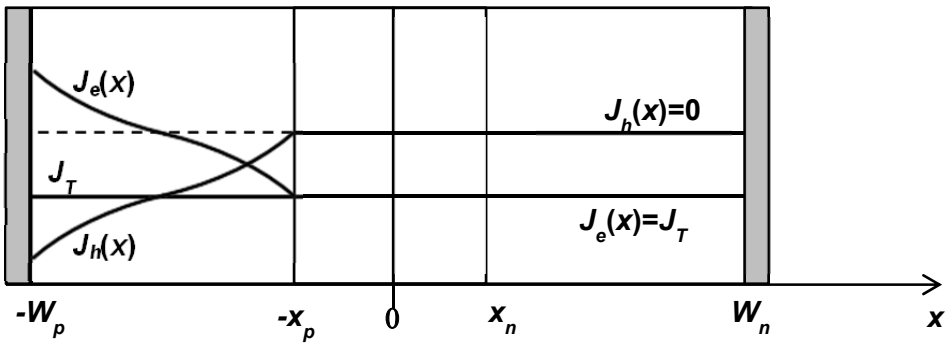
We know the electron current for $-W_p \le x \le x_n$, and therefore we also know the hole current over the same range. The holes, being the majority carriers, flow by both diffusion and drift. The diffusion current of the holes on the p-side can be found as follows. Charge neutrality implies that on the p-side, $\Delta p(x) = \Delta n(x)$. The hole diffusion current is then,

$$J_{h,diff}(x) = -q\,D_{h1}\frac{\partial \Delta p}{\partial x} = -q\,D_{h1}\frac{\partial \Delta n}{\partial x} = -\frac{D_{h1}}{D_{e1}}J_e(x)$$

Since we know that total hole current on the p-side, we can find the hole drift current on the p-side as follows,

$$J_h(x) = J_T - J_e(x)$$
$$\Rightarrow J_{h,drift}(x) + J_{h,diff}(x) = J_T - J_e(x)$$
$$\Rightarrow J_{h,drift}(x) = J_T - J_e(x) - J_{h,diff}(x)$$
$$\Rightarrow J_{h,drift}(x) = J_T - \left[1 - \frac{D_{h1}}{D_{e1}}\right]J_e(x)$$

The electron and hole currents are sketched in the Figure below.



At $x = -x_p$ electrons reaching the depletion region edge are swept away across the depletion region to the n-side of the junction by the electric field in the junction. The holes diffusing towards the depletion region edge cannot cross the depletion region and are left behind. Majority carrier drift current (i.e hole drift current) ensures that these left behind holes don't pile up at $x = -x_p$ but come out at $x = -W_p$.

3.2.5 External Current:

The current in the external circuit is,

$$I_L = -AJ_T$$

$$I_L = qG_L AL_{e1}\left[\frac{\cosh\left(\frac{W_p - x_p}{L_{e1}}\right) - 1}{\sinh\left(\frac{W_p - x_p}{L_{e1}}\right)}\right]$$

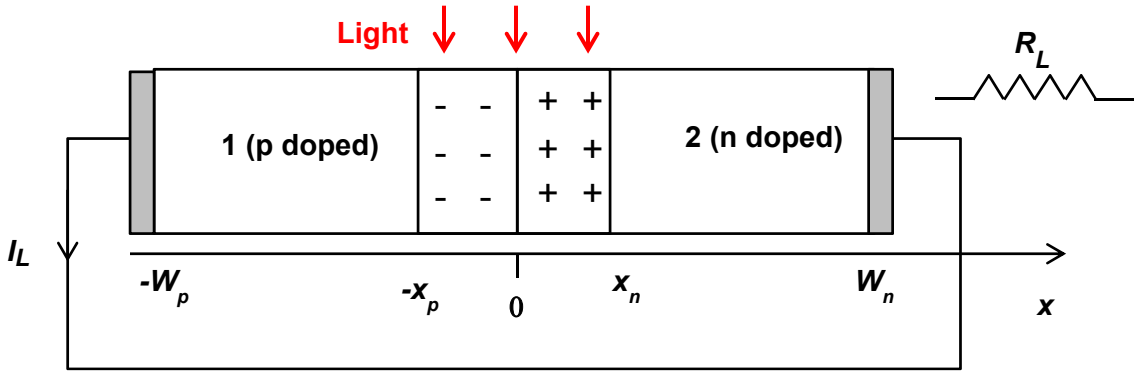
$$= qG_L A\left(\frac{W_p - x_p}{2}\right) \qquad \{\text{Short base limit}$$

If one ignores electron-hole recombination inside the photodetector (shot base limit) then the total charge per second coming out of the photodetector is equal to half the number of electron-hole pairs generated per second inside the device. This makes physical sense since in the short base limit only half of the

photogenerated electrons are able to cross the junction and contribute to the current. The other half recombine at the left metal contact. If we allow for electron-hole recombination inside the photodetector then the external current would get reduced further.

3.2.6 Case II: Photogeneration in the Junction Depletion Region:

We now consider the case when electron-hole pairs are generated by light inside the junction depletion region, as shown in the Figure below.



3.2.7 Electron and Hole Currents:

Since the depletion region width is typically small and since photogenerated electrons and holes are quickly swept out of the depletion region by the large electric field present inside the depletion region, it is safe to ignore electron-hole recombination inside the depletion region. Assuming uniform light illumination inside the depletion region and assuming steady state, we have,

$$-\frac{1}{q}\frac{\partial J_e(x)}{\partial x} = G_L$$

$$\Rightarrow J_e(-x_p) - J_e(x_n) = q\int_{-x_p}^{x_n} G_L\, dx = qG_L(x_n + x_p)$$

$$\frac{1}{q}\frac{\partial J_h(x)}{\partial x} = G_L$$

$$\Rightarrow J_h(x_n) - J_h(-x_p) = q\int_{-x_p}^{x_n} G_L\, dx = qG_L(x_n + x_p)$$

The electric field, present inside the depletion region, sweeps the electrons towards the n-side and the holes towards the p-side. Consequently, it must be that,

$$J_e(-x_p) = 0$$

$$J_h(x_n) = 0$$

This implies,

$$J_h(-x_p) = J_e(x_n) = -qG_L(x_n + x_p)$$

The above equation is just a statement of the fact that all the photogenerated electrons more towards the n-side and all the photo generated holes move towards the p-side.

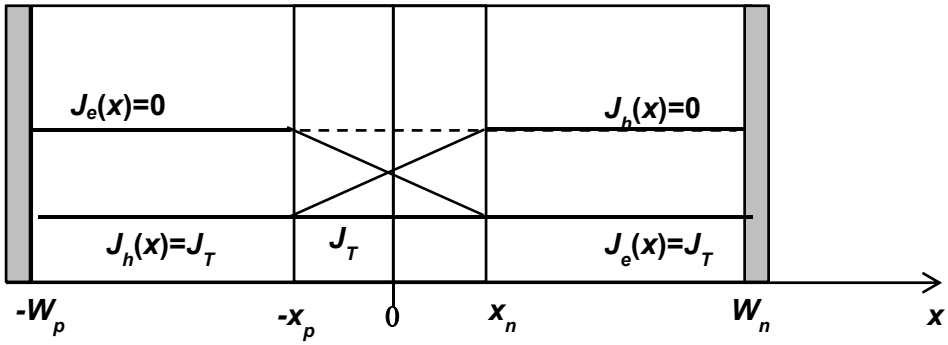
3.2.8 Total Current:

The total current is, as always, obtained by knowing both the electron current and the hole current at any one point in the device,

$$J_T = J_e(x_n) + J_h(x_n) = J_e(-x_p) + J_h(-x_p)$$
$$= -qG_L(x_n + x_p)$$

The current in the external circuit is,

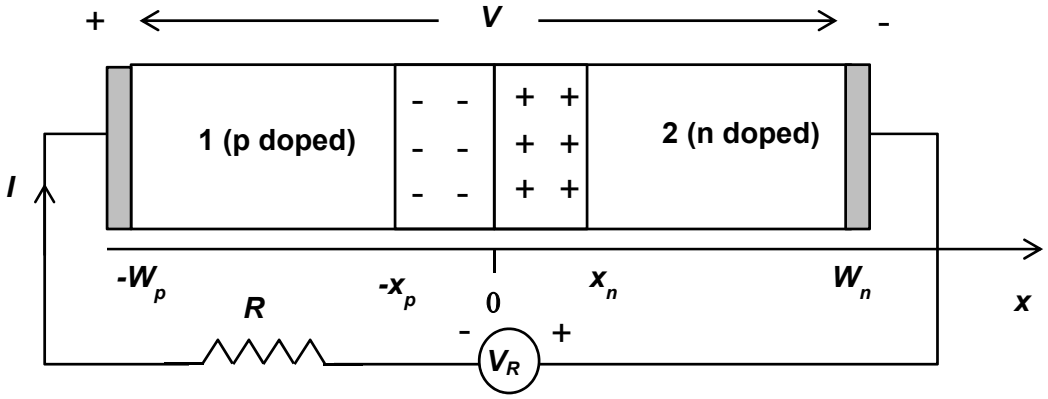
$$I_L = -AJ_T = qG_L\left(x_n + x_p\right)$$



In this case, all the photogenerated electron-hole pairs contribute to the current in the external circuit. The lesson here is that for higher efficiency detector it is far better to have photogeneration take place inside the depletion region than in the quasineutral regions. In most practical photodetectors, photogeneration takes place inside the depletion region as well as inside both the quasineutral regions. The contributions from all these components can be added algebraically to get the total photocurrent I_L.

3.2.9 Photodetector Circuits and Electrical Characteristics:

A pn photodetector is usually connected in the configuration shown in the Figure below.



The junction is reversed biased with voltage V_R to increase the electric field in the depletion region. The current in the circuit has two components:

i) The current due to the intrinsic electrical behavior of a diode and given as,

$$I_o\left(e^{\frac{qV}{KT}} - 1\right)$$

ii) The current due to photogenerated electron-hole pairs (i.e. I_L).

These two components can be added to get the total current in the external circuit,

$$I = I_o\left(e^{\frac{qV}{KT}} - 1\right) - I_L \qquad (1)$$

In writing the above Equation, a 'superposition' principle has been used; the effects of a voltage bias V and light can be superposed. This might not make much sense since a pn diode is not a linear device. A

detailed analysis shows the currents due to the junction bias and due to photogeneration are indeed additive.

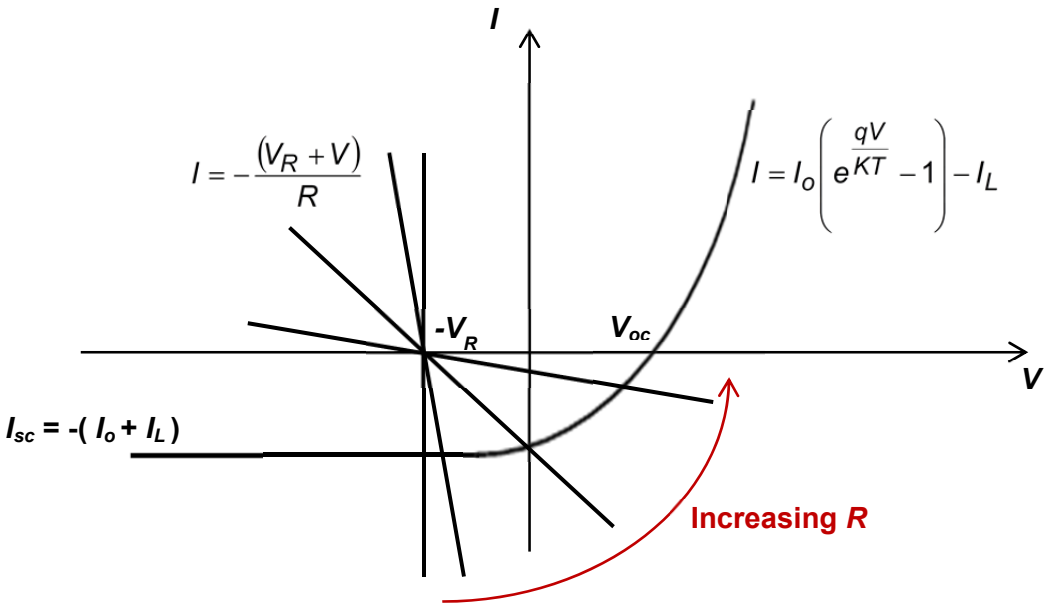
Kirchhoff's voltage law gives us another equation,

$$IR + V_R + V = 0 \qquad (2)$$

A graphical solution of Equations (1) and (2) is shown in the Figure below. Where the 'load line' corresponding to Equation (2) is plotted against Equation (1) for different values of the external circuit resistance R. In the ideal case, when $R = 0$, and the junction is sufficiently reverse biased, the current in the external circuit consists of two parts; the current $-I_L$ due to photogeneration and the current $-I_o$ due to the reverse biased pn junction. This latter unwanted current is called the 'dark' current of the photodetector since it will be present even in the absence of light. Recall that this dark current is due to electron-hole generation in a reverse biased pn diode. The dark current limits the light sensitivity of a photodetector.

3.2.10 Short Circuit Current:

The $R = 0$ case defines the short circuit current I_{sc} of the detector as $-(I_o + I_L)$.



z

When the external resistance R is large, current from the photogeneration process can cause enough potential to drop across the resistor that the pn junction becomes forward biased. This is also undesirable and leads to a decrease the external circuit current since the current of the forward biased pn junction will be in direction opposite to the current due to photogeneration. When $R = \infty$, the forward bias diode current completely cancels the current due to photogeneration.

3.2.11 Open Circuit Voltage:

The $R = \infty$ case defines the open circuit voltage of the detector. It can be computed as follows,

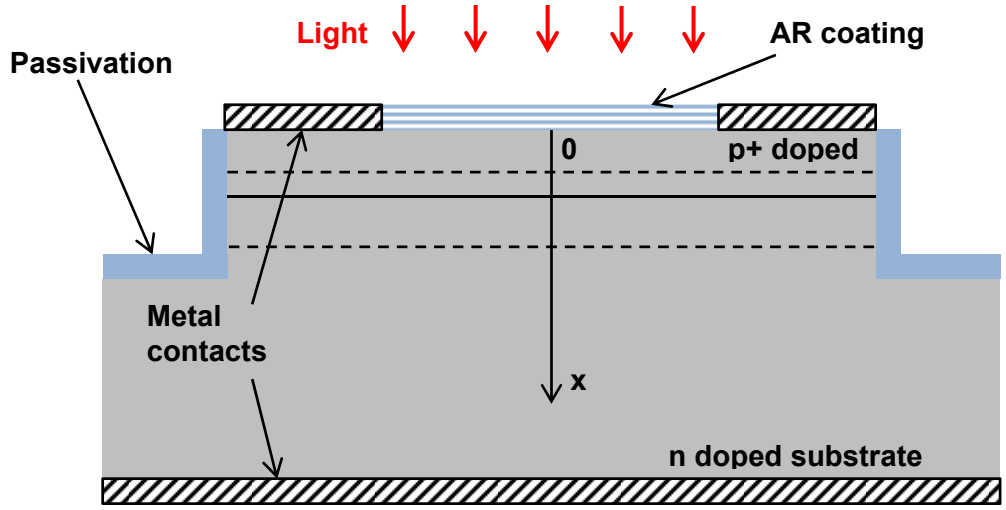
$$0 = I_o\left(e^{\frac{qV_{oc}}{KT}} - 1\right) - I_L$$

$$\Rightarrow V_{oc} = \frac{KT}{q}\ln\left(\frac{I_L}{I_o} + 1\right)$$

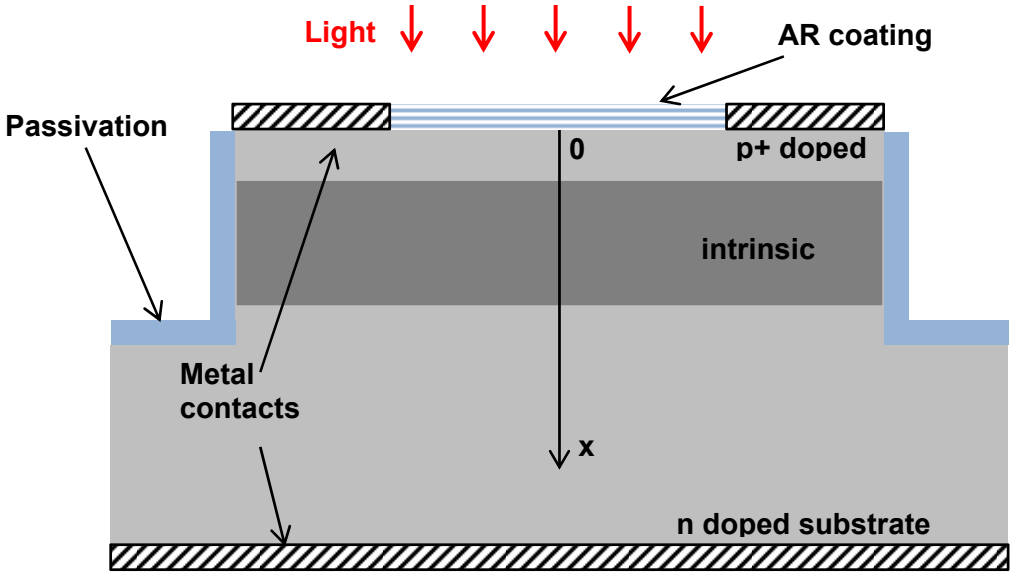
3.2.12 Practical Issues and Real Device Designs:

The structure of an actual pn junction photodiode is shown below. Light is incident from the top. The top surface has an anti-reflection (AR) coating to minimize light reflection. Good AR coatings can reduce reflectivity to values as low as 10^{-4} or 10^{-6}. Different semiconductor layers of the device are grown by MBE or MOCVD (for III-V semiconductors). The top metal contact is usually in the form of a circular ring. The top p+ layer is highly doped and kept thin so that most of the light is absorbed in the depletion region. The light intensity in the semiconductor decays as, $I(x) = I_o e^{-\alpha x}$. The photogeneration rate $G_L(x)$ at depth x in the device is therefore,

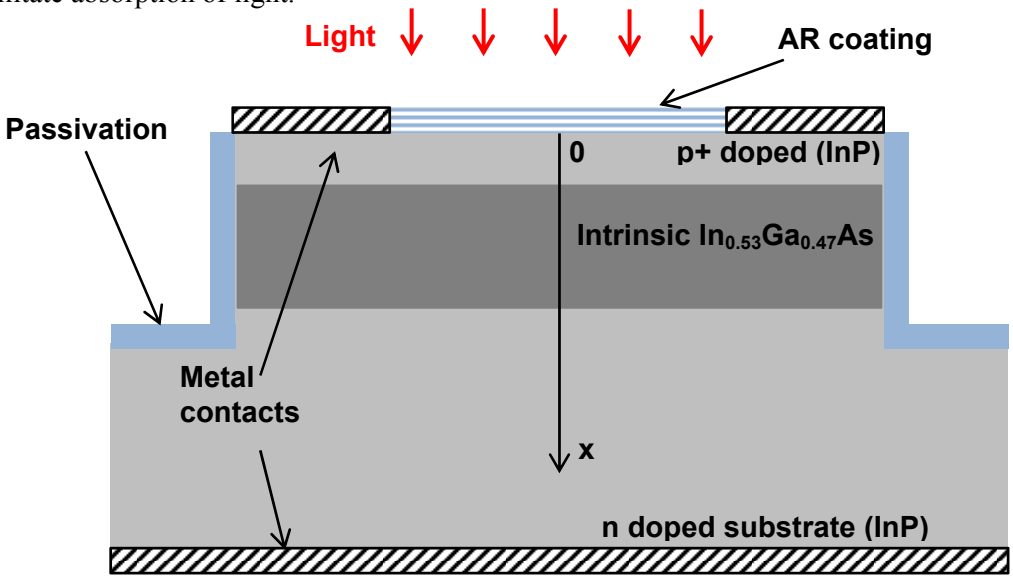
$$G_L(x) = \alpha \frac{I_o}{\hbar\omega} e^{-\alpha x}$$



In order to get most of the light absorbed in the depletion region (as opposed to in the quasineutral regions), one needs a depletion region of thickness of the order of $1/\alpha$. This can be accomplished more easily with a pin structure instead of a pn structure, as show below.



The intrinsic layer is typically a few microns thick. In order to completely eliminate light absorption in the top quasineutral region heterostructure pin photodiodes are used. The Figure below shows the structure of a commercial InGaAs/InP heterostructure pin photodetector for 1.55 μm wavelength light for optical fiber communications. In these devices, the top p+ layer and the bottom n+ layer and the substrate are wide bandgap semiconductors (InP in the Figure below) that do not absorb photons at the wavelength of interest. The intrinsic light absorption layer (InGaAs in the Figure below) has a smaller bandgap to facilitate absorption of light.



3.2.13 Some Figures of Merit of Photodiodes:

The **responsivity** (units: Amps/watt) of a photodiode is defined as,

$$R = \frac{I_L}{P_{inc}}$$

where P_{inc} is the total light power (units: energy per second) incident on the detector. The incident intensity I_{inc} is,

$$I_{inc} = \frac{P_{inc}}{A}$$

where A is the detector area. Part of the incident light is reflected. If the reflectivity of the top surface of the detector is R, then the intensity of light in the semiconductor is,

$$I(x=0) = I_{inc}(1-R)$$

Inside the semiconductor, light decays as,

$$I(x) = I(x=0)e^{-\int_0^x \alpha(x')dx'}$$

where $\alpha(x)$ is the position dependent absorption coefficient - position dependent because different layers could have different values for α. The photogeneration rate is,

$$G_L(x) = \frac{\alpha(x)}{\hbar\omega} I(x=0)e^{-\int_0^x \alpha(x')dx'}$$

The **quantum efficiency** (also called **the external quantum efficiency**) of a photodiode is defined as,

$$\eta_{ext} = \frac{I_L/q}{P_{inc}/\hbar\omega} = \frac{\hbar\omega}{q}\frac{I_L}{P_{inc}}$$

$\eta = 1$ if every photon incident on the photo detector ends up producing one electron in the external circuit. If one assumes a heterostructure pin photodiode with a depletion region of thickness W then the upper limit for the quantum efficiency is approximately given as,

$$\eta \leq (1-R)\left(1-e^{-\alpha W}\right)$$

The **internal quantum efficiency** of a photodiode is defined as,

$$\eta_{\text{int}} = \frac{I_L/q}{P_{abs}/\hbar\omega} = \frac{\hbar\omega}{q}\frac{I_L}{P_{abs}}$$

Here, P_{abs} is the power which is absorbed in the photodetector and leads to photogeneration. The internal quantum efficiency is a measure of the ability of the photodiode to convert all photogenerated carriers into external current.

3.2.14 Dark Current:

The dark current of a photodetector is the current present even if there no light. Recall that in a reversed biased photodiode the circuit equations are,

$$I = I_o\left(e^{\frac{qV_R}{KT}} - 1\right) - I_L$$

$$IR + V_R + V = 0$$

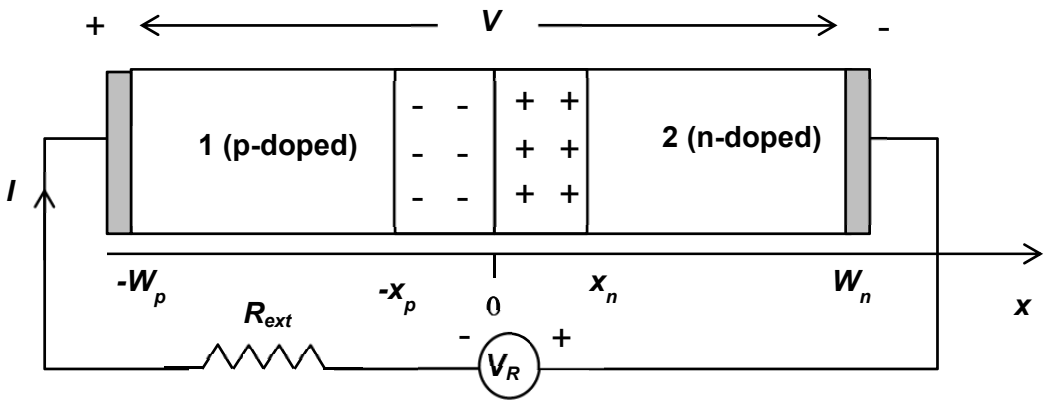
when V_R is large (as is always the case), and R is small, $I = -(I_o + I_L)$. So even if $I_L = 0$ (i.e. no light) the current in the external circuits is equal to $-I_o$ and is called the dark current. This is just the current of a pn diode in heavy reverse bias. Recall that the current in is a reverse biased pn diode is due to thermal generation of electron-hole pains in the quasineutral and in the depletion regions. The dark current of a photovoltaic detector is undesirable as it does not let you detect weak light signals that would produce currents I_L smaller than the dark current I_o. Dark current is also a source of noise (we will not discuss noise in this course). Recall that for a homojunction pn diode (ignoring generation and recombination in the quasineutral regions),

$$I_o \;\alpha\; An_i^2 \left\{ \text{where } \; n_i^2 = \sqrt{N_c N_v}\, e^{\frac{-E_g}{2kT}} \right.$$

Therefore, dark current scales with the area of the device and dark current will be larger for materials with smaller bandgaps. That is why long wavelength detectors are less sensitive than short wavelength detectors. Also note that the dark current will increate with temperature (more thermal generation at higher temperatures).

3.2.15 Photodetector Bandwidth:

In this section we discuss how fast can one operate a photodiode. The speed of operation of a photodetector is important for high data rate fiber-optic communication links. Consider a photodetector connected as shown in the Figure below. We assume that the photodiode is well designed and all photogeneration takes place inside the junction depletion region. The photodetector responsivity is R.



We assume that the optical power incident of the photodiode is time dependent,

$$P_{inc}(t) = P_o + \mathrm{Re}\left\{P(f)e^{j2\pi ft}\right\} \qquad \left\{|P(f)| << P_o\right\}$$

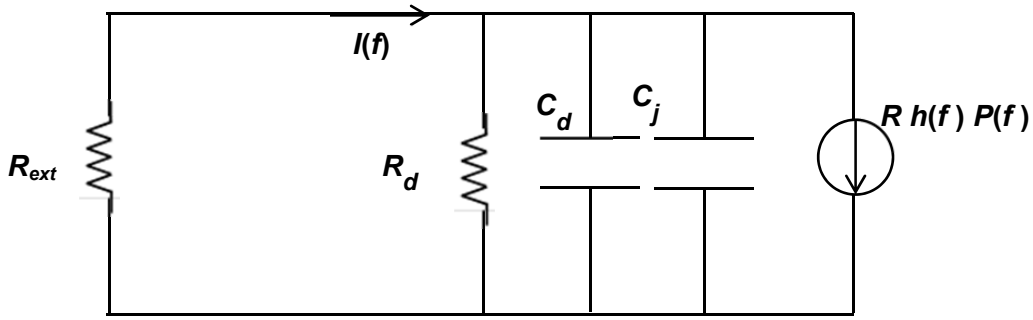
The incident optical power has a dc part P_o and a time-varying part $P(f)$ at a frequency f. Suppose the current $I(t)$ in the external circuit is,

$$I(t) = I_o + \mathrm{Re}\left\{I(f)e^{j2\pi ft}\right\} \left\{|I(f)| << I_o\right\}$$

We know that,

$$I_o = R P_o$$

We need to find $I(f)$. The small signal electrical model of a reverse biased pn diode is shown in the Figure below. The current source represents the current due to photogeneration inside the depletion region of the diode.



Here,

R_d = Diode differential resistance at the bias point $= \dfrac{kT}{qI_o}e^{-qV/KT} == \dfrac{kT}{qI_o}e^{qV_R/KT}$

C_j = Diode junction capacitance

C_d = Diode diffusion capacitance due to charge storage inside the quasineutral regions and can be ignored in a reverse biased diode. We will neglect it.

$h(f)$ = A frequency dependent function that describes the intrinsic frequency response of the photogenerated carriers inside the photodiode.

Since the photodiode is reverse biased, the diode differential resistance R_d is very large (typically several tens of mega Ohms to several hundreds of mega Ohms) and therefore $R_d >> R_{ext}$. From the circuit,

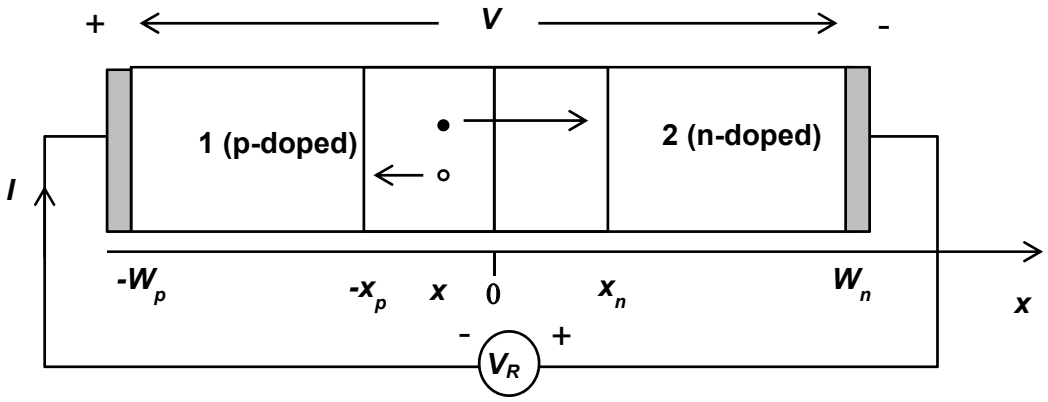
$$I(f) = Rh(f)P(f)\frac{R_d}{R_d + R_{ext} + j2\pi f\, C_j R_d R_{ext}} \approx Rh(f)P(f)\frac{1}{1 + j2\pi f\, C_j R_{ext}} \qquad (1)$$

We still need to determine the function $h(f)$ that describes the intrinsic frequency response of the photogenerated carriers inside the photodiode. The limiting behavior of $h(f)$ is not difficult to argue. We must have,

$$h(f \to 0) \to -1$$
$$h(f \to \infty) \to 0$$

Suppose the carriers are generated inside the junction depletion region. It takes some time for these carriers to come out and contribute to the current in the external circuit and this delay is captured by the function $h(f)$. The extrinsic frequency response of the photodiode is captured by the factor $(1 + j2\pi f\, C_j R_{ext})^{-1}$ and depends on the circuit of the photodiode. The photodetector response will decrease with frequency when the frequency f exceeds the value $1/2\pi C_j R_{ext}$.

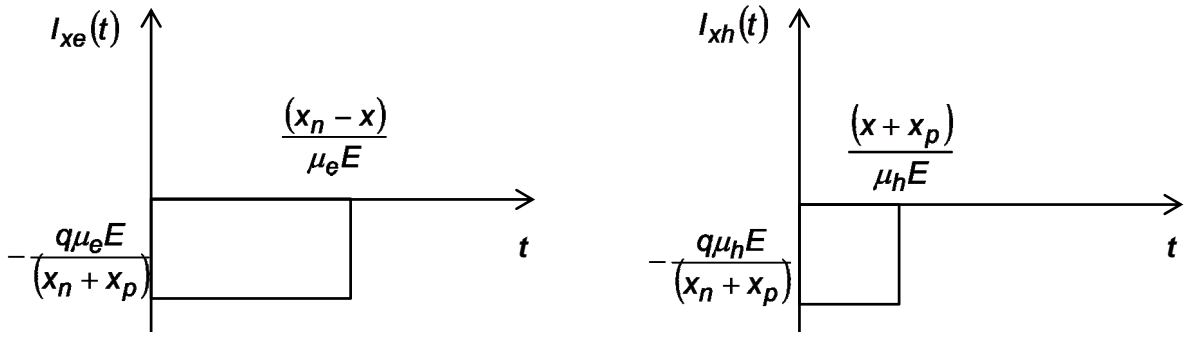
Next, we will compute $h(f)$ for a heterostructure pin diode. As mentioned previously, $h(f)$ is determined by how fast electron-hole pairs generated inside the depletion region make it out as current in the external circuit. For simplicity we will assume that $R_{ext} = 0$ and the only limitation to the speed of operation of the device comes from $h(f)$.



Consider a single electron-hole pair generated at point x inside the junction depletion region at time $t = 0$. We assume that the electric field in the junction is uniform and equals E and the electron and hole mobilities throughout the depletion regions are μ_e and μ_h, respectively. As the electron moves towards the n-side with a drift-velocity $\mu_e E$ and the hole moves towards the p-side with velocity $\mu_h E$, the current flowing in the external circuit (due to capacitive couplings) is,

$$-\frac{q\mu_e E}{x_n + x_p} - \frac{q\mu_h E}{x_n + x_p}.$$

The above result is the famous Ramo-Shockley theorem. The current for each carrier flows for a short duration till that carrier reaches the edge of the depletion region. The current pulse in the external circuit from this photogenerated electron-hole pair is depicted in the Figure below.



The total charge that flows in the external circuit from this one electron-hole pair is,

$$\int_o^\infty [I_{xe}(t) + I_{xh}(t)]\, dt = -q\,\frac{\mu_e E}{(x_n + x_p)} \cdot \frac{(x_n - x)}{\mu_e E} - q\,\frac{\mu_h E}{(x_n + x_p)} \cdot \frac{(x + x_p)}{\mu_h E} = -q$$

Electron-hole pairs are generated at all x uniformly. So, on the average, the current waveform is,

$$\langle I(t)\rangle = \frac{1}{(x_n + x_p)} \int_{-x_p}^{x_n} [I_{xe}(t) + I_{xh}(t)]\, dx = -\frac{q}{\tau_e}\left(1 - \frac{t}{\tau_e}\right) \quad \rightarrow 0 \le t \le \tau_e$$

$$-\frac{q}{\tau_h}\left(1 - \frac{t}{\tau_h}\right) \quad \rightarrow 0 \le t \le \tau_h$$

The electron and hole transit times, τ_e and τ_h, are defined as follows,

$$\tau_e = \frac{(x_n + x_p)}{\mu_e E}$$

$$\tau_h = \frac{(x_n + x_p)}{\mu_h E}$$

The average current waveform from one photogenerated electron-hole pair can be approximated as,

$$\langle I(t)\rangle = -\frac{q}{2\tau_e} e^{-\frac{t}{\tau_e}} - \frac{q}{2\tau_h} e^{\frac{t}{\tau_h}} \quad (t \ge 0)$$

$\langle I(t)\rangle$ as given above is actually the impulse response $h(t)$ of the system. So, $h(t) = \langle I(t)\rangle / q$. The Fourier transform of the Impulse response gives us the function $h(f)$,

$$h(f) = -\left[\frac{1/2}{1 + j2\pi f\tau_e} + \frac{1/2}{1 + j2\pi f\tau_h}\right]$$

If the photogeneration rate is not uniform but given by $G_L(x)$ then a better way to average the current waveforms is,

$$\langle I(t)\rangle = \frac{\int_{-x_p}^{x_n} [I_{xe}(t) + I_{xh}(t)]\, G_L(x) dx}{\int_{-x_p}^{x_n} G_L(x) dx}$$

Regardless of the shape of $G_L(x)$, the poles in $h(f)$ are roughly related to the time it takes the electrons and the holes to cross the depletion region. Typically the holes have smaller mobilities compared to the electrons and move slower. The complete frequency response of the photodiode can then be written as,

$$I(f) = Rh(f)P(f)\frac{1}{1+j2\pi f\ C_j R_{ext}}$$

$$= -RP(f)\left[\frac{1/2}{1+j2\pi f\tau_e} + \frac{1/2}{1+j2\pi f\tau_h}\right]\left[\frac{1}{1+j2\pi f\ C_j R_{ext}}\right]$$

The above expression shows the frequency limitations of a photodiode comes from both the intrinsic as well as the extrinsic (circuit level) factors. The intrinsic frequency limitations pose a fundamental limit.

3.2.16 Bandwidth-Efficiency Tradeoff in Photodiodes:

The quantum efficiency of an ideal pin heterostructure diode is limited by,

$$\eta \le (1-R)\left(1-e^{-\alpha W}\right)$$

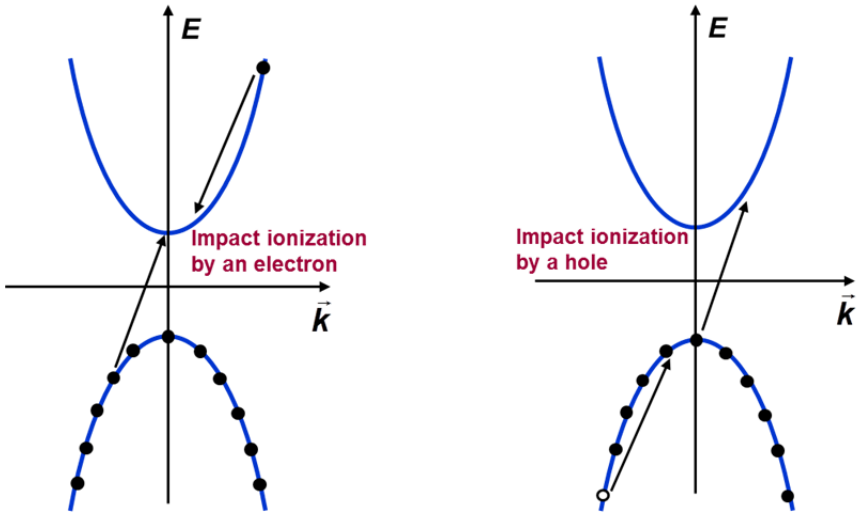
The quantum efficiency η increases as the total depletion region width $W = (x_n + x_p)$ increases. On the other hand, the intrinsic frequency bandwidth is equal to the smaller one of $1/2\pi\tau_e$ or $1/2\pi\tau_h$. This intrinsic bandwidth decreases as W increases (since it takes longer for the photogenerated carriers to cross the depletion region). Therefore, there is a fundamental tradeoff between efficiency and bandwidth of a photodiode. Also note that the extrinsic frequency bandwidth, given by $1/2\pi C_j R_{ext}$, increases as W increases. For high efficiency design, W is chosen very large. For high speed design, W is chosen small but not so small that the extrinsic bandwidth becomes too small to be useful.

3.3 Avalanche Photodiodes

Avalanche detectors are typically used to detect weak light signals when intrinsic current gain is required out of a photodiode. The basic principle used in avalanche photodiodes is that if photogenerated electrons and holes are accelerated to high energies in an electric field then they can acquire enough energies to create additional electron-hole pairs by a process known impact ionization.

3.3.1 Impact Ionization in Semiconductors

Impact ionization is a process in which either a high kinetic energy electron or a high kinetic energy hole loses its kinetic energy and creates and additional electron-hole pair. These processes are depicted in the Figure below.



The impact ionization process must conserve momentum and energy. These conservation rules imply that the minimum kinetic energy needed for electron and hole impact ionization processes are,

$$E_{\text{min}}(\text{electron}) = E_g\left(1+\frac{m_e}{m_e+m_h}\right)$$

$$E_{\text{min}}(\text{hole}) = E_g\left(1+\frac{m_h}{m_e+m_h}\right)$$

Impact ionization rates are characterized by impact ionization coefficients α_e and α_h. These are defined as follows,

α_e = The electron ionization coefficient (units: 1/cm) defined as the number of electron-hole pairs created by one electron in unit distance of travel.

α_h = The hole ionization coefficient (units: 1/cm) defined as the number of electron-hole pairs created by one hole in unit distance of travel.

In general, $\alpha_e \neq \alpha_h$. These ionization coefficients depend strongly on the magnitude E of the electric field since it is the electric field which accelerates the carriers thereby giving them enough energy to cause impact ionization. The electric field dependence is generally exponential,

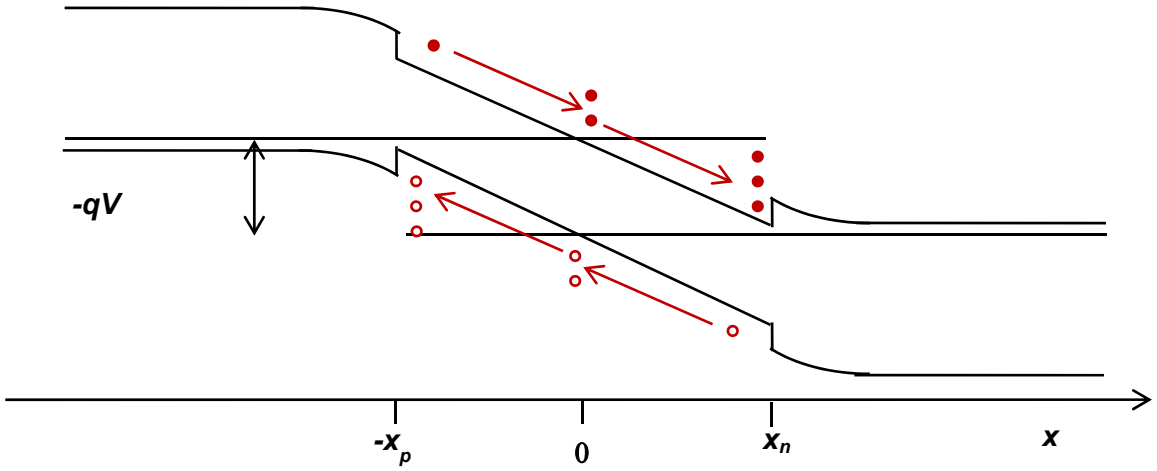
$$\alpha_e = A_e e^{-C_e/E^{\gamma}}$$

$$\alpha_h == B_h e^{-C_h/E^{\delta}}$$

The additional factors appearing are material constants.

3.3.2 Avalanche Photodiode Operation in Steady State:

Consider a strongly reverse biased heterostructure pin photodiode with a very wide intrinsic region for light absorption, as shown below.



The equations for electron and the hole currents assuming no generation-recombination other than generation due to impact ionization and light,

$$-\frac{\partial J_e(x)}{\partial x} = q[G_e(x)+G_L(x)] = \alpha_e|J_e(x)| + \alpha_h|J_h(x)| + qG_L(x)$$
$$\Rightarrow \frac{\partial J_e(x)}{\partial x} = \alpha_e J_e(x) + \alpha_h J_h(x) - qG_L(x) \quad (1)$$

$$\frac{\partial J_h(x)}{\partial x} = q[G_h(x)+G_L(x)] = \alpha_e|J_e(x)| + \alpha_h|J_h(x)| + qG_L(x)$$
$$\Rightarrow \frac{\partial J_h(x)}{\partial x} = -\alpha_e J_e(x) - \alpha_h J_h(x) + qG_L(x) \quad (2)$$

Note that, as always, the total current, given as,

$$J_T = J_e(x) + J_h(x)$$

is always independent of position in steady state. Using, $J_h(x) = J_T - J_e(x)$, in Equation (1) we get,

$$\frac{\partial J_e(x)}{\partial x} - (\alpha_e - \alpha_h)J_e(x) = +\alpha_h J_T - qG_L(x)$$

After integrating one obtains,

$$J_e(x_n) = J_e(-x_p)e^{(\alpha_e-\alpha_h)(x_n+x_p)} - \int_{-x_p}^{x_n}[qG_L(x') - \alpha_h J_T]e^{(\alpha_e-\alpha_h)(x_n-x')}\,dx'$$

Assuming constant illumination ($G(x) = G_L$),

$$J_e(x_n) = J_e(-x_p)e^{(\alpha_e-\alpha_h)(x_n+x_p)} - (qG_L - \alpha_h J_T)\frac{\left[e^{(\alpha_e-\alpha_h)(x_n+x_p)} - 1\right]}{(\alpha_e - \alpha_h)}$$

Assume, as in a regular pin heterostructure photodiode,

$$J_e(-x_p) = 0$$
$$J_h(x_n) = 0.$$

The total current in steady state is,

$$J_T = J_e(x_n) + J_h(x_n) = J_e(x_n)$$

$$\Rightarrow J_T = -(qG_L - \alpha_h J_T)\frac{\left(e^{(\alpha_e-\alpha_h)(x_n+x_p)} - 1\right)}{(\alpha_e - \alpha_h)}$$

$$\Rightarrow J_T = -qG_L\frac{\left(e^{(\alpha_e-\alpha_h)(x_n+x_p)} - 1\right)}{(\alpha_e - \alpha_h) - \alpha_h\left(e^{(\alpha_e-\alpha_h)(x_n+x_p)} - 1\right)}$$

In the absence of impact ionization we would have obtained,

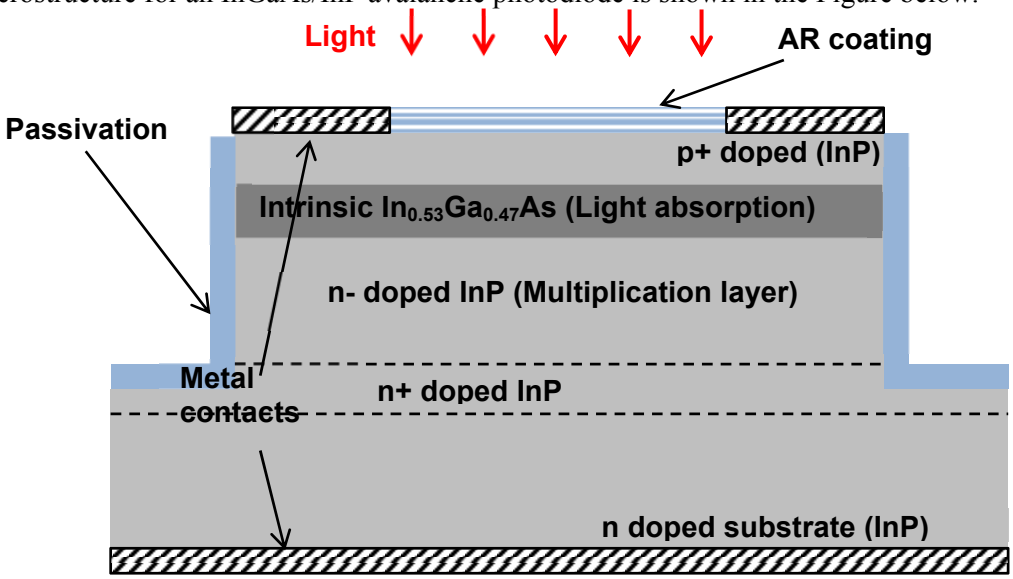
$$J_T = -qG_L(x_n + x_p)$$

So the multiplication gain M of the photodetector due to impact ionization is,

$$M = \frac{J_T}{qG_L(x_n + x_p)} = \frac{1}{(x_n + x_p)}\frac{\left(e^{(\alpha_e-\alpha_h)(x_n+x_p)} - 1\right)}{(\alpha_e - \alpha_h) - \alpha_h\left(e^{(\alpha_e-\alpha_h)(x_n+x_p)} - 1\right)}$$

Avalanche photodetectors are good for high gain but low speed applications. Photogeneration does not usually takes place in the entire high field region where carrier multiplication occurs. Most avalanche

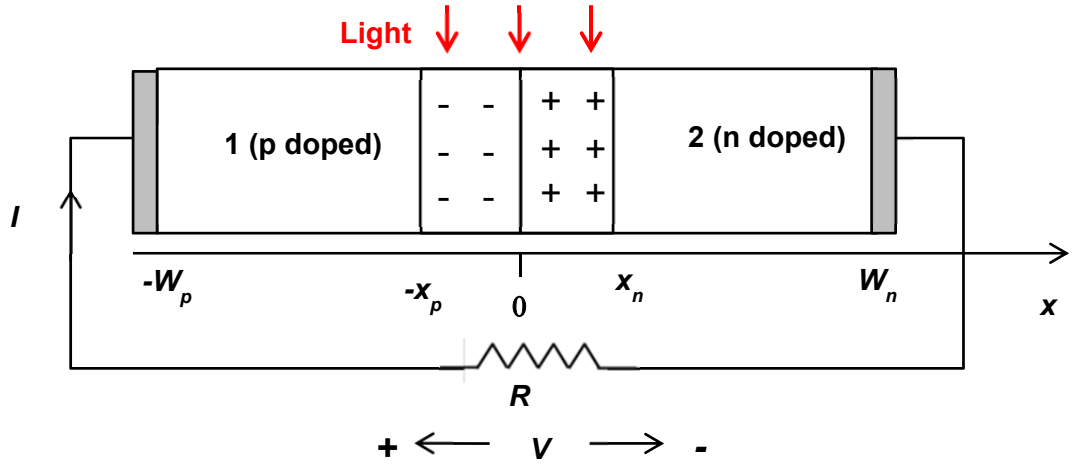
photodiodes have separate regions for light absorption and carrier multiplication. A typical heterostructure for an InGaAs/InP avalanche photodiode is shown in the Figure below.



3.4 Solar Cells

3.4.1 A pn Junction Solar Cell

Photodiodes can be used as solar cells to convert solar energy to electrical energy. Consider the solar cell connected in a circuit, as shown below.

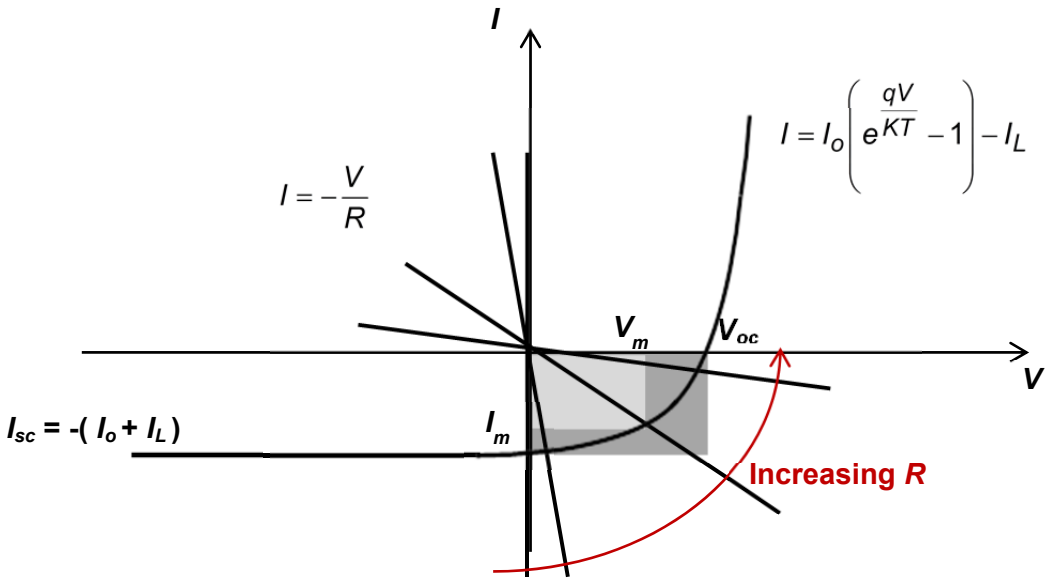


The relevant equations for the cell are,

$$I = I_o\left(e^{\frac{qV}{KT}} - 1\right) - I_L \qquad (1)$$

$$IR + V = 0 \qquad (2)$$

Graphic solutions to these equations are depicted in the Figure below for different values of the resistor R. The solutions, corresponding to the intersection of the curves, represent the operating points of the cell. Note that the pn junction in a solar cell is always forward biased.



If the external resistor R is too large, the voltage delivered by the cell will be large but the current will be small and therefore the power delivered to the resistor, which the product of the current and the voltage, will be small. If the external resistor R is too small, the voltage delivered by the cell will be small but the current will be large and therefore the power delivered to the resistor will again be small. Therefore, there is an optimal value of R which provides the maximum power transfer from the cell to the external resistor. This optimal value can be obtained as follows. First, we maximize the power delivered to the resistor,

$$\frac{d}{dV}\left\{\left[I_o\left(e^{\frac{qV}{KT}} - 1\right) - I_L\right](-V)\right\} = 0$$

$$\Rightarrow e^{\frac{qV_m}{KT}}\left(1 + \frac{qV_m}{KT}\right) = \frac{I_L}{I_o} + 1$$

Here, V_m is the voltage that maximizes the output power of the cell and the corresponding current is,

$$I_m = I_o\left(e^{\frac{qV_m}{KT}} - 1\right) - I_L$$

The power delivered to the resistor is equal to the area of the small lightly shaded rectangle in the Figure above. Once we know the optimal operating point, we can choose the resistor R such that the optimal operating point is the solution of Equations (1) and (2). The optimal value of R is,

$$\frac{1}{R} = \frac{qI_o}{KT}e^{\frac{qV_m}{KT}} = \frac{1}{R_{dm}}$$

Therefore, the optimal value of R is equal to the differential resistance R_{dm} of the diode at the optimal operating point. Note that the power output of the solar cell satisfies,

$$P_{out} = -I_m V_m < -I_{sc}V_{oc}$$

The product $-I_{sc}V_{oc}$ corresponds to the area of the large dark shaded rectangle in the Figure above. The fill factor FF of a solar cell is defined as the ratio of the areas of the two rectangles,

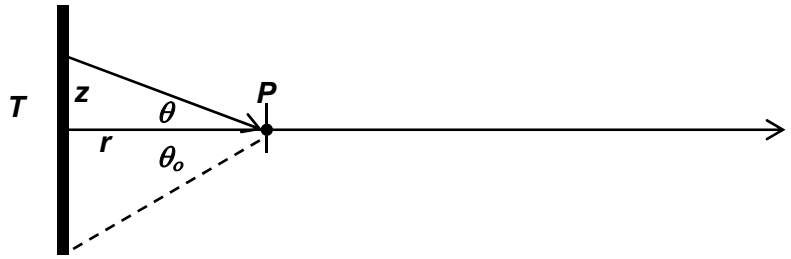
$$FF = \frac{-I_m V_m}{-I_{sc}V_{oc}}$$

Needless to say, a fill factor unity would be highly desirable. The sharper the diode turn-on behavior the larger is the fill factor.

3.5 Fundamental Limits on Solar Energy Conversion with Solar Cells

3.5.1 Sun as a Black Body Radiator

Near the surface of Earth the spectrum of the radiation from the Sun is almost the same as that of a black body at a temperature of $T_S = 5760\,\text{K}$. The solar radiation can thus be modeled as black body radiation. We therefore need to look at the characteristics of black body radiation. Consider a black body source at temperature T, as shown below.



We need to calculate the energy flux per unit area at the point P at a distance r from the source. We need only the component of the photon flux normal to the surface placed at that point (as shown in the Figure). Each mode of radiation with wavevector $\vec{q}$ emitted from the black body will have a photon occupancy given by the Bose-Einstein factor,

$$n(\vec{q}) = \frac{1}{e^{\hbar\omega(\vec{q})/KT} - 1}$$

The point P will receive all radiation from a source point as long as the radiation from that source point has a wavevector in the direction of P. The total photon flux per unit area at P can then be written as an integral,

$$F_N = 2 \times \int_0^{\theta_o} \sin\theta \cos\theta \, d\theta \int_0^{\infty} \frac{2\pi q^2 dq}{(2\pi)^3} \, c\, n(q)$$

The upper limit for the angular integration is the half-angle subtended by the source at the observation point. We can convert the integral into an integral over photon frequencies using the photon density of states,

$$F_N = 2 \times \int_0^{\theta_o} \sin\theta \cos\theta \, d\theta \int_0^{\infty} \frac{2\pi q^2 dq}{(2\pi)^3} \, c\, n(q) = c \frac{\sin^2\theta_o}{4} \int_0^{\infty} d\omega \, g_p(\omega)\, n(\omega)$$

where,

$$g_p(\omega) = \frac{\omega^2}{\pi^2 c^3}$$

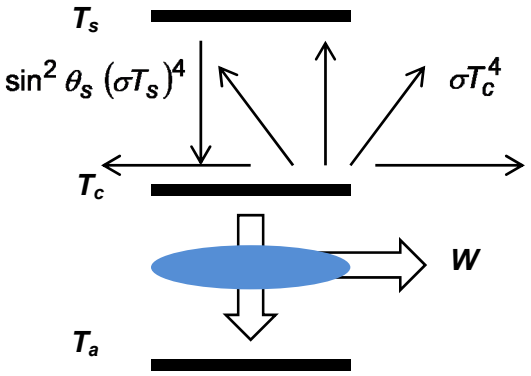
The energy flux at the point P is then,

$$F_E = c \frac{\sin^2\theta_o}{4} \int_0^{\infty} d\omega \, \hbar\omega \, g_p(\omega)\, n(\omega) = \sin^2\theta_o \, \frac{\pi^2}{60} \frac{(KT)^4}{c^2\hbar^3} = \sin^2\theta_o \left(\sigma T^4\right)$$

where σ is the Stephan-Boltzmann constant with a value equal to 5.67×10^{-8} Watt/m^2-K^4. The half-angle subtended by the sun when viewed from the Earth is about 0.267 degrees. Therefore the solar power flux near Earth is around 1350 Watt/m^2.

3.5.2 Fundamental Limit for Solar Energy Conversion Efficiency by any Device

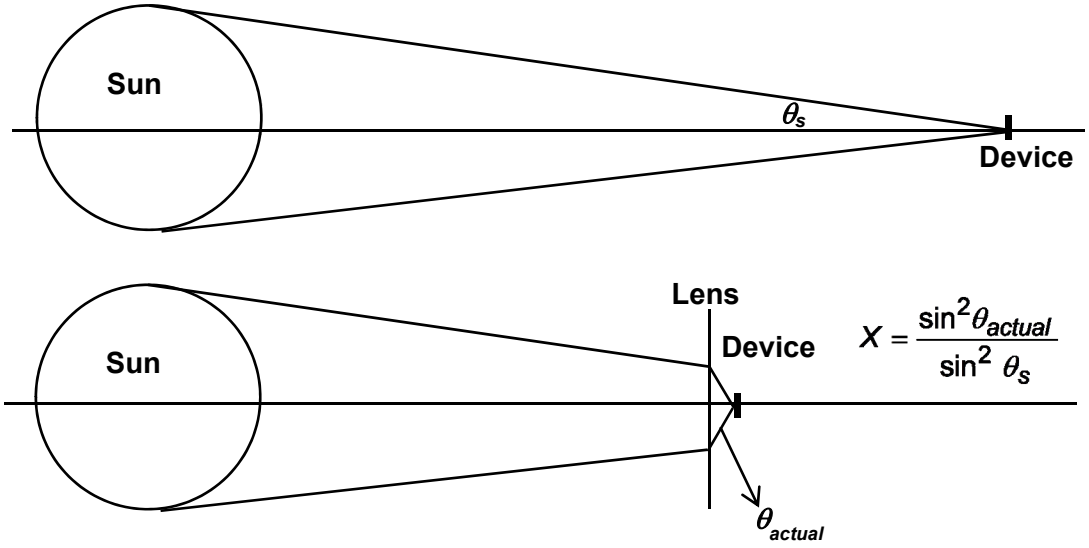
Here we consider the maximum efficiency of a device that converts solar energy into useful work. We consider the following scheme.



A device at temperature T_c absorbs solar energy, radiates a part of it away, and then converts the rest into useful work W by a heat engine operating between the temperature T_c and the ambient temperature on Earth $T_a = 300K$. The net radiative energy absorbed by the device is $\sin^2\theta_s\left(\sigma T_s^4\right) - \sigma T_c^4$. Since the maximum efficiency of the heat engine is $\left(1 - T_a/T_c\right)$ the efficiency of conversion is,

$$\eta = \frac{\left(\sin^2\theta_s\left(\sigma T_s^4\right) - \sigma T_c^4\right)\left(1 - \frac{T_a}{T_c}\right)}{\sin^2\theta_s\left(\sigma T_s^4\right)}$$

One can increase the efficiency by focusing or concentrating sunlight. Focusing can increase the effective half-angle θ_s of the Sun from 0.267 degrees to the maximum possible value of 90 degrees.



The amount of light concentration is described by the quantity X defined as,

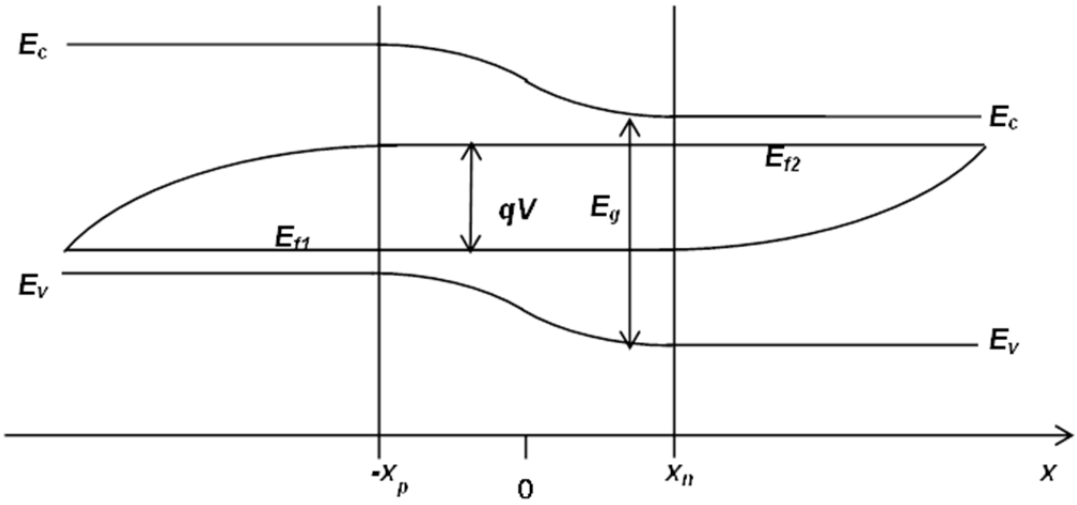
$$X = \frac{\sin^2\theta_{actual}}{\sin^2\theta_s}$$

The maximum possible value is 4.6×10^{4}. Using this value of X, one can maximize the energy conversion efficiency with respect to the temperature T_c of the device. Using this procedure, a maximum efficiency of close to 85% is achieved for $T_c = 2434\ \text{K}$. A conversion efficiency of 85% represents the maximum possible efficiency for solar energy conversion.

3.5.3 Fundamental Limit for Solar Energy Conversion by a Photodiode

We now consider the maximum possible solar energy conversion efficiency of a photodiode. We assume a photodiode that is suitably designed to absorb all incident photons with energies larger than its bandgap E_g and is transparent to all photons with energies smaller than its bandgap. The photodiode is kept at temperature $T_a = 300\ \text{K}$.

During operation, the photodiode will be operating in forward bias with a voltage V across the pn junction. The radiation emitted by the photodiode at temperature $T_a = 300\ \text{K}$ will not have a thermal spectrum and the photon occupancy in the emitted radiation modes will not be given by the Bose-Einstein factor. This is because the photodiode itself is not in thermal equilibrium during operation. To see this more clearly, assume that emitted radiation is in equilibrium with the electrons inside the photodiode (but not in thermal equilibrium!).



From Chapter 3, we have the condition for equilibrium between all emission and absorption processes,

$$n(\omega)\left[f_c(\vec{k}) - f_v(\vec{k})\right] + f_c(\vec{k})\left[1 - f_v(\vec{k})\right] = 0 \quad \text{for} \quad E_c(\vec{k}) - E_v(\vec{k}) = \hbar\omega$$

$$\Rightarrow n(\omega) = \frac{f_c(\vec{k})\left[1 - f_v(\vec{k})\right]}{f_v(\vec{k}) - f_c(\vec{k})} \quad \text{for} \quad E_c(\vec{k}) - E_v(\vec{k}) = \hbar\omega$$

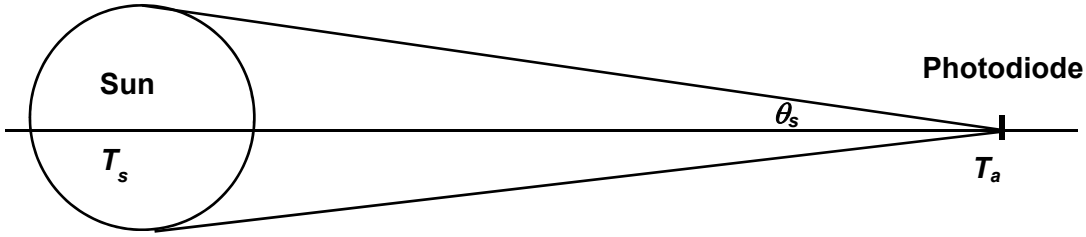
$$\Rightarrow n(\omega) = \frac{1}{e^{(\hbar\omega - qV)/KT_a} - 1} \neq \frac{1}{e^{\hbar\omega/KT_a} - 1}$$

For simplicity, we define quantities related to photon flux and energy flux, $F_N(T,V,\omega_{\min})$ $F_E(T,V,\omega_{\min})$, respectively, as,

$$F_N(T,V,\omega_{\min}) = \frac{c}{4}\int_{\omega_{\min}}^{\infty} d\omega \; g_p(\omega)\frac{1}{e^{(\hbar\omega - qV)/KT} - 1}$$

$$F_E(T,V,\omega_{\min}) = \frac{c}{4}\int_{\omega_{\min}}^{\infty} d\omega\, \hbar\omega\, g_p(\omega)\, \frac{1}{e^{(\hbar\omega - qV)/KT} - 1}$$

We consider a photodiode solar cell of area A facing the Sun, as shown below.



The radiation power per unit area incident on the cell from the Sun is,

$$A\sin^2\theta_s\; F_E(T_s,0,0)$$

The photon flux per unit area incident on the cell from the Sun and absorbed by the cell is,

$$A\sin^2\theta_s\; F_N(T_s,0,E_g/\hbar)$$

The photon flux per unit area incident on the cell from the surrounding ambient and absorbed by the cell is,

$$A\left(1-\sin^2\theta_s\right)F_N(T_a,0,E_g/\hbar)$$

The photon flux per unit area emitted by the cell is,

$$AF_N(T_a,V,E_g/\hbar)$$

If we assume that the internal quantum efficiency of the cell is 100% and each photon absorbed by the cell results in one electron flow in the external circuit then the current can be written as,

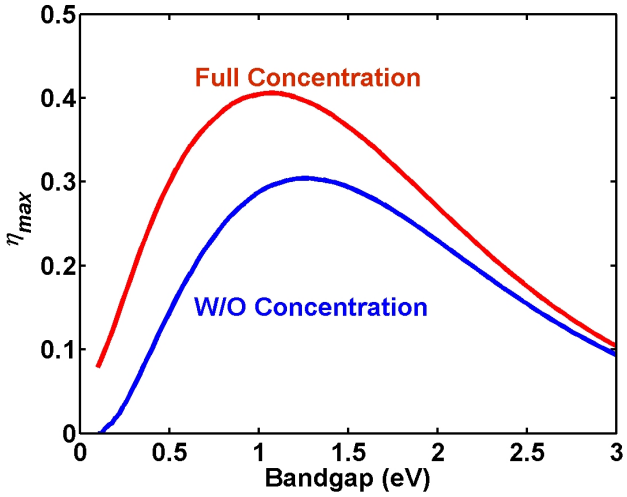
$$I = qA\sin^2\theta_s\; F_N(T_s,0,E_g/\hbar) + qA\left(1-\sin^2\theta_s\right)F_N(T_a,0,E_g/\hbar) - qAF_N(T_a,V,E_g/\hbar)$$

The power delivered to some external resistor would then be IV. The energy conversion efficiency of the cell is therefore,

$$\eta = \frac{IV}{A\sin^2\theta_s\; F_E(T_s,0,0)}$$

$$= \frac{qA\left[\sin^2\theta_s\; F_N(T_s,0,E_g/\hbar) + \left(1-\sin^2\theta_s\right)F_N(T_a,0,E_g/\hbar) - F_N(T_a,V,E_g/\hbar)\right]V}{A\sin^2\theta_s\; F_E(T_s,0,0)}$$

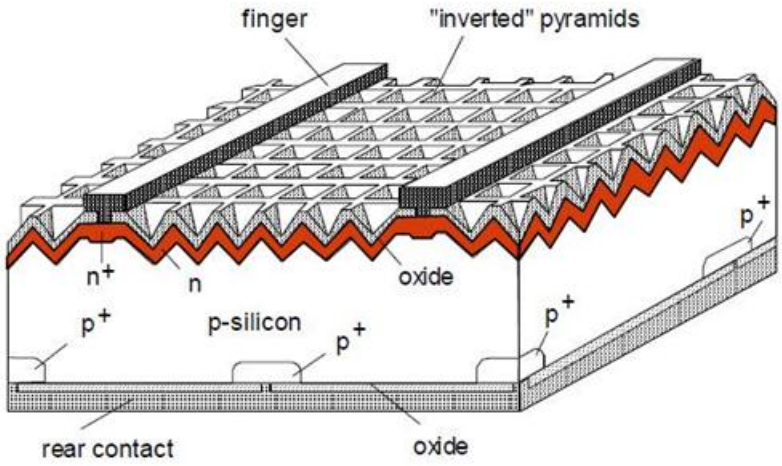
To find out the maximum value of η one must first maximize the product IV with respect to the voltage V and then use the above Equation to calculate η. The resulting values of η are plotted in the Figure below as a function of the semiconductor bandgap E_g. The Figure shows that in the absence of light concentration a maximum efficiency of ~31% is reached at a bandgap of ~1.27 eV and with full concentration a maximum efficiency of ~41% is reached at a bandgap of ~1.1 eV. Light concentration increases efficiency because it increases the absorbed photon flux compared to the emitted photon flux in the numerator in the Equation above. However, some caution needs to be exercised in interpreting these results. Under full concentration the heat load on the device would be immense and we have assumed that the device temperature remains fixed at $T_a = 300\ \text{K}$. Exceptionally good heat sinks would need to be used to keep the device from heating up under full concentration. In addition, most electron-hole recombination rates are highly nonlinear functions of the electron and hole densities and increase rapidly when the carrier densities become large. When light concentration is increased, the result can be poor device internal quantum efficiencies because of increased recombination rates of photogenerated carriers. The Table below shows the typical reported performances of various homojunction solar cells.



Typical Performances of Semiconductor Photocells (Green at al., Prog. Photovolt: Res. Appl., 17, 85 (2009))

Material	Voc (V)	Jsc (Amp/cm^2)	FF (%)	Efficiency (%)
Crystalline Si	0.705	42.7	82.8	25.0
Crystalline GaAs	1.045	29.7	84.7	26.1
Poly-Si	0.664	38.0	80.9	20.4
a-Si	0.859	17.5	63.0	9.5
$CuInGaSe_2$ (CIGS)	0.716	33.7	80.3	19.4
CdTe	0.845	26.1	75.5	16.7

3.5.4 Practical Cell Designs and Modules

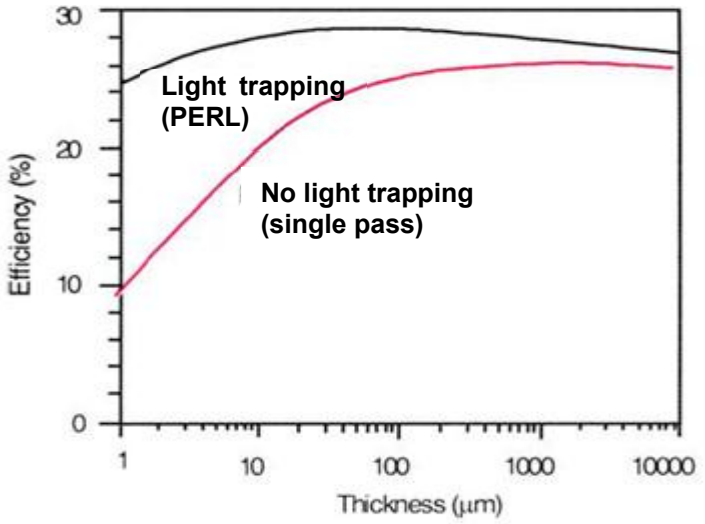


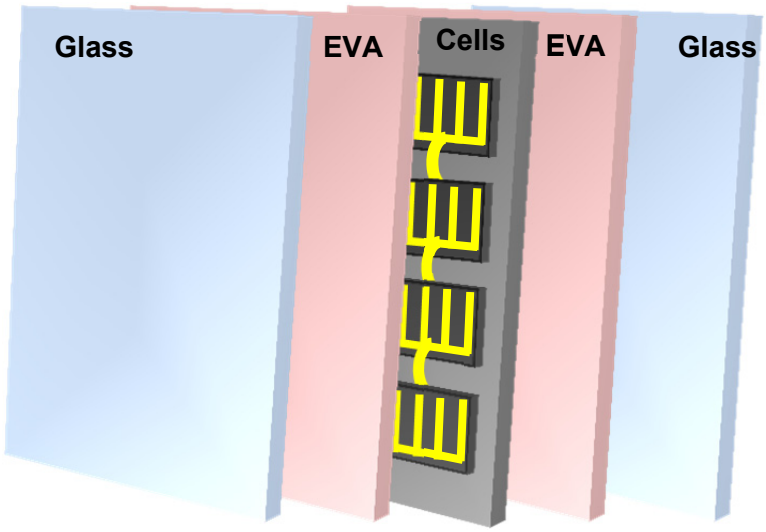
The PERL Si solar cell (Green et al., 1994). Efficiency ~24-25%

The structure of the Si PERL cell (Passivated Emitter, Rear Locally Diffused Cell) after Martin et al. (1994) is shown in the Figure above. It employs a textured front surface to trap light, front surface passivation to reduce surface electron-hole recombination, surface anti-reflection coating, rear surface passivation to reduce electron-hole recombination at the rear surface as well, and highly doped n+ and p+ regions to make good quality metal contacts. The rear metal contact also provides a highly reflecting

surface and helps trap light. The top metal lines occupy a small area in order to not block light from entering the pn junction underneath. The PERL designs have achieved efficiencies between of around 24%-25%.

Note that the absorption coefficient of silicon in the near-IR and visible region of the spectrum ranges from values smaller than 10^3 1/cm to values larger than 10^4 1/cm. Therefore, the absorption lengths range from values smaller than 1 μm to values larger than 10 μm. The widths of the depletion regions in pn junctions are of the order of 1 μm. Therefore, only a small fraction of the solar radiation entering the device is absorbed inside the depletion region in a single pass. The top textured surface and the rear metal reflector cause the light entering the device to bounce back and forth between the front and rear surfaces several times thereby effectively increasing the length of the light absorbing region. The increase in the optical path length through the device is by a factor $\sim 4n^2$, where n is the refractive index of the semiconductor. The thicknesses of the n and p quasineutral regions are kept as small as possible to achieve high quantum efficiencies. Ideally, the thickness of the device (the entire silicon region) should not be much larger than the thickness of the junction depletion region. The Figure below shows the benefits of light trapping in a silicon solar cell with a textured front surface and a rear metal reflector (as in the PERL cell). As expected, the efficiency improvements are large for small device thicknesses.





A solar cell module

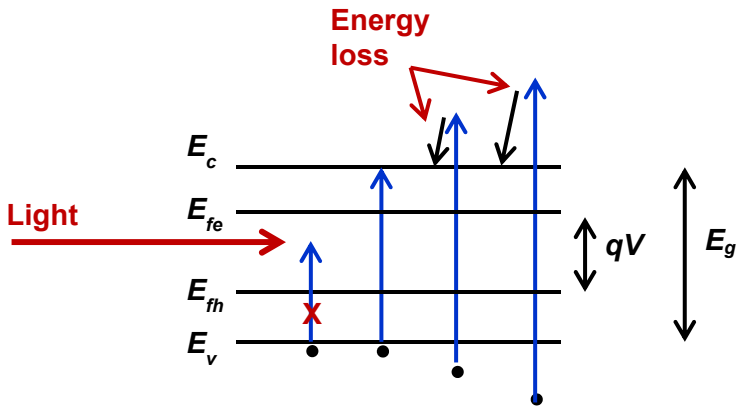
In a solar cell module, several cells (typically ~36) are connected electrically in series. A weather-proof packaging consisting of glass and polymer laminates in used to sandwich the cells, as shown in the Figure

above. The packaging is robust with most manufacturers offering warranties of 10-25 years. We have seen earlier in Section 3.4 that the pn junction of a solar cell needs to be electrically matched to the external resistor in order to deliver the maximum efficiency. When pn junctions are connected electrically in series, as in a silicon solar cell module, ideal matching is not possible. The efficiency of a cell in a module is always less than that of a standalone cell. Best results are obtained when the characteristics of all pn junctions connected in series are identical.

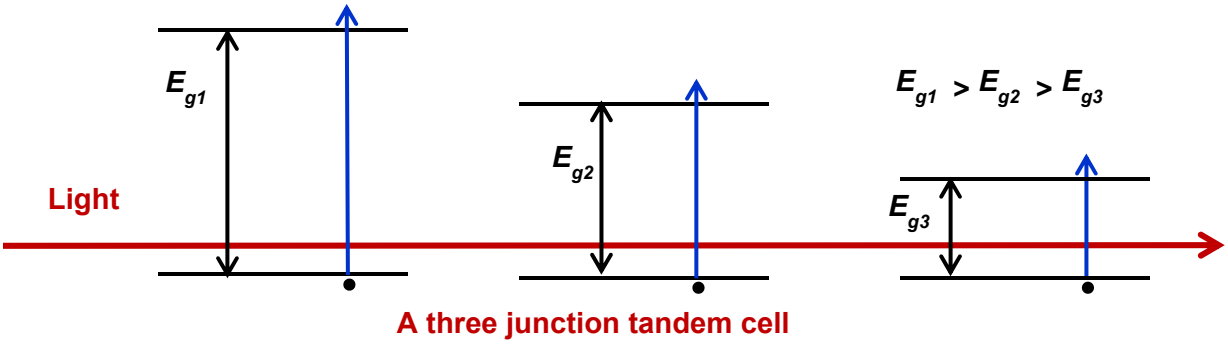
3.5.5 Increasing the Efficiencies of Solar Cells

The best possible efficiency of a solar energy converter is close to 85%, as discussed earlier, whereas the best possible theoretical efficiency of a semiconductor solar cell even under full concentration is only around 41%. The question arises which mechanisms are responsible for efficiency loss and what designs can improve the solar cell efficiencies to get to the 85% thermodynamic limit.

The Figure below shows the main mechanisms responsible for energy loss in a typical solar cell. Solar radiation has photons with energies in a wide spectrum. Photons with energies below the bandgap are not absorbed resulting in energy loss. Photons with energies much larger than the bandgap excite electrons much above the conduction band edge. The electrons subsequently lose energy to phonons and relax to the band edge. Finally, when all electrons come out of one terminal of the solar cell their energy with respect to the other terminal is qV and $qV < qV_{oc} < E_g$. Only under strong photoexcitation (strong light concentration, for example) will qV approach E_g.

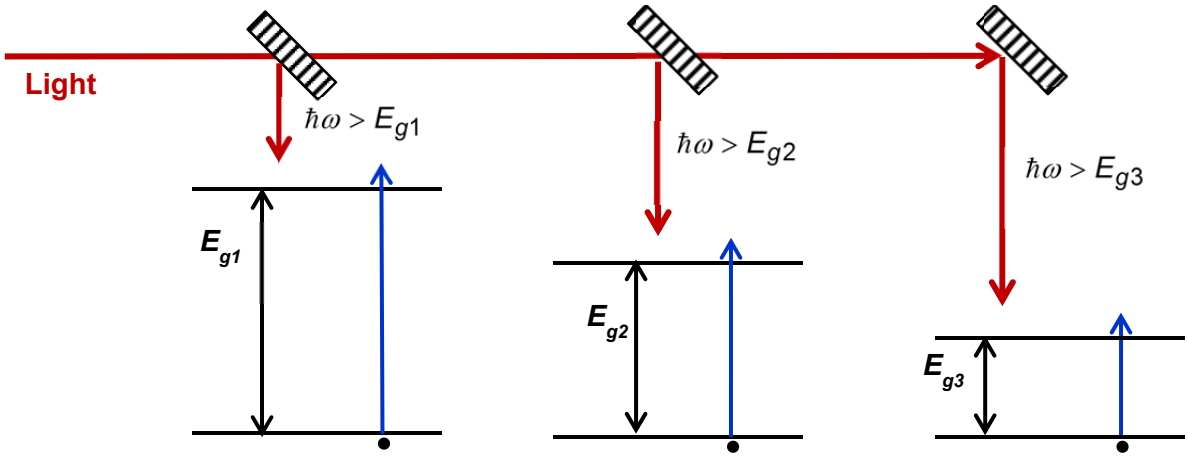


Efficiencies can be improved if all the above energy losses are removed. One way to fix the problem is to use semiconductor pn junctions with different bandgaps connected electrically in series either in the same device or in different devices. Two different schemes to implement this idea are shown in the Figures below.

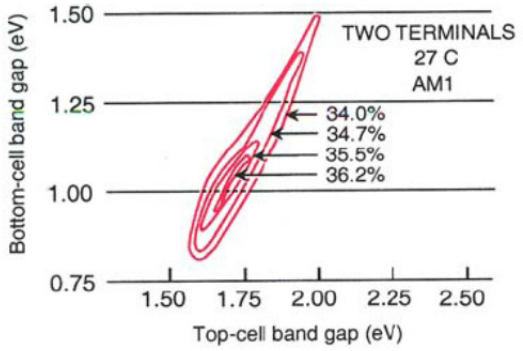


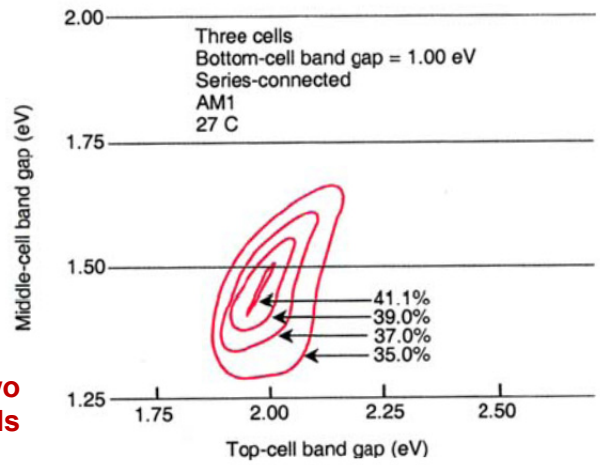
A three junction tandem cell

A three junction tandem cell with spectral filtering



The Figures below show the theoretical maximum efficiencies achievable with two and three junction solar cell devices. A maximum efficiency of 36.2% is achievable with two junctions having bandgaps of ~1.75 eV and ~1.1 eV, and a maximum efficiency of 41.1% is achievable with three junctions having bandgaps of ~1.95 eV, 1.45 eV, and 1.0 eV.



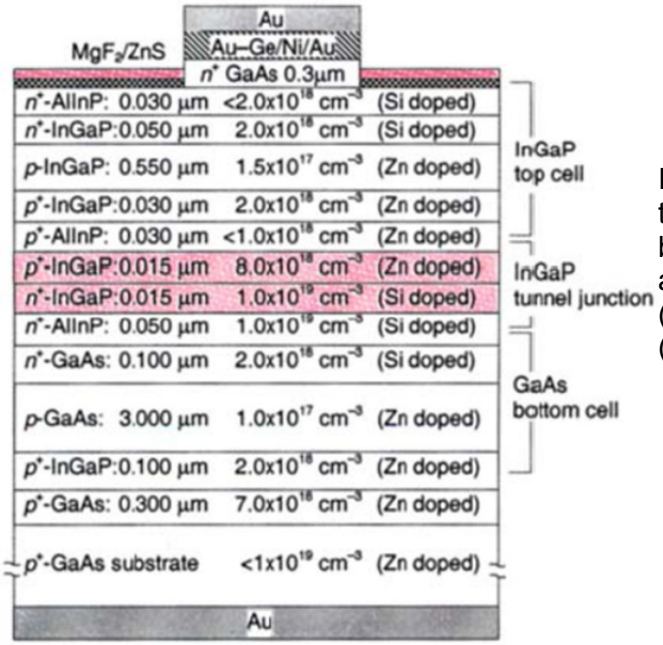


Maximum theoretical efficiencies for two junction and three junction tandem cells (Fan et al., 1982) (w/o concentration)

There are several problems that arise when different kinds of pn junctions are connected electrically in series. The biggest challenge is that the IV characteristics of pn junctions of different materials are difficult, if not impossible, to match. Second, materials which have the bandgaps needed to achieve the maximum efficiencies in a tandem cell are not necessarily lattice matched and cannot be grown together while maintaining a high crystal quality. Third, cost of tandem cells can also become too large for many applications. One of the best results is described below.

3.5.6 InGaP/GaAs Two-Junction Tandem Cell:

The tandem combination of an $In_{0.5}Ga_{0.5}P$ top cell with a direct band gap of ~1.9 eV and a GaAs bottom cell with a direct band gap of ~1.43 eV has a theoretical efficiency of 34%. The lattice parameters and current of these cells can both be matched. These cells have demonstrated a maximum efficiency of 30.3% (w/o concentration) (Takamoto et al., 1997). The device layer structure (not to scale) is shown in the Figure below. The two junctions are connected electrically in series via a reverse biased Esaki tunnel junction. The substrate used in GaAs.



InGaP/GaAs two-junction tandem cell with a reversed biased Esaki tunnel junction and an efficiency of 30.3% (w/o concentration) (Takamoto et al., 1997).

The Figure below is useful for designing devices. It shows the bandgaps (and the corresponding light wavelength) versus the lattice constants of different semiconductors.