Semiconductor Optoelectronic Devices

Cheng Wang

Phone: 20685263
Office: SIST 401E
wangcheng1@shanghaitech.edu.cn

上海科技大学
ShanghaiTech University

Part III Photodetectors

Important relations

- Electric field vs. charge density (Possion's equation)

$$\frac{\partial E}{\partial x} = \frac{\rho}{\varepsilon}$$

$$\rho(x) = q(p - n + N_D^+ - N_A^-)$$

- Potential vs. electric field

$$E = -\frac{\partial \phi}{\partial x}$$

- Potential energy vs. potential

$$\xi = -q\phi$$

Important relations

- Current density vs. carrier drift & diffusion (Carrier transport equation)

$$J_n = qu_n nE + qD_n \frac{\partial n}{\partial x}$$

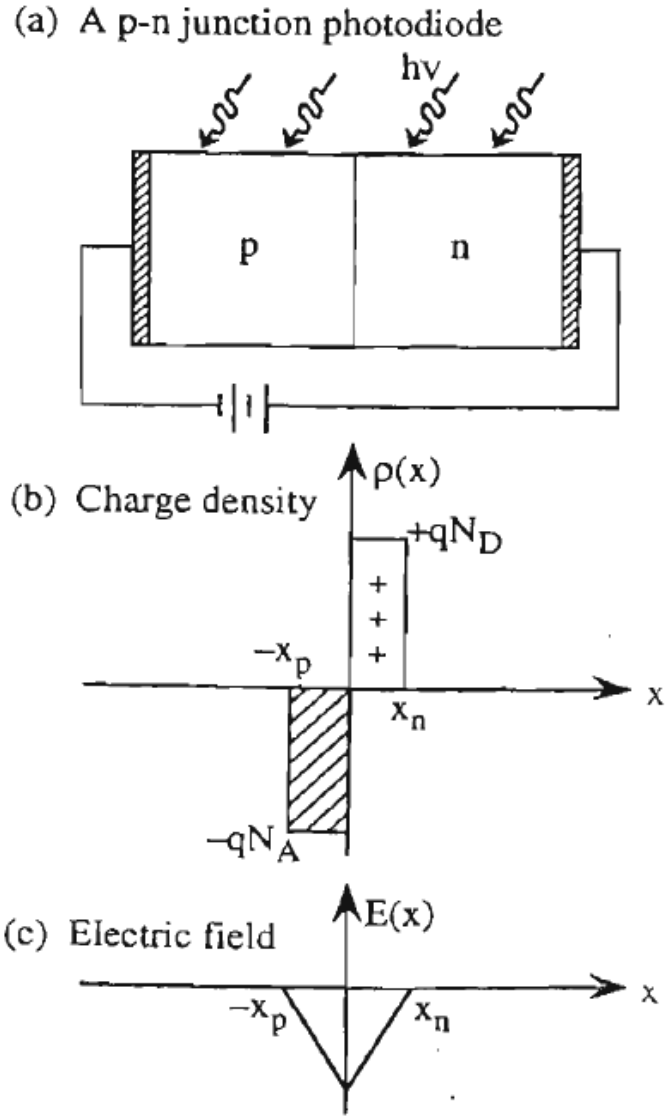
$$J_p = qu_p pE - qD_p \frac{\partial p}{\partial x}$$

- Carrier density dependence on time (Continuity equation)

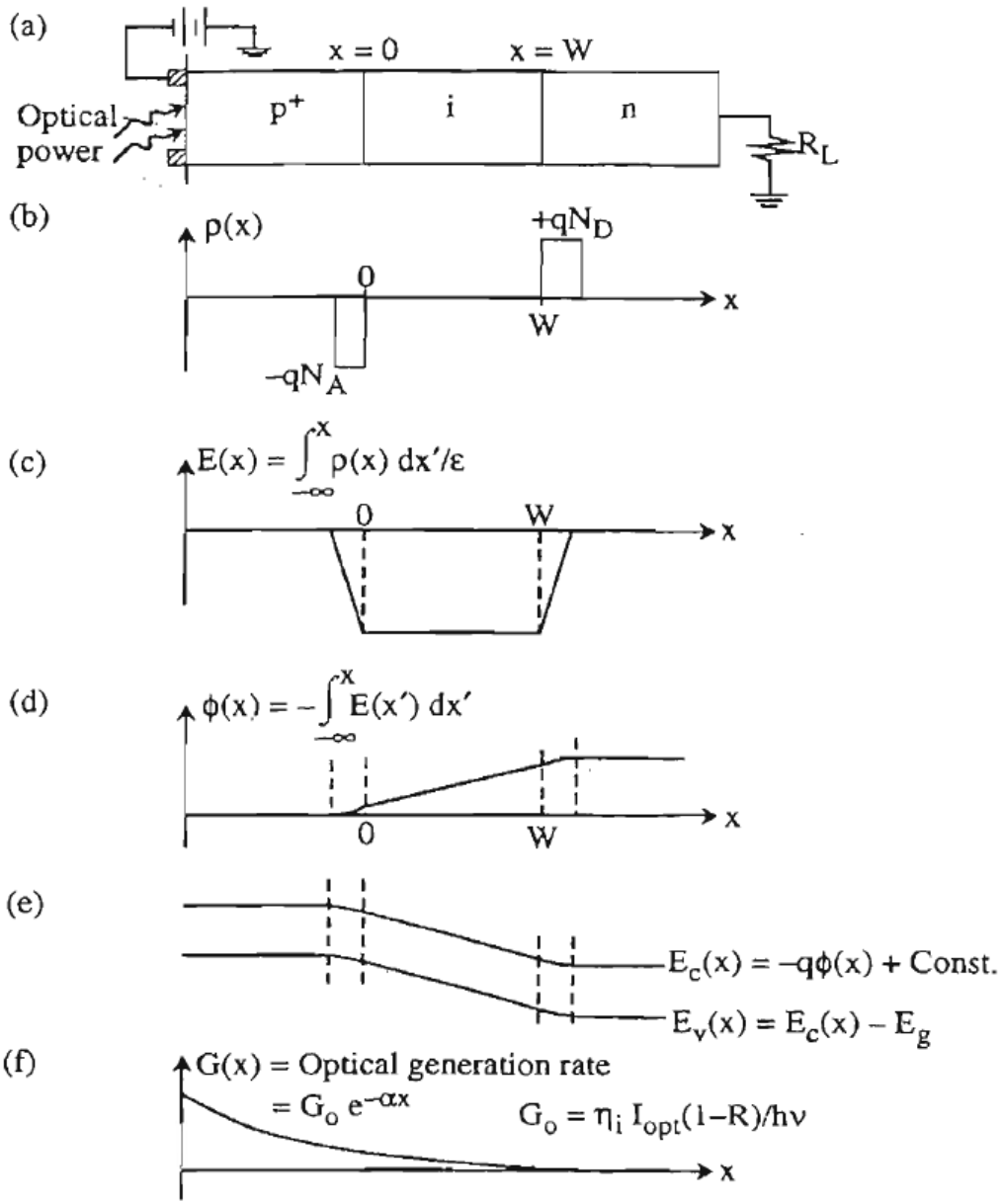
$$\frac{\partial n}{\partial t} = G_n - R_n + \frac{1}{q}\frac{\partial J_n}{\partial x}$$

$$\frac{\partial p}{\partial t} = G_p - R_p - \frac{1}{q}\frac{\partial J_p}{\partial x}$$

Pn PDs



Pin PDs



Lecture 26

- pin photodiode frequency response

  ---Transit-limited bandwidth

  ---Capacitance-limited bandwidth

Transit-limited bandwidth

- Neglecting the diffusion current, and assuming constant electric field, the continuity equations in the deplete region becomes

$$\frac{\partial p}{\partial t} = G_o(x,t) - \frac{1}{q}\frac{\partial J_h}{\partial x}$$
$$\frac{\partial n}{\partial t} = G_o(x,t) + \frac{1}{q}\frac{\partial J_n}{\partial x},$$

- The drift current density

$$J_h = q v_{h,\text{sat}} p, \quad J_n = q v_{n,\text{sat}} n$$

Transit-limited bandwidth

- In the frequency domain, the continuity equations are

$$j\omega p(x) = G_o(x) - v_{h,\text{sat}}\frac{dp(x)}{dx} = G_o(0)e^{-\alpha x} - v_{h,\text{sat}}\frac{dp(x)}{dx}$$

$$j\omega n(x) = G_o(x) + v_{n,\text{sat}}\frac{dn(x)}{dx} = G_o(0)e^{-\alpha x} + v_{n,\text{sat}}\frac{dn(x)}{dx},$$

- The boundary conditions are

$$p(W) = 0, n(0) = 0$$

- That is, in the reverse-bias pi junction, the minority electron density is zero, and the minority hole density in the in junction is also zero.

Transit-limited bandwidth

- The solution for the hole density,

$$p(x) = \frac{G_o(0,\omega)}{j\omega - \alpha v_{h,\text{sat}}} e^{-\alpha W}\left[e^{-\alpha(x-W)} - e^{-\frac{j\omega(x-W)}{v_{h,\text{sat}}}}\right]$$

- The solution for the electron density,

$$n(x) = \frac{G_o(0,\omega)}{j\omega + \alpha v_{n,\text{sat}}}\left[e^{-\alpha x} - e^{\frac{j\omega x}{v_{n,\text{sat}}}}\right]$$

- The drift currents are then,

$$J_h(x) = q v_{h,\text{sat}} p(x) = \frac{q v_{h,\text{sat}} G_o(0,\omega)}{j\omega - \alpha v_{h,\text{sat}}} e^{-\alpha W}\left[e^{-\alpha(x-W)} - e^{-\frac{j\omega(x-W)}{v_{h,\text{sat}}}}\right]$$

$$J_n(x) = q v_{n,\text{sat}} n(x) = \frac{q v_{n,\text{sat}} G_o(0,\omega)}{j\omega + \alpha v_{n,\text{sat}}}\left[e^{-\alpha x} - e^{\frac{j\omega x}{v_{n,\text{sat}}}}\right].$$

Transit-limited bandwidth

- The total current is the sum of the drift current and the displacement current,

$$J_t(\omega) = J_h + J_n + \mathrm{j}\omega\epsilon_s E(x,\omega)$$

- The integration gives,

$$\begin{aligned}
\int_0^W J_t(\omega)\,\mathrm{d}x &= W J_t(\omega) = \int_0^W \left[J_h(x) + J_n(x) + \mathrm{j}\omega\epsilon_s E(x,\omega)\right]\mathrm{d}x \\
&= \frac{q v_{h,\mathrm{sat}} G_o(0,\omega)}{\mathrm{j}\omega - \alpha v_{h,\mathrm{sat}}} \mathrm{e}^{-\alpha W} \int_0^W \left(\mathrm{e}^{-\alpha(x-W)} - \mathrm{e}^{-\frac{\mathrm{j}\omega(x-W)}{v_{h,\mathrm{sat}}}}\right)\mathrm{d}x \\
&\quad + \frac{q v_{n,\mathrm{sat}} G_o(0,\omega)}{\mathrm{j}\omega + \alpha v_{n,\mathrm{sat}}} \int_0^W \left(\mathrm{e}^{-\alpha x} - \mathrm{e}^{\frac{\mathrm{j}\omega x}{v_{n,\mathrm{sat}}}}\right)\mathrm{d}x + \mathrm{j}\omega\epsilon_s \int_0^W E(x,\omega)\,\mathrm{d}x \\
&= \frac{q v_{h,\mathrm{sat}} G_o(0,\omega)}{\mathrm{j}\omega - \alpha v_{h,\mathrm{sat}}} \mathrm{e}^{-\alpha W} \left[\frac{\mathrm{e}^{-\alpha(x-W)}}{-\alpha} - \frac{\mathrm{e}^{-\frac{\mathrm{j}\omega(x-W)}{v_{h,\mathrm{sat}}}}}{-\mathrm{j}\omega/v_{h,\mathrm{sat}}}\right]_0^W \\
&\quad + \frac{q v_{n,\mathrm{sat}} G_o(0,\omega)}{\mathrm{j}\omega + \alpha v_{n,\mathrm{sat}}} \left[\frac{\mathrm{e}^{-\alpha x}}{-\alpha} - \frac{\mathrm{e}^{\frac{\mathrm{j}\omega x}{v_{n,\mathrm{sat}}}}}{\mathrm{j}\omega/v_{n,\mathrm{sat}}}\right]_0^W + \mathrm{j}\omega\epsilon_s\,[-V]_0^W .
\end{aligned}$$

Transit-limited bandwidth

- The total current in the frequency domain is then given by

$$J_t(\omega) = \frac{1}{W}\frac{q v_{h,\mathrm{sat}} G_o}{\mathrm{j}\omega - \alpha v_{h,\mathrm{sat}}} \mathrm{e}^{-\alpha W}\left(\frac{\mathrm{e}^{\alpha W}-1}{\alpha} + \frac{1-\mathrm{e}^{\frac{\mathrm{j}\omega W}{v_{h,\mathrm{sat}}}}}{\mathrm{j}\omega / v_{h,\mathrm{sat}}}\right)$$
$$+ \frac{q v_{n,\mathrm{sat}} G_o}{\mathrm{j}\omega + \alpha v_{n,\mathrm{sat}}}\left(\frac{1-\mathrm{e}^{-\alpha W}}{\alpha} + \frac{1-\mathrm{e}^{\frac{\mathrm{j}\omega W}{v_{n,\mathrm{sat}}}}}{\mathrm{j}\omega / v_{h,\mathrm{sat}}}\right) + \mathrm{j}\omega\frac{\epsilon_s}{W}[V(0) - V(W)]$$

- The carrier transit time is defined as

$$\tau_{dr,n} = \frac{W}{v_{n,\mathrm{sat}}}, \quad \tau_{dr,h} = \frac{W}{v_{h,\mathrm{sat}}}$$

Transit-limited bandwidth

- Using the optical generation rate expression,

$$G_o(0,\omega) = \eta_Q \frac{(1-R)}{Ahf} \alpha P_{in}(\omega)$$

- The PD total current becomes

$$\begin{aligned} I_t(\omega) &= \alpha W \frac{q}{hf} \eta_Q (1-R) P_{in}(\omega) \\ &\quad \times \left\{ \frac{e^{-\alpha W} - 1}{\alpha W (\alpha W - j\omega\tau_{dr,h})} + e^{-\alpha W} \frac{e^{j\omega\tau_{dr,h}} - 1}{j\omega\tau_{dr,h} (\alpha W - j\omega\tau_{dr,h})} \right. \\ &\quad \left. + \frac{1 - e^{-\alpha W}}{\alpha W (j\omega\tau_{dr,n} + \alpha W)} + \frac{1 - e^{j\omega\tau_{dr,n}}}{j\omega\tau_{dr,n} (j\omega\tau_{dr,n} + \alpha W)} \right\} + j\omega \frac{A\epsilon_s}{W} V_A(\omega) \\ &= -I_L(\omega) + j\omega C V_A(\omega). \end{aligned}$$

Transit-limited bandwidth

- For frequency → 0, the total current reduces to the photocurrent,

$$I_t(0) = -I_L(0) = -\frac{q}{hf}\eta_Q\,(1-R)\,P_{in}(0)\left[1-\exp\left(-\alpha W\right)\right]$$

- The small-signal photocurrent in the frequency domain is finally given by

$$I_L(\omega) = \alpha W \frac{q}{hf}\eta_Q\,(1-R)\,P_{in}(\omega) \times \left\{\frac{1}{\alpha W - \mathrm{j}\omega\tau_{dr,h}}\left[\frac{1-\mathrm{e}^{-\alpha W}}{\alpha W} + \mathrm{e}^{-\alpha W}\frac{1-\mathrm{e}^{\mathrm{j}\omega\tau_{dr,h}}}{\mathrm{j}\omega\tau_{dr,h}}\right] - \frac{1}{\alpha W + \mathrm{j}\omega\tau_{dr,n}}\left[\frac{1-\mathrm{e}^{-\alpha W}}{\alpha W} + \frac{1-\mathrm{e}^{\mathrm{j}\omega\tau_{dr,n}}}{\mathrm{j}\omega\tau_{dr,n}}\right]\right\}.$$

Transit-limited bandwidth

- The normalized responsivity of pin PDs is

$$\mathfrak{r}(\omega) = \frac{I_L(\omega)}{I_L(0)} = \frac{1}{\alpha W - j\omega\tau_{dr,h}}\left[\frac{1}{\alpha W} + \frac{1 - e^{j\omega\tau_{dr,h}}}{j\omega\tau_{dr,h}}\frac{1}{e^{\alpha W} - 1}\right] - \frac{1}{\alpha W + j\omega\tau_{dr,n}}\left[\frac{1}{\alpha W} + \frac{1 - e^{j\omega\tau_{dr,n}}}{j\omega\tau_{dr,n}}\frac{e^{\alpha W}}{e^{\alpha W} - 1}\right].$$

- When the diode is thick enough, alpha*W>>1,

$$|\mathfrak{r}(\omega)| \approx \left|\frac{\sin\left(\frac{\omega\tau_{dr,n}}{2}\right)}{\frac{\omega\tau_{dr,n}}{2}}\right|$$

- The corresponding bandwidth is given by

$$f_{3\text{dB},tr} = \frac{2 \times 1.391}{2\pi}\frac{1}{\tau_{dr,n}} = 0.443\frac{v_{n,\text{sat}}}{W}$$

Transit-limited bandwidth

- This 3-dB bandwidth is the transit-time limited cutoff frequency, which depends on the transit time of the minority carriers generated in the most illuminated part of the intrinsic region, close to the p surface side, such carriers are the electrons.
- For a PD illuminated from the back n side, the 3-dB bandwidth becomes

$$f_{3\mathrm{dB},tr} = 0.443\frac{v_{h,\mathrm{sat}}}{W}$$

- Because holes are slower than electrons, illumination should come from the p side to maximize the device speed.

Assuming, on the other hand, that both carriers have the same transit time we obtain the approximate expression

$$f_{3\mathrm{dB},tr} \approx \frac{1}{2.2\tau_t},$$

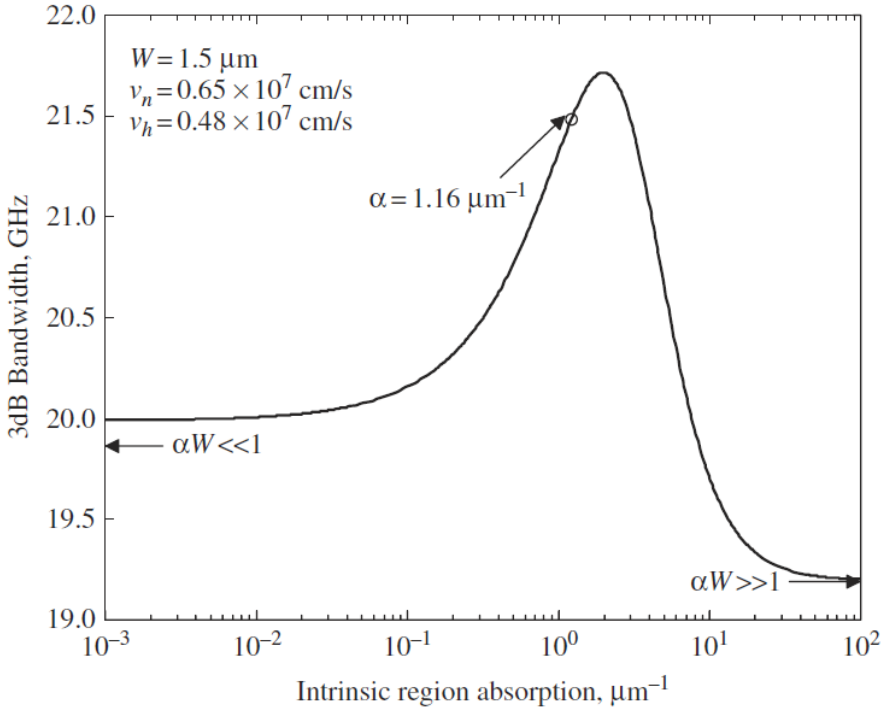
where τ_t is the electron or hole transit time.

Transit-limited bandwidth

- When the diode is thin enough, the 3-dB bandwidth

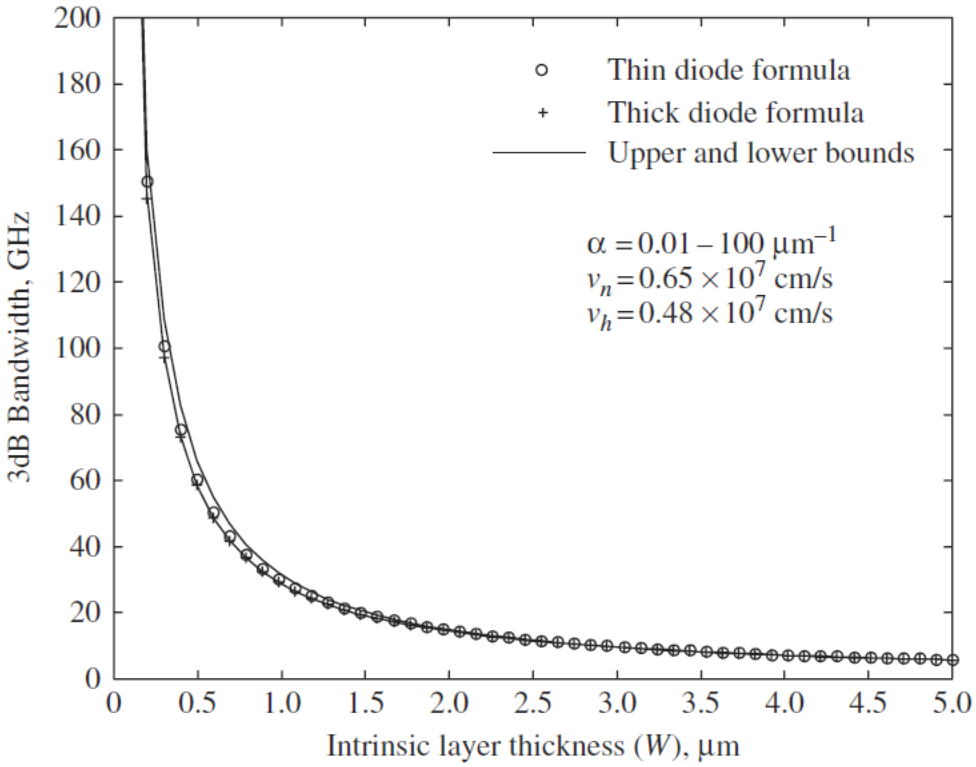
$$f_{3\text{dB},tr} = \frac{3.5\bar{v}}{2\pi W}, \quad \text{where} \quad \frac{1}{\bar{v}^4} = \frac{1}{2}\left(\frac{1}{v_{n,\text{sat}}^4} + \frac{1}{v_{h,\text{sat}}^4}\right)$$

- The bandwidth is weakly dependent on the absorption coefficient.



Transit-limited bandwidth

- While the bandwidth is inversely proportional to intrinsic region width



- Values in excess of 100 GHz can be obtained for very thin diodes, but the responsivity will be low due to the low efficiency. Thus, there will be a speed-efficiency trade-off.

Lecture 26

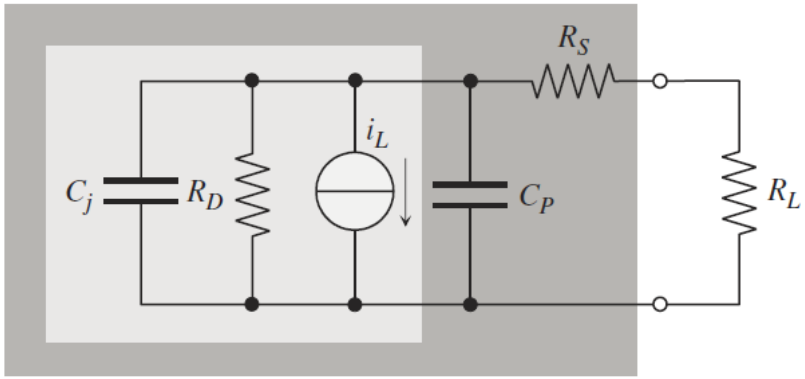
- pin photodiode frequency response

  ---Transit-limited bandwidth

  ---Capacitance-limited bandwidth

Capacitance-limited bandwidth

- The photodiode can be seen as a photocurrent generator with the following equivalent circuit,



- C_p is external parasitic capacitance, R_s is series parasitic resistance; R_L is the load resistance.
- C_j is intrinsic capacitance, R_D is parallel diode resistance. Generally,

$$R_D \gg R_s, R_L$$

Capacitance-limited bandwidth

- The impedance-limited bandwidth is given by,

$$f_{3\text{dB},RC} \approx \frac{1}{2\pi RC},$$

where

$$R \approx R_S + R_L, \quad C \approx C_j + C_p, \quad C_j = \frac{\epsilon_s A}{W}.$$

- The total cutoff frequency resulting from the transit time and the RC effect is approximately,

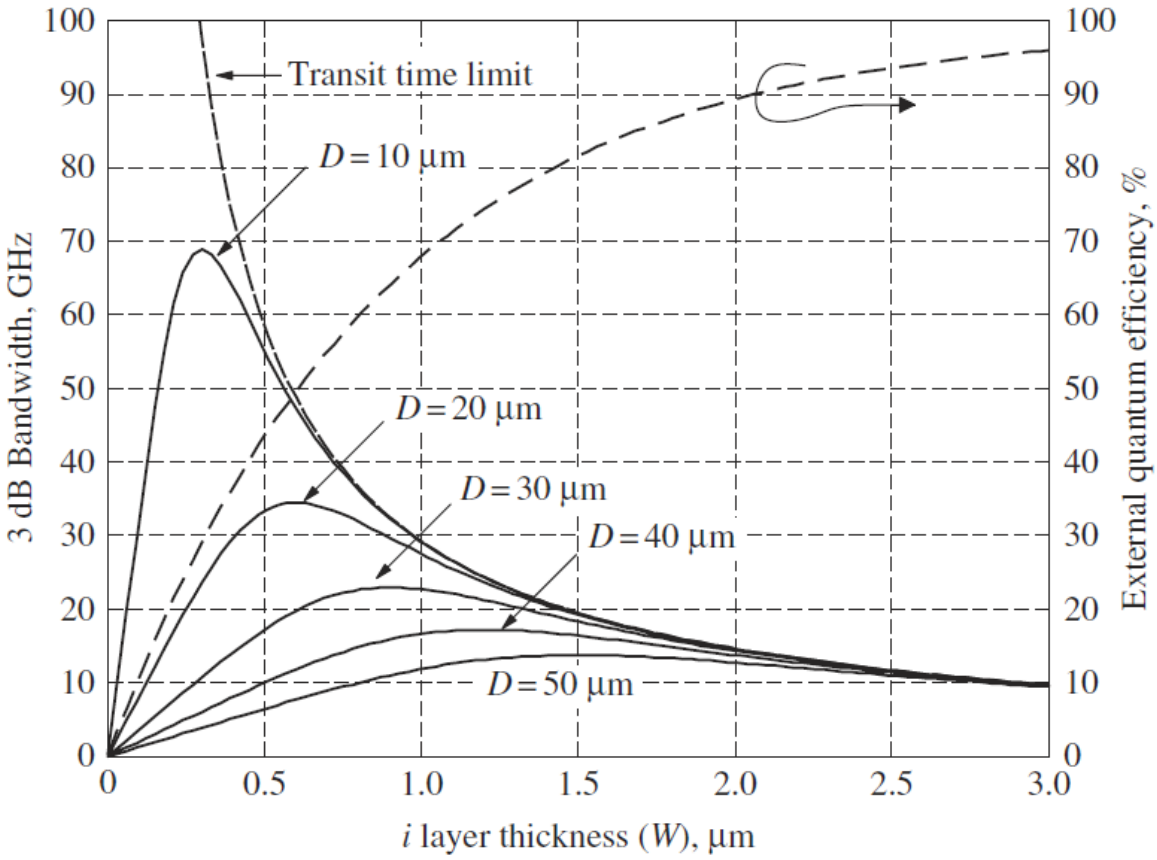
$$f_{3\text{dB}} \approx \frac{1}{\sqrt{f_{3\text{dB},RC}^{-2} + f_{3\text{dB},tr}^{-2}}}$$

Bandwidth-efficiency trade-off

In a vertically illuminated photodiode, optimization of the external quantum efficiency suggests $W >> L_\alpha = 1/\alpha$; moreover, the detection area A should be large in order to improve coupling with the external source (e.g., an optical fiber). However, increasing W increases the RC-limited bandwidth (since it decreases the junction capacitance) but decreases the transit time-limited bandwidth. Increasing the device area has no influence on the transit time-limited bandwidth but makes the capacitance larger and therefore decreases the RC-limited bandwidth. Keeping the device area A constant, we therefore have $f_{3dB,RC} \propto W$ but $f_{3dB,tr} \propto 1/W$. Since $f_{3dB} < \min(f_{3dB,RC}, f_{3dB,tr})$, the total bandwidth is dominated by $f_{3dB,RC} \propto W$ (low W) or $f_{3dB,tr} \propto 1/W$ (large W). The total bandwidth then first increases as a function of W, then decreases; f_{3dB} therefore has a maximum, which shifts toward smaller values of W and larger cutoff frequencies with decreasing A. At the same time, the efficiency always increases with W.

As a consequence, high-frequency operation (high f_{3dB}) requires small-area diodes, with small W and increasingly poor efficiency.

Bandwidth-efficiency trade-off



- Commercial high-speed diodes exhibit responsivities around 0.7-0.9 A/W and have active region thicknesses W< 1 um.

Lecture 27

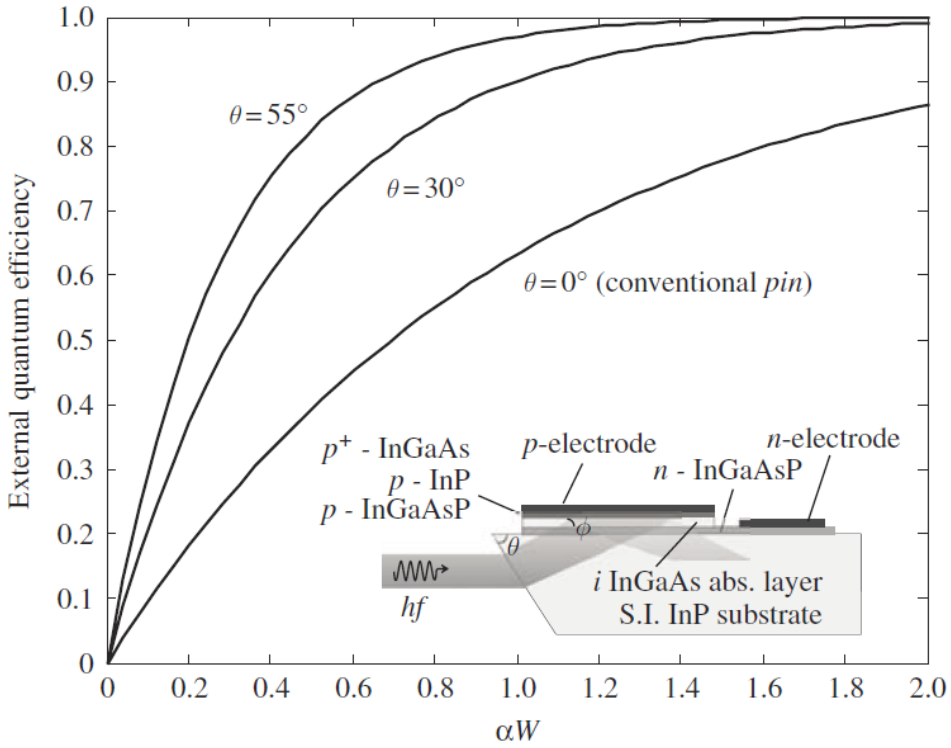
- Advanced pin photodiodes

Advanced pin PDs

Although conventional (vertical) *pin* photodiodes can today achieve bandwidths well in excess of 40 Gbps, such structures are close to the limit performances. Their main limitation, i.e., the efficiency–bandwidth trade-off, can be overcome by properly modifying the device design. From a physical standpoint, the efficiency–bandwidth trade-off originates from the fact that photons and photogenerated carriers run parallel. This implies that increasing the thickness of the absorption layer (and, therefore, the efficiency) automatically leads to an increase of the transit time. However, if the angle between the power flux and the collected photocarrier path is made larger than zero, the absorption region can be made wider while preserving a low transit time. Alternatively, if photons are made to cross the active region several times, absorption can be high even with a low value of the photocarrier transit time.

Refracting facet PDs

- RFPD Uses a refracting facet to increase the effective absorption thickness, while the transit time is unchanged.



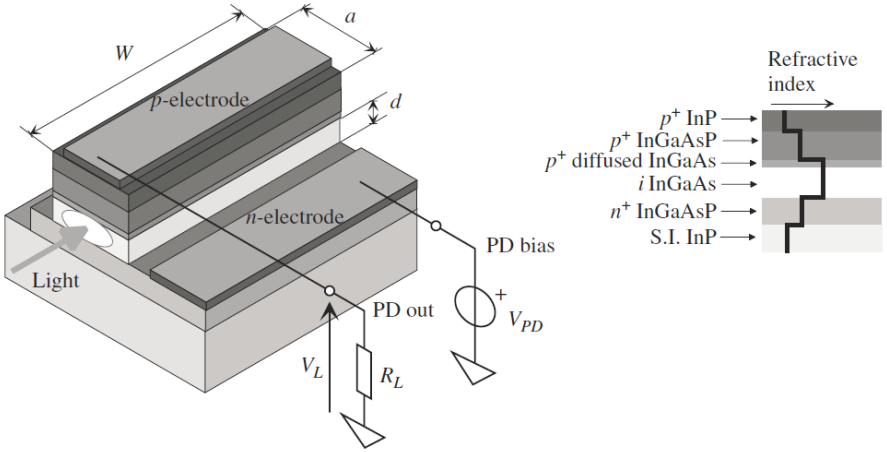
$$W_{\text{eff}} = \frac{2W}{\cos\theta}$$

Resonant cavity PDs

In *resonant cavity detectors*, light is made to cross the absorption layer several times by inserting such a layer in a vertical optical resonator confined by front and back Bragg reflectors (a grating obtained by stacking several semiconductor layers of low and high refractive index). This solution increases the efficiency but, since the cavity is resonant, the optical bandwidth is narrow.

Waveguide PDs

- In waveguide PDs, the photon flux and the carrier motion are orthogonal. Light is guided by an optical waveguide made of an intrinsic narrow-gap semiconductor layer, sandwiched between two highly doped widegap layers.
- The transit time is determined by d, the photon absorption is by the waveguide length W, and the capacitance is low due to the large width a of the waveguide.

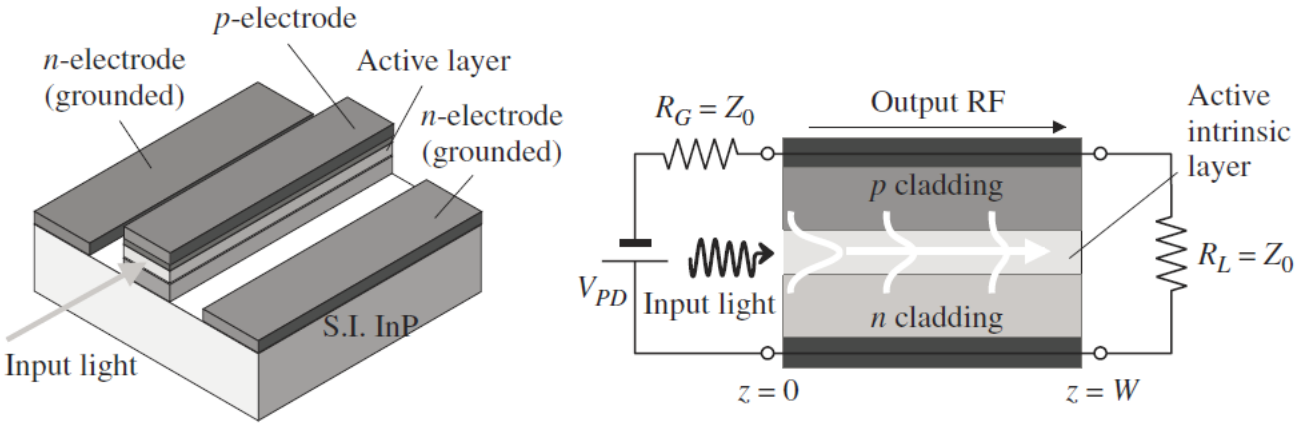


Schematic structure of an InGaAs waveguide photodiode (left) and details of the epitaxial structure (right) showing the guiding refractive index profile.

- The disadvantage is that the optical field is not completely confined by the narrowgap waveguide core, but partly extends into the widegap cladding, where absorption is negligible.

Traveling-wave PDs

- The high-speed operation of waveguide photodiodes is mainly limited by RC cutoff; further bandwidth improvements can be obtained by turning the RF electrodes into transmission line. This can be done by properly feeding the signal from one (input) end of the structure and by loading the other (output) end with a matched load; the detector becomes a nonconventional coplanar waveguide, continuously loaded by the pin reverse-biased capacitance. The resulting traveling wave or distributed photodiode is able to overcome the RC limitations of waveguide PDs.



Distributed waveguide photodiode structure (left) and equivalent circuit in terms of optical and RF coupled waveguides (right) showing the generator and load conditions.

Traveling-wave PDs

- A traveling-wave structure is defined when the electrical and optical waves co-propagate in the same direction. Thus, both interact over some distances as they propagate down the waveguide.

- The bandwidth of such structure depends strongly on the matching between the optical phase velocity and the electrical phase velocity.

- The characteristics that distinguish the traveling-wave PD from the waveguide PD are (1) the electrical waveguide is in parallel with the optical waveguide; (2) a matched electrical termination at the device output end (while the input termination can be open-circuit or matched).

Velocity-matched TW PDs

- In practical implementations, the ideal version of the traveling-wave photodetector is affected by heavy RF losses; moreover, obtaining velocity matching is critical.
- While the optical refractive index is close to the active material index, the microwave effective refractive index can be lower (due to quasi-TEM averaging between the air and the dielectric) or, in some structures, much larger (and also strongly frequency-dependent) because of the slow-wave effect.

$$\epsilon_{\text{eff}} \approx \frac{\epsilon_{RF} + 1}{2} < \epsilon_o$$

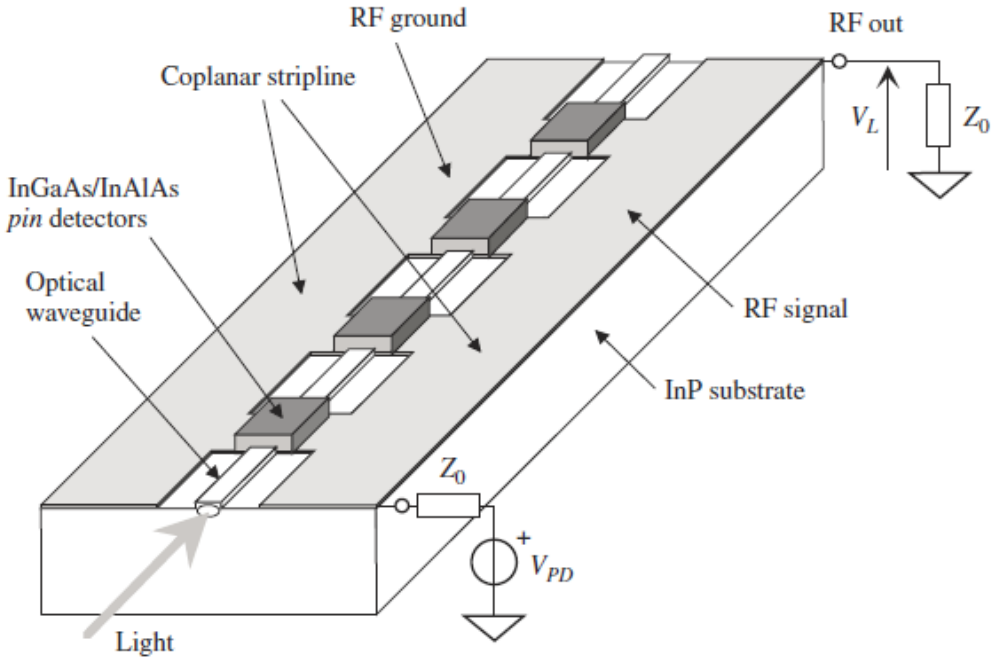
Velocity-matched TW PDs

- To decrease RF losses, probably the most important limitation to high-speed operation, periodically loaded distributed detectors have been proposed; the RF line is a quasi-TEM coplanar stripline periodically loaded with pin photodiodes. Owing to the wider RF conductors , the RF losses decrease dramatically and the effective length of the detector can be increased. However, over such a long structure, velocity matching becomes an issue.

- The effective permittivity for RF signal is smaller than for optical signal, so the former propagates faster.

Velocity-matched TW PDs

- By periodically loading the RF waveguide with PDs, the PD capacitance can increase the RF effective permittivity, and hence slow down the RF wave to match the velocity of the light.
- As opposed to TW PDs, veolocity-matched PD can achieve bandwidths well in excess of 100 GHz.

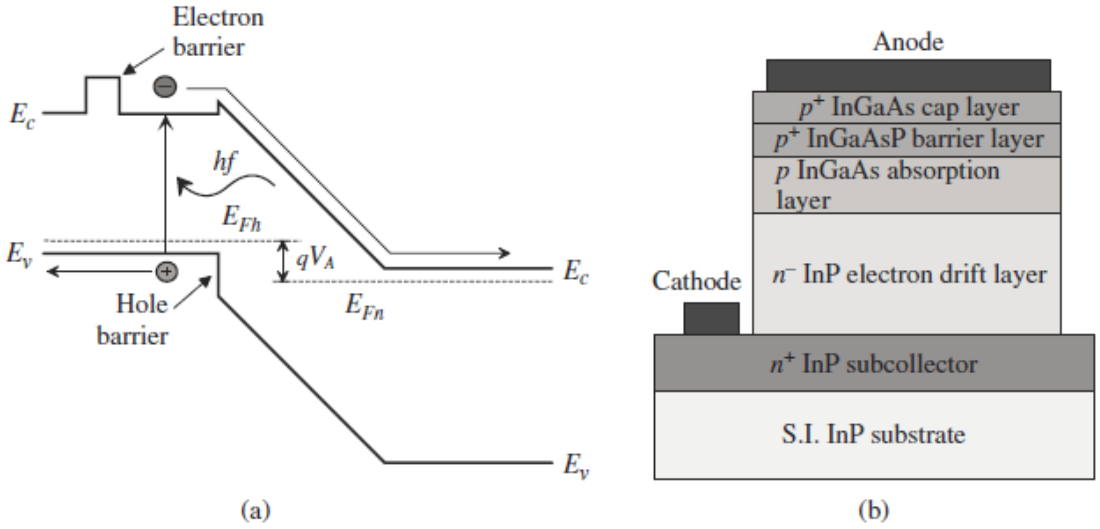


Distributed photodetector consisting of a RF waveguide periodically loaded with photodiodes.

Uni-traveling carrier PDs

Uni-traveling carrier photodiodes (UTC-PDs) are a high-speed evolution of conventional, vertically illuminated *pin* photodiodes. In the conventional *pin*, absorption takes place in the intrinsic layer, where electrons and holes are generated. The transit time is negatively affected by the lower velocity of holes vs. electrons; moreover, if the absorption layer thickness is reduced to improve the transit time, the efficiency and the *RC*-limited bandwidth decrease. UTC-PDs overcome such limitations in two ways. Firstly, photons are absorbed in a *p* layer where the field is low, but holes are collected by the *p* contact, while electrons quickly diffuse into an intrinsic or *n*− high-field drift region, where they undergo quasi-ballistic (i.e., not affected by scattering events) motion to the collector. Due to the very thin absorption layer, the transit time depends only on the electron drift through the intrinsic layer, while holes do not play a role; however, due to quasi-ballistic motion, the electron transit time is low. Secondly, the electron drift layer is thick enough to limit the diode capacitance, thus increasing the *RC*-limited bandwidth.

Uni-traveling carrier PDs



Simplified band diagram of uni-traveling carrier photodiode (UTC-PD) under reverse bias and backside illumination (a); layer structure of UTC-PD (b). Adapted from [44], Fig. 1.

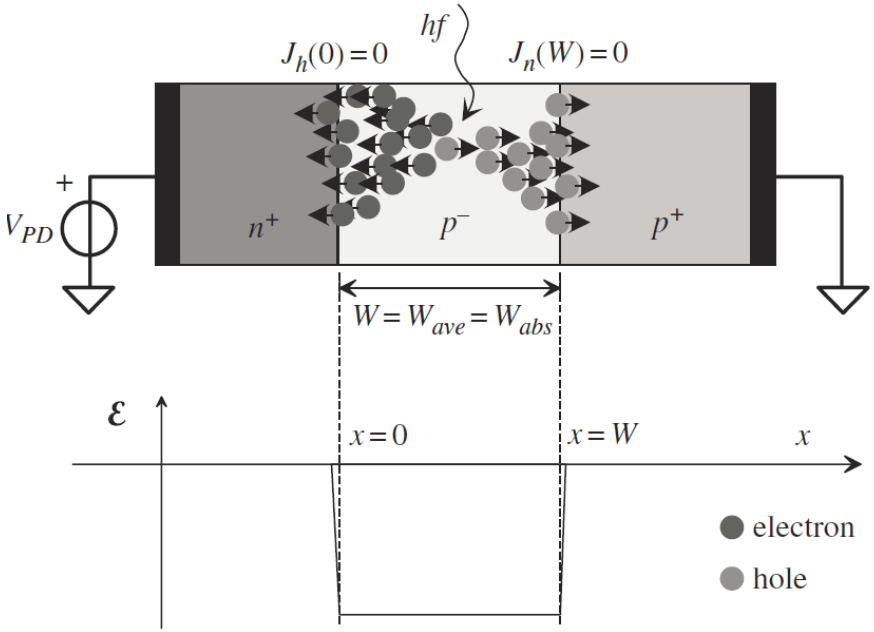
Note the InGaAsP electron barrier layer and the hole barrier introduced by the InGaAs–InP valence band discontinuity. Typical absorption layers are thin in order not to decrease the response speed due to diffusion in the neutral absorption layer, with reported widths of the order of 100–200 nm, leading to rather poor responsivities (around 0.15 mA/mW); backside illumination is exploited to improve the photon collection and avoid photons being absorbed by the electron barrier layer. UTC-PDs developed by NTT Photonics Labs have demonstrated record performances for 1.55µm detectors, with 3 dB bandwidth of 310 GHz. The estimated average electron velocity is of the order of 3×10^7 cm/s, thus demonstrating velocity overshoot effects and quasi-ballistic transport in the drift region. With respect to distributed and waveguide photodetectors, UTC-PD have a much simpler, vertical structure and comparable speed; typical responsivities are lower, however.

Lecture 28

- APD responsivity

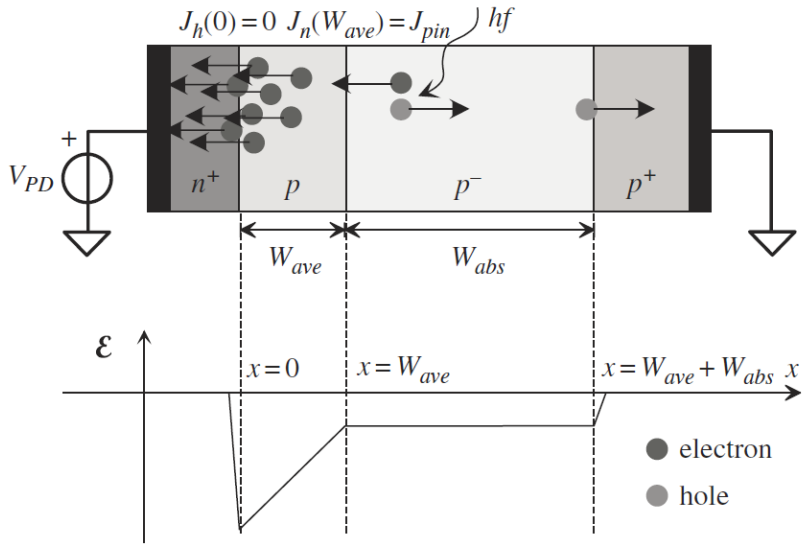
APDs

Avalanche photodiodes (APDs) exploit the avalanche multiplication of photogenerated carriers through *impact ionization (碰撞电离)* to amplify the detector current and improve the device sensitivity. A larger photocurrent (for the same illumination level) is obtained with respect to *pin* photodiodes, but also higher noise. In principle, the device includes two regions: the *generation region* (low to medium electric field) and the *multiplication region* (high field). The two regions can be physically the same, as in *conventional APDs*, or can be separated (SAM-APD: *separate absorption and multiplication APD*). In the APD the avalanche is triggered by electrons; the dual case (avalanche triggered by holes) is also possible;

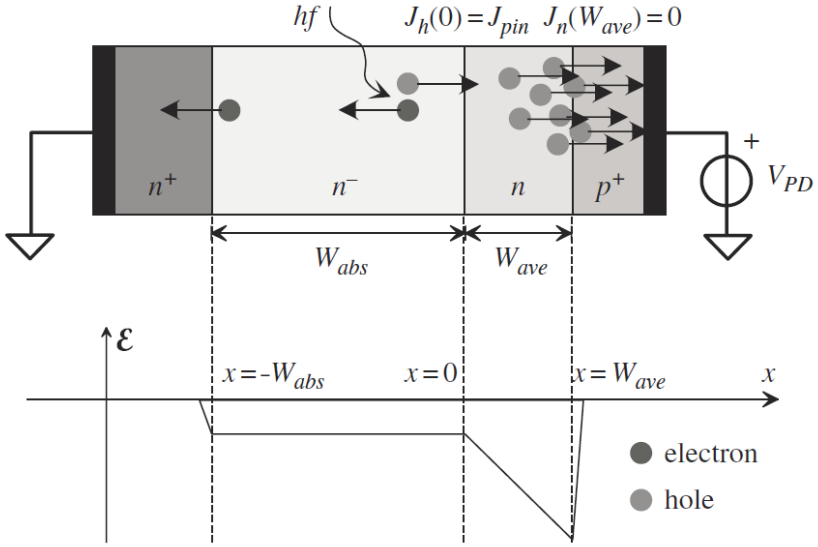


Conventional APD (above) and electric field profile (below).

APDs



SAM-APD (electron-triggered avalanche) with electric field profile.



SAM-APD (hole-triggered avalanche) with electric field profile.

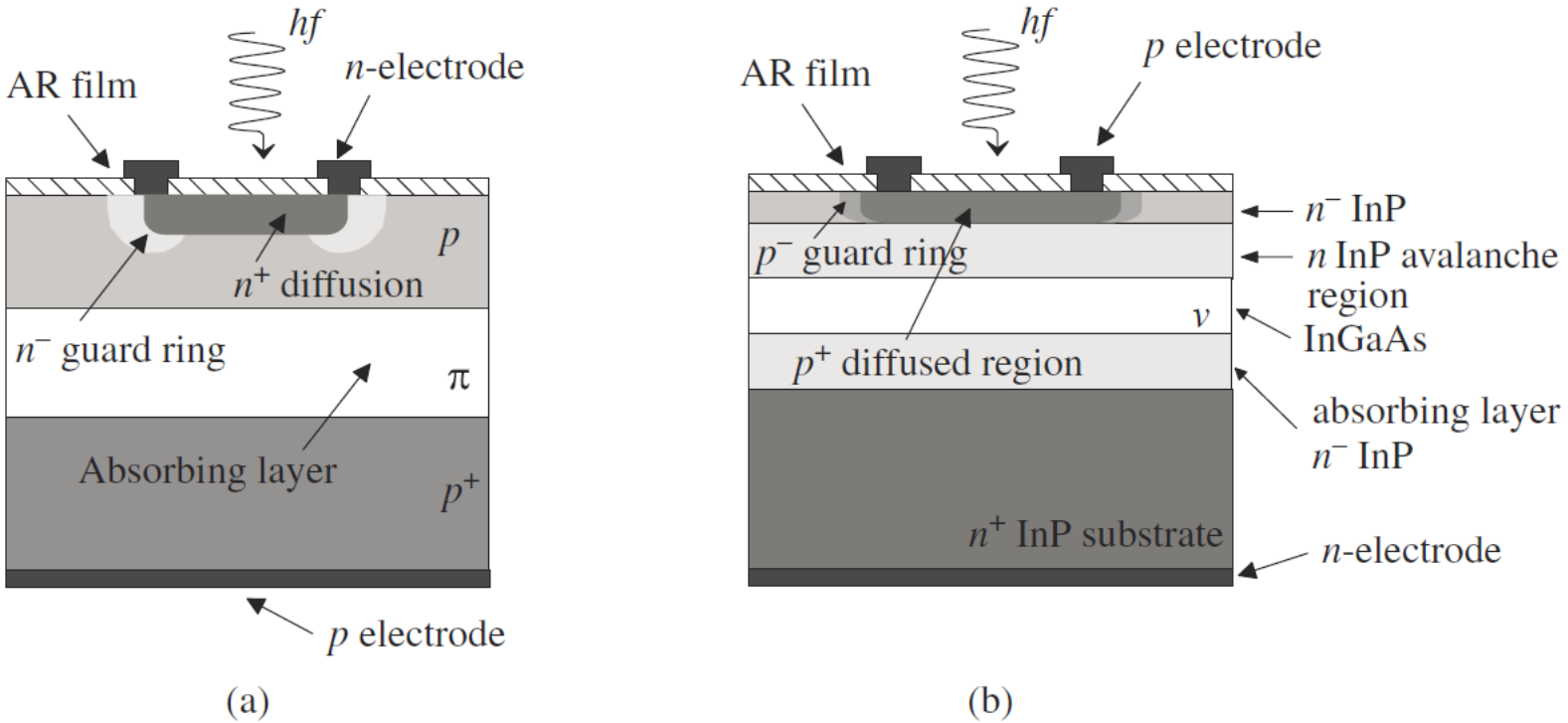
In the SAM-APD the absorption region is intrinsic, and multiplication takes place either in an *n*+ *p* junction at the device left-hand side, or in a *np*+ junction at the device right-hand side,. In the first case, the hole current at the left-hand side depletion edge is negligible, $J_h(0) = 0$, while the electron current at the beginning of the avalanche region is $J_n(W_{av}) = J_{pin}$, where J_{pin} is the *primary current* photogenerated by the absorption region alone. In the second case, the electron current at the right-hand side depletion edge is negligible, $J_n(W_{av}) = 0$, while the hole current at the beginning of the avalanche region is $J_n(0) = J_{pin}$, where J_{pin} is the primary current photogenerated by the absorption region alone。

APDs

Although the device analysis shows that the responsivity increases with increasing avalanche multiplication, which is, in turn, larger if the hole and electron ionization coefficients are similar (i.e., if both carriers contribute to the avalanche process), large values of the multiplication factor are inconvenient because they increase noise and decrease speed. Optimum noise and speed are achieved when the avalanche process is almost unipolar: this happens when the ionization coefficient of the *avalanche triggering carrier* is much larger than the ionization coefficient of the other carrier.
For some materials, this condition is met naturally; it can be artificially introduced through bandstructure engineering and the use of superlattices. From the technological standpoint, early APDs were Si- or Ge-based homojunction devices, with large operating voltages (e.g., in excess of 50 V). Today, InP-based long-wavelength APDs are available, and also the operating APD voltage has considerably decreased, thanks to device downsizing.

APDs

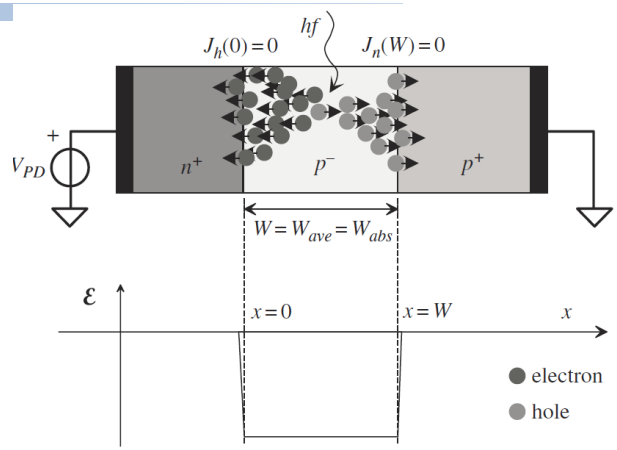
Examples of APD structures are shown in the following figures; note the guard ring introduced into the junction. The structure is typically vertical (top or bottom illuminated), although waveguide APDs also have been proposed.



(a) Si-based APD; (b) InGaAs long-wavelength APD.

APD responsivity

- For conventional APD, consider a DC condition, in which avalanche generation is triggers, in a region of length W and electric field E, by the injection of a hole current in x=0 and an electron current in x=W.
- Neglecting all generation and recombination terms, we have



Conventional APD (above) and electric field profile (below).

$$\frac{dJ_n}{dx} = -\alpha_h J_h - \alpha_n J_n$$

$$\frac{dJ_h}{dx} = \alpha_h J_h + \alpha_n J_n.$$

- The ionization coefficient is defined as the number of electron-hole pairs generated per unit distance per incident electron/hole.

- J_n decreases while J_h increases with increasing x.
- Carrier current is mainly drift current due to the large electric field, while diffusion current is negligible.

APD responsivity

- The total current density $J=J_n+J_h$ is constant vs. x, then

$$\frac{\mathrm{d}J_h}{\mathrm{d}x} - (\alpha_h - \alpha_n) J_h = \alpha_n J.$$

- Assume the boundary conditions are: hole incident current $J_h(0)$, and electron incident current $J_n(W)$. We obtain

$$J_h(x) = \mathrm{e}^{\int_0^x (\alpha_h - \alpha_n)\,\mathrm{d}\xi} \left[J_h(0) + J \int_0^x \alpha_n \mathrm{e}^{-\int_0^\xi (\alpha_h - \alpha_n)\,\mathrm{d}\eta}\,\mathrm{d}\xi \right]$$

- Note that the ionization coefficient alpha is generally functions of x in the presence of a nonuniform electric field.
- The total current density can be given at x=W,

$$J = J_n(W) + J_h(W) = J_n(W) + \mathrm{e}^{\int_0^W (\alpha_h - \alpha_n)\,\mathrm{d}\xi} \times \left[J_h(0) + J \int_0^W \alpha_n \mathrm{e}^{-\int_0^\xi (\alpha_h - \alpha_n)\,\mathrm{d}\eta}\,\mathrm{d}\xi \right]$$

APD responsivity

- Then, we obtain the current density as

$$J = \frac{e^{\int_0^W (\alpha_h - \alpha_n)\,d\xi} J_h(0) + J_n(W)}{1 - e^{\int_0^W (\alpha_h - \alpha_n)\,d\xi} \int_0^W \alpha_n e^{-\int_0^\xi (\alpha_h - \alpha_n)\,d\eta}\,d\xi} \equiv \frac{N}{D}$$
$$= M_h J_h(0) + M_n J_n(W),$$

Where M is the multiplication factor.

Thus, in the presence of a finite injected electron or hole current, the total current becomes larger, and can diverge due to the carrier avalanche generation when D→0.

$$D = e^{\int_0^W (\alpha_h - \alpha_n)\,d\xi} \left[1 - \int_0^W \alpha_h e^{-\int_0^\xi (\alpha_h - \alpha_n)\,d\eta}\,d\xi\right]$$

APD responsivity

- So the avalanche breakdown condition, corresponding to infinite current is,

$$\int_0^W \alpha_h e^{-\int_0^x (\alpha_h - \alpha_n)\, d\xi}\, dx \equiv I_{h,\text{ion}} = 1$$

Where $I_{h,ion}$ is the hole ionization integral. Similarly, the electron ionization integral is

$$I_{n,\text{ion}} = \int_0^W \alpha_n e^{\int_x^W (\alpha_h - \alpha_n)\, d\xi}\, dx$$

- Correspondingly, the multiplication factors are

$$M_h = \frac{1}{1 - \int_0^W \alpha_h e^{-\int_0^\xi (\alpha_h - \alpha_n)\, d\eta}\, d\xi} = \frac{1}{1 - I_{h,\text{ion}}}$$

$$M_n = \frac{1}{1 - \int_0^W \alpha_n e^{\int_\xi^W (\alpha_h - \alpha_n)\, d\eta}\, d\xi} = \frac{1}{1 - I_{n,\text{ion}}}.$$

APD responsivity

- If one of the two injected current vanishes, we say the avalanche multiplication is triggered by holes ($J_n(W)=0$) or by electrons ($J_h(0)=0$).
- In case the ionization coefficients for holes and electrons are the same, then the ionization integrals become

$$I_{n,\text{ion}} = \int_0^W \alpha_n \, \mathrm{d}x = I_{h,\text{ion}} = \int_0^W \alpha_h \, \mathrm{d}x$$

- So the avalanche condition is simplified to

$$\int_0^W \alpha_n \, \mathrm{d}x = \int_0^W \alpha_h \, \mathrm{d}x = 1$$

Condition $\alpha_n \approx \alpha_h$ is approximately verified in some compound semiconductors, while in Si the ionization rates are significantly different. In many cases we can approximately assume that the avalanche ionization coefficients are different but *constant over the avalanche region*.

APD responsivity

- Then, the ionization integrals simplify as

$$I_{n,\text{ion}} = \frac{\alpha_n}{\alpha_h - \alpha_n}\left[\mathrm{e}^{(\alpha_h-\alpha_n)W} - 1\right]$$
$$I_{h,\text{ion}} = \frac{\alpha_h}{\alpha_n - \alpha_h}\left[\mathrm{e}^{(\alpha_n-\alpha_h)W} - 1\right];$$

and the multiplication factors become

$$M_n = \frac{1}{1 - I_{n,\text{ion}}} = \frac{\alpha_h - \alpha_n}{\alpha_h - \alpha_n \mathrm{e}^{(\alpha_h-\alpha_n)W}} = \frac{(1 - k_{hn})\,\mathrm{e}^{\alpha_n(1-k_{hn})W}}{1 - k_{hn}\mathrm{e}^{\alpha_n(1-k_{hn})W}}$$
$$M_h = \frac{1}{1 - I_{h,\text{ion}}} = \frac{\alpha_n - \alpha_h}{\alpha_n - \alpha_h \mathrm{e}^{(\alpha_n-\alpha_h)W}} = \frac{k_{hn} - 1}{k_{hn}\mathrm{e}^{\alpha_n(1-k_{hn})W} - 1},$$

where the ionization coefficient ratio

$$k_{hn} = \frac{\alpha_h}{\alpha_n}.$$

has been introduced.

APD responsivity

- For electron-triggered SAM APD, $J_h(0)=0$, so

$$J = M_n J_{pin} \qquad J_{pin} = J_n(W)$$

- Thus, the APD responsivity is

$$\Re^n_{\text{SAM-APD}} = M_n \Re_{pin} = M_n \frac{q\eta_Q}{hf}(1-R)\left(1-\mathrm{e}^{-\alpha W_{abs}}\right)$$

$$\Re^h_{\text{SAM-APD}} = M_h \Re_{pin} = M_h \frac{q\eta_Q}{hf}(1-R)\left(1-\mathrm{e}^{-\alpha W_{abs}}\right)$$

Where R_{pin} is the responsivity of the equivalent pin.

Therefore, the responsivity of APD can be larger than the ideal maximum value.

APD responsivity

- Introducing the optical generation, the current density becomes

$$\frac{\mathrm{d}J_n}{\mathrm{d}x} = -\alpha_h J_h - \alpha_n J_n - qG_o$$
$$\frac{\mathrm{d}J_h}{\mathrm{d}x} = \alpha_h J_h + \alpha_n J_n + qG_o,$$

- Accounting for the constant total current,

$$\frac{\mathrm{d}J_h}{\mathrm{d}x} - (\alpha_h - \alpha_n) J_h = \alpha_n J + qG_o.$$

- Using the boundary conditions, $J_h(0)=0$ and $J_n(W)=0$, we have

$$J_h(x) = \mathrm{e}^{\int_0^x (\alpha_h - \alpha_n)\,\mathrm{d}\xi} \times \left[J_h(0) + J \int_0^x \alpha_n \mathrm{e}^{-\int_0^\xi (\alpha_h - \alpha_n)\,\mathrm{d}\eta}\,\mathrm{d}\xi + qG_o \int_0^x \mathrm{e}^{-\int_0^\xi (\alpha_h - \alpha_n)\,\mathrm{d}\eta}\,\mathrm{d}\xi \right]$$

APD responsivity

- The total current density,

$$J = J_h(W) = e^{\int_0^W (\alpha_h - \alpha_n)\,dx} \times \left[J \int_0^W \alpha_n e^{-\int_0^x (\alpha_h - \alpha_n)\,d\xi}\,dx + qG_o \int_0^W e^{-\int_0^x (\alpha_h - \alpha_n)\,d\xi}\,dx \right]$$

$$= qG_o \frac{e^{\int_0^W (\alpha_h - \alpha_n)\,d\xi} \int_0^W e^{-\int_0^\xi (\alpha_h - \alpha_n)\,d\eta}\,d\xi}{1 - e^{\int_0^W (\alpha_h - \alpha_n)\,d\xi} \int_0^W \alpha_n e^{-\int_0^\xi (\alpha_h - \alpha_n)\,d\eta}\,d\xi} = qG_o W M_o = M_o J_{pin}$$

- The photocurrent multiplication factor is

$$M_o = \frac{1}{W} \frac{\int_0^W e^{-\int_0^\xi (\alpha_h - \alpha_n)\,d\eta}\,d\xi}{1 - \int_0^W \alpha_h e^{-\int_0^\xi (\alpha_h - \alpha_n)\,d\eta}\,d\xi}$$

APD responsivity

- For the same ionization coefficients, the multiplication factor becomes,

$$M_o = \frac{1}{1 - \int_0^W \alpha_h \, \mathrm{d}x}$$

- Assume the ionization coefficients are constant over x, then

$$M_o = \frac{1}{W} \frac{\mathrm{e}^{(\alpha_n - \alpha_h)W} - 1}{\alpha_n - \alpha_h \mathrm{e}^{(\alpha_n - \alpha_h)W}} = \frac{1}{W\alpha_n} \frac{\mathrm{e}^{\alpha_n(1-k_{hn})W} - 1}{1 - k_{hn}\mathrm{e}^{\alpha_n(1-k_{hn})W}}$$

- Thus, the responsivity of APD is

$$\Re_{\mathrm{APD}} = M_o \Re_{pin}$$

APD responsivity

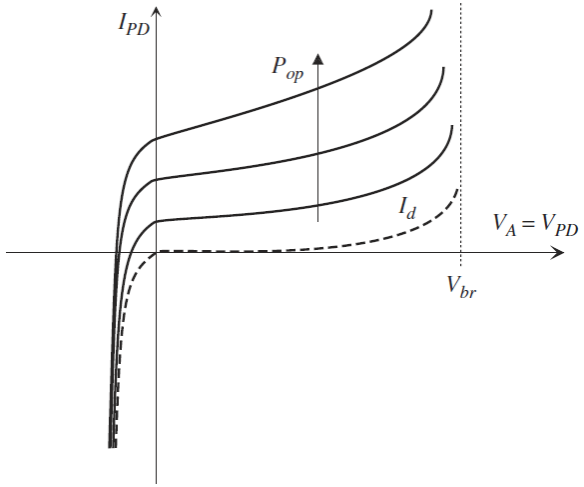
In both the conventional APD and the SAM-APD the responsivity is enhanced due to the effect of avalanche multiplication. From a technological standpoint, the SAMAPD allows for separate optimization of the avalanche region and of the absorption region, while in the conventional APD the high-field avalanche region width should be $W > L_{\alpha}$; as a result, the APD voltage may be very large, unless compound semiconductors are used. The separate optimization of the avalanche region thickness and of the absorption region length can be carried out in principle by exploiting waveguide photodetector structures

APD responsivity

- The APD multiplication factor (M_n, M_h, or M_0) increases with increasing voltage. The empirical relation is

$$M \approx \left[1 - \left(\frac{V_A - R_s I}{V_{br}}\right)\right]^{-n}$$

where R_s is the parasitic series resistance, V_{br} is the diode breakdown voltage, and V_A is the applied reverse voltage. Practical values for M typically are < 10 to decrease noise, lower the bias voltage, and improve the device reliability.



V–I characteristics of an APD.

APD responsivity

- The multiplication factor of SAM APDs is usually larger than the conventional APD.
- For k_{hn}=0,

$$M_n = e^{\alpha_n W} > M_o = \frac{e^{\alpha_n W} - 1}{\alpha_n W}$$

- For k_{hn}=1,

$$M_o = \frac{1}{1 - \alpha_n W} = M_n$$

- From a physical standpoint, this is because in the SAM-APD the whole primary photocurrent (from the most effective avalanching carrier, the electrons) is injected entirely at the edges of the avalanche region, while in the conventional APD the photocurrent injection takes place only gradually within the avalanche region (also acting as the absorption region). To optimize the multiplication factor in a SAM-APD the avalanche triggering (injected) carrier should be in principle the one with the larger ionization coefficient. If, on the other hand, k_{hn} = 1, we have $M_o = M_h = M_n$, i.e., all structures are equivalent concerning the multiplication factor.

Lecture 29

- APD noise

APD noise

- In pin diode, the noise can be modeled as shot noise, and the power spectrum of the noise current (neglecting dark current) is given by

$$S_{i_n,pin} = 2q I_L = 2q\Re_{pin} P_{in}$$

- The APD current is given by the equivalent pin current multiplied by the multiplication coefficient, and the shot noise spectrum seems to be

$$i_{L,\mathrm{APD}} = M i_{L,pin}$$

$$S_{i_n,\mathrm{APD}} = M^2 S_{i_n,pin}$$

- However, the avalanche carriers generates additional noise with respect to the simple multiplication of the shot noise.

APD noise

- The actual noise of APD is

$$S_{i_n,\mathrm{APD}} = M^2 F S_{i_n,pin} = 2qM^2 F \Re_{pin} P_{in} = 2qMF\Re_{\mathrm{APD}} P_{in},$$

Where F is the excess noise factor, which is more than 1.

- The excess noise is strongly dependent on the ratio of the ionization coefficients.
- For conventional APD, the short-circuit noise is

$$S_{\delta i}^{o} = 2q I_L M_o F_o(M_o)$$

$$M_o = \frac{1}{W\alpha_n} \frac{\mathrm{e}^{\alpha_n(1-k_{hn})W} - 1}{1 - k_{hn}\mathrm{e}^{\alpha_n(1-k_{hn})W}}$$

$$F_o = \frac{(1+\alpha_h W M_o)(1+\alpha_n W M_o)}{M_o}$$

APD noise

- For electron-triggered SAM APD,

$$S_{\delta i}^{\mathrm{SAM}_n} = 2q\,I_L\,M_n\,F_n(M_n)$$

$$M_n = \frac{(1-k_{hn})\,\mathrm{e}^{\alpha_n(1-k_{hn})W_{av}}}{1-k_{hn}\mathrm{e}^{\alpha_n(1-k_{hn})W_{av}}}$$

$$F_n = k_{hn}M_n + (1-k_{hn})\left(2-M_n^{-1}\right)$$

$$= M_n\left[1-(1-k_{hn})\left(1-M_n^{-1}\right)^2\right]$$

- For hole-triggered SAM APD,

$$S_{\delta i}^{\mathrm{SAM}_h} = 2q\,I_L\,M_h\,F_h(M_h)$$

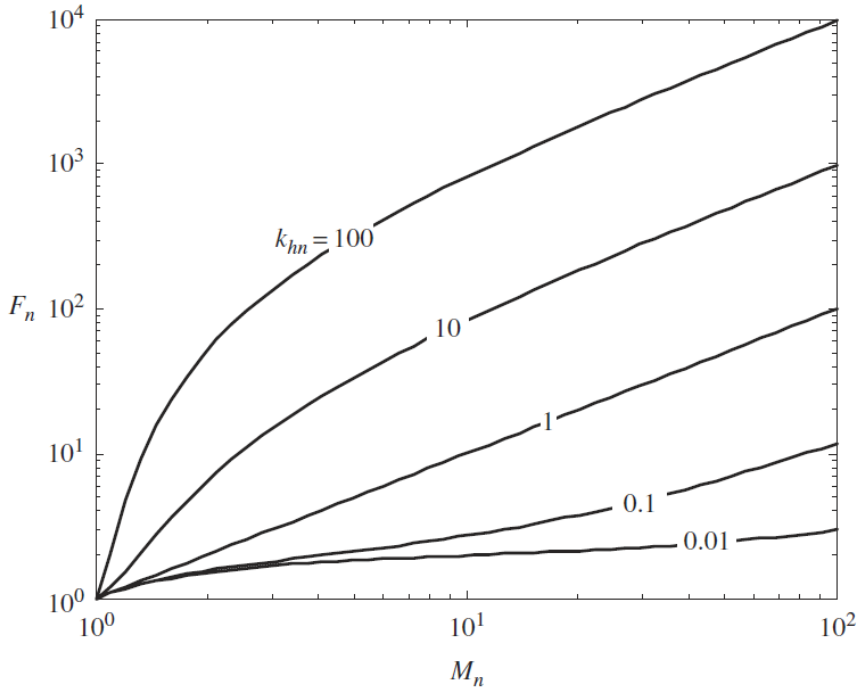
$$M_h = \frac{k_{hn}-1}{k_{hn}\mathrm{e}^{\alpha_n(1-k_{hn})W_{av}}-1}$$

$$F_h = k_{nh}M_h + (1-k_{nh})\left(2-M_h^{-1}\right)$$

$$= M_h\left[1-(1-k_{nh})\left(1-M_h^{-1}\right)^2\right]$$

APD noise

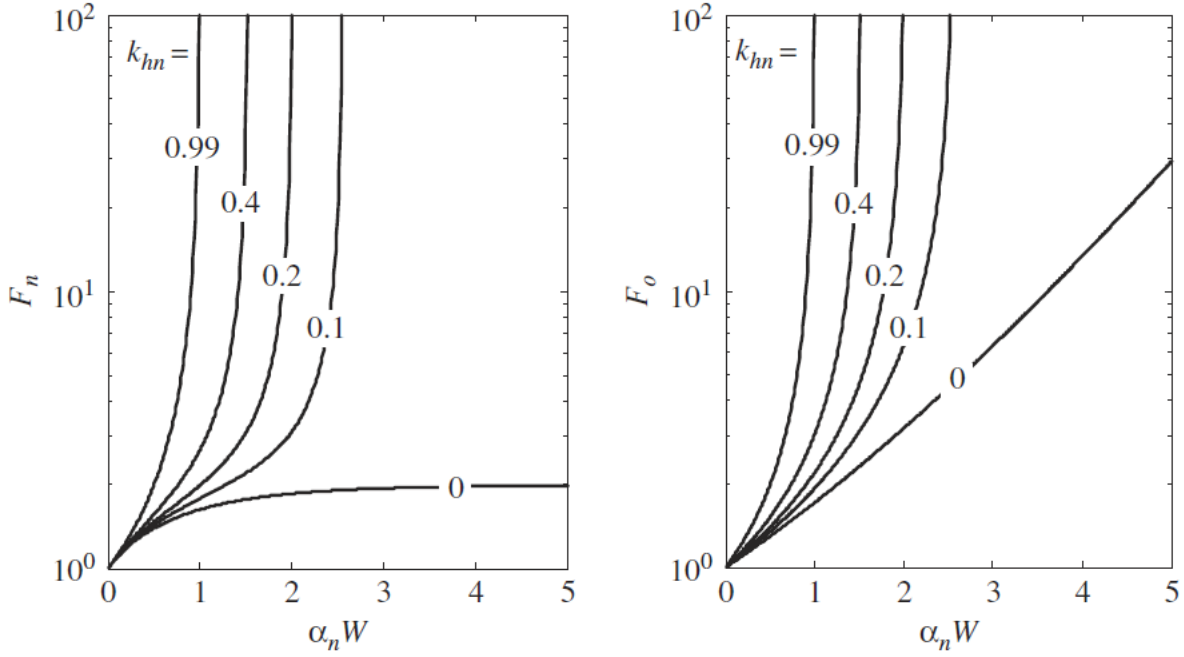
- The excess noise factor increases if the electrons and holes have similar ionization coefficients, and also increases with increasing multiplication factor.



Excess noise factor F_n as a function of the electron multiplication coefficient M_n; the parameter is k_{hn}.

APD noise

- For low k_{hn}, the SAM APD is able to reach lower excess noise factor than the conventional APD.



Excess noise factor for a SAM-APD (electron triggered) and conventional APD as a function of $\alpha_n W$.

APD noise

- SAM-APDs therefore, beside allowing to optimize the width of the absorption region so that $W_{abs} > L_{\alpha}$ without requiring unrealistically large driving voltages, are also potentially less noisy than conventional APDs; moreover, in SAM-APDs the avalanche region can be implemented with a larger gap material.

- Independent of the material, low-voltage operation leads to low multiplication factor, but also low excess noise.

Lecture 30

- APD frequency response

APD frequency response

- The frequency response of APD is related to several intrinsic mechanism delaying the carriers before being collected.

- The electron transit time in the absorption region:

$$\tau_{tr} = W_{abs}/v_{n,\mathrm{sat}}.$$

- The avalanche buildup time (or avalanche delay), which can be expressed as

$$\tau_A \approx \frac{r\,(k_{hn})\,M_n W_{av} k_{hn}}{v_e},$$

where $r\,(k_{hn})$ is a slowly varying function of k_{hn} such as $r\,(1) = 1/6$, $r(0) = 1$, and v_e is an effective velocity defined as

$$v_e = \frac{v_{n,\mathrm{sat}} v_{h,\mathrm{sat}}}{v_{n,\mathrm{sat}} + v_{h,\mathrm{sat}}}.$$

APD frequency response

- The transit time of avalanche generated holes:

$$\tau_{tr,Ah} = \frac{W_{abs} + W_{av}}{v_{h,\mathrm{sat}}}$$

- The total delay of the SAM APD is

$$\tau = \tau_{tr} + \tau_A + \tau_{tr,Ah} = \frac{W_{abs}}{v_{n,\mathrm{sat}}} + \frac{M_n W_{av} k_{hn}}{v_{n,\mathrm{sat}}} + \frac{W_{abs} + W_{av}}{v_{h,\mathrm{sat}}}$$

- The intrinsic cutoff frequency is approximately,

$$f_T \approx \frac{1}{2.2\tau}$$

- The total delay is minimized if $k_{hn} \ll 1$, that is, the ionization coefficient of the electron is much larger than the hole in the electron-trigger case.

APD frequency response

Due to the additional delay mechanisms, the APD is typically slower than the *pin*; moreover, the device response is faster when only the avalanche triggering carriers (e.g., electrons) ionize. In fact, if both electrons and holes have the same ionization probability, an electron pulse injected into the high-field region generates holes which, instead of simply being collected in the *p*-side, in turn generate secondary electrons, whose delay before collection is of course larger than the delay of the electrons directly generated by the initial pulse. Secondary electrons generate new holes, and so on; if the multiplication factor is finite, the process finally dies out with a slow tail. As a result, a larger number of carriers is collected with respect to the $k_{hn} = 0$ case (and, indeed the multiplication factor is larger), but such carriers will be collected over a longer time interval leading to an output current pulse with a large spread in time. This finally amounts to a slower response.

Example 4.4

Good luck for your future on Optoelectronics!