Purpose of a Receiver

• The receiver fulfils the function of optoelectronic conversion of an input optical signal into an output electrical signal (data stream). The purpose is to recover the data transmitted over the optical fibre communication channel.

• In fibre communication systems the optical detector is invariably a semiconductor p-i-n photodiode or avalanche photodiode. In free-space systems (sometimes used/proposed to be used to connect buildings) other detectors such as the vacuum photomultiplier may be used.

Desirable Properties of the Optical Detector

- Wavelength response Matched to the optical wavelength used for transmission.
- High optoelectronic conversion efficiency Maximises output signal to noise ratio efficiency & low noise.
- Fast response rate To maximise the system bit.
- High reliability Long mean time to failure (*MTTF*).
- Small size & low cost Minimise system cost and integrate with electronic circuitry.

The Photoelectric Effect

- Optical absorption of a photon within a semiconductor material occurs when the photon has sufficient energy to generate an electron-hole pair, i.e. E_p = hf ≥ E_g.
- If an external electric field is applied across the semiconductor, the photogenerated carriers will be swept away causing a photocurrent, I_{ρ} , proportional to the optical power, P_{in} , to flow.

$$I_p = R(\lambda)P_{in}$$

where

 $R(\lambda) AW^{-1}$ is the optical wavelength dependent detector responsivity.

The Photoelectric Effect

- As with optical sources, this is a quantum process and we can define a quantum efficiency, η

$$\eta = \frac{\text{Carrier generation rate}}{\text{Photon incidence rate}} = \frac{I_p / q}{P_{in} / hf} = R \frac{hf}{q}$$

then

$$R = \frac{\eta q}{hf} \approx \eta \lambda (\mu m) / 1.24$$

The Photoelectric Effect

Hence, the responsivity is directly proportional to the optical wavelength until photon energy, $E_p = hf < E_g$ when it falls rapidly to zero, as shown in figure.



The Photoelectric Effect

 As for the optical fibre, we can define an absorption coefficient, a, to calculate the amount of power absorbed over a distance, D, of the material. The transmitted power is then given by

$$P(D) = P_{in}e^{-\alpha D}$$

and the absorbed power is

$$P_{abs} = P_{in} \left[1 - e^{-\alpha D} \right]$$

The efficiency is thus given by

$$\eta = \frac{P_{abs}}{P_{in}} = \left[1 - e^{-\alpha D}\right]$$

If $\alpha=0$, $\eta=0$ but if $\alpha D >>1$ then $\eta->1$

The Photoelectric Effect

• Figure below shows the absorption coefficients for several semiconductors as a function of wavelength. For each, there is a cut-off value of wavelength, λ_c , for which the absorption coefficient approaches zero. This is the wavelength value which satisfies:



The Photoelectric Effect

- These materials are therefore only useful as optical detectors for wavelengths below the cut-off, i.e. $\lambda < \lambda_c$.
- For a good quantum efficiency (≈1) the absorption coefficient, α, should be at least 10³ cm⁻¹ and thus the absorption depth, D, of the material must be ≥ 10µm.
- When no light is absorbed there is still a low but non-zero conductivity and hence, a so-called *dark current* flows at all times which is unrelated to the optical power and thus reduces the signal to noise ratio.

Photovoltaic Detectors

• A reverse biased p-n junction can be used as an optical detector.



- The dark current is opposed by the depletion region and built-in electric field. However, it is still non-zero but is reduced from mA to nA.
- Any photo-electrically generated carrier pairs, however will be rapidly separated by the large internal field and a photocurrent can flow. This is the basis of all photodiode optical detectors.

Photovoltaic Detectors

• Figure below shows a typical reverse biased p-n diode and illustrates the small amount of light absorbed within the depletion region.



The electric field across the depletion region is much greater than that across the p or n regions as their conductivity is higher. Carriers generated outside the depletion region diffuse apart slowly, whereas carriers generated within the depletion region drift apart much faster under the influence of the electric field. Consequently, it would be beneficial to minimise the width of the p and n regions and to maximise the width of the depletion region in the depletion region and the speed of response are maximised.

Photovoltaic Detectors

The separate drift and diffusion times give rise to the pulse distortion illustrated in figure. For a typical p-n junction total width, W ≈10µm, a drift velocity of 10⁵ ms⁻¹ gives a transit time of ≈0.1ns and thus a maximum bit rate ≈ 1Gbs⁻¹. The limiting factor is pulse broadening due to the slow diffusion speed ≈ 10³ ms⁻¹.



p-i-n Photodiode

 The diffusion contribution can be reduced by reducing the p and n region widths and increasing the depletion width by inserting an additional intrinsic semiconductor layer between the p and n layers forming a *p-i-n photodiode*, as shown in figure.



- The intrinsic section has a high resistance hence, most of the bias potential is dropped across it.
- To design the optimum width one must make a compromise between increasing W to increase responsivity as η increases, but also increasing the transit time, *t*, and therefore reducing the maximum bit rate.

p-i-n Photodiode

- Semiconductors such as Si and Ge (indirect bandgap) have $\eta \rightarrow 1$ for $W = 20 50 \mu m$, but $\tau > 0.2 ns$.
- Direct bandgap semiconductors such as InGaAs can achieve $\tau \approx 30$ -50 ps and bandwidth of 3 - 5 GHz for $W \approx 3$ - 5 µm. Bandwidths up to 70 GHz have been achieved but this requires $W \approx 1$ µm and thus the responsivity is low. InGaAs has $\lambda_c = 1.65$ and is thus ideal for use in optical communications receivers at 1.3 and 1.55µm.



Avalanche Photodiode (APD)

 Leakage across the diode junction gives rise to a non-zero dark current. Hence, the minimum detectable optical signal and signal to noise ratio (SNR) are limited. For

$$SNR > 1$$
 $P_{in} > I_{dark}/R$

• For a p-i-n photodiode the maximum responsivity is

$$R_{max} = q/hf$$
 when $(\eta=1)$

• This can be increased using an *avalanche photodiode*.

Avalanche Photodiode (APD)

- APDs have internal current gain via the mechanism of impact ionisation.
- If photo-generated carriers experience a sufficiently large electric field, they are accelerated to high velocity. The kinetic energy can reach a level such that collisions with atoms ionises them, that is sufficient energy is transferred to promote another valence electron to the conduction band leaving behind another hole.
- Carrier pair generations and impact ionisations occur in a cascade manner and a large photocurrent gain can be achieved.

Avalanche Photodiode (APD)

• The excess carrier pair generation is governed by impact ionisation parameters, α_e and α_h , which depend on both the material and the applied field. Figure shows parameters for several semiconductors.



• The fields appear very large however, across the width of a typical *APD* a field of 10MVm⁻¹ represents a voltage of around 100V.

Avalanche Photodiode (APD)

• Figure below shows a schematic and actual design for an *APD*. The only difference from a standard p-i-n photodiode is an extra thin high voltage gain layer of p type material between the intrinsic and n regions.



Avalanche Photodiode (APD)

- Avalanche gain is an intrinsically noisy process and thus *M* must be the mean *APD* gain.
- The bandwidth of an *APD* is lower than for a p-i-n photodiode. The electrical bandwidth is found to be proportional to *1/M*, hence the design is again a compromise between high responsivity and bandwidth. Gains of M≈100 are easily obtainable.
- Silicon APDs can be used at $\lambda = 0.85 \mu m$ with low noise and bit rates up to 1Gb/s.
- InGaAs APDs can be used at 1.3 and 1.55µm and have relatively high noise and low bandwidth. This can be improved using heterostructures and with an InP gain region and adding a further InGaAsP layer of intermediate bandgap. These separate absorption, grading and multiplication layer (SAGM) APDs have a gain >12 and a gain-bandwidth product up to 70GHz

Receiver Components



Front End

- This performs the initial optoelectronic conversion of the optical bit stream into an electrical bit stream. It comprises the photodiode (pin or *APD*) and a preamplifier (combined in a pin-FET) coupled to the optical fibre channel via a pigtail or coupling optics.
- The two most common front end designs are shown in figure.



Front End

In a *high impedance* front end, the voltage input to the pre-amplifier is increased by the load resistance, R_L. Therefore, if R_L is large this design has high sensitivity, but low bandwidth, ∆f as

$$\Delta f = \frac{1}{2\pi R_L C_T}$$

where

- C_{τ} is the total input capacitance.

• The *transimpedance* design uses the load resistor to provide negative feedback to an inverting amplifier. Hence, the input resistance is reduced by a factor equal to the amplifier gain, *G* and the bandwidth thus increased by *G*. This is the most commonly used configuration.

The Linear Channel

- This consists of the main high gain amplifier plus a low pass filter.
- The automatic gain control (*AGC*) circuit gives a constant average output voltage irrespective of the average input optical power (as long as it is above a certain minimum value).
- The total receiver noise is proportional to the receiver bandwidth. Consequently, the insertion of a low pass filter places a limit on the bandwidth and limits noise.
- Note that if *∆f* < *B* i.e. the pulse width is greater than the bit slot, then intersymbol interference (*ISI*) will occur.

Data Recovery

- This section has two purposes
 - 1. Clock recovery circuits isolate a signal at a frequency equal to the bit rate which is used to synchronise the bit decision process.
 - 2. A threshold detector is used to determine whether a particular bit is a logic '0' or logic '1'.
- The optimum threshold levels and bit sampling time can be determined from an *eye pattern* for the receiver as shown in figure. This superimposes '0' and '1' bits degraded by noise. How 'open' the eye appears indicates the degree of differentiation between the two bit states.



Eye Diagram

See these two web pages for more information on interpretation:

http://www.scientificarts.com/logo/logos.html

http://www.complextoreal.com/chapters/eye.pdf

Receiver Noise

- Noise Sources
 - The total time dependent photocurrent can be written as:

$$I(t) = I_P(t) + I_{dark} + i_s(t) + i_T(t)$$

where

• $I_P(t)$ is the signal photocurrent \propto Incident optical power, P_{in}

- • I_{dark} is the constant dark current ∞ Stray light and thermal carrier recombinations but not P_{in}
- •*i*_s(*t*) is the **shot noise** current contribution due to random carrier generation Mean square shot noise current = $\bar{i}_s(t)^2 = 2q(I_P + I_{dark})\Delta f$
- • $i_T(t)$ is the thermal noise contribution due to the random thermal motion of electrons within resistors (also called **Johnson or Nyquist noise**)

Mean square thermal noise current = $\bar{i}_T(t)^2 = 4k_BT(1/R_L)F\Delta f$

where *F* is the *noise figure* of the pre-amplifier.

Receiver Noise

- Noise Equivalent Power (*NEP*)
 - The NEP is defined as the minimum optical power per unit bandwidth required to produce a signal to noise ratio, SNR =1. The inverse of this quantity, known as the detectivity is sometimes quoted by manufacturers.

Receiver Noise

• Signal to Noise Ratios – pin Receivers

$$SNR = \frac{Average \ signal \ power}{Average \ noise \ power} = \frac{I_P^2}{\sigma_{total}^2}$$

in most practical cases, thermal noise is dominant: $\sigma_{\rm T} >> \sigma_{\rm S}$, and from $I_{\rm P} = RP_{\rm in}$

$$SNR = \frac{R^2 P_{in}^2 R_L}{4k_B TF \Delta f}$$

$$NEP = \frac{P_{in(SNR=1)}}{\sqrt{\Delta f}} = \sqrt{\frac{4k_B TF}{R^2 R_L}}$$

Typically, *NEP*≈1-10 pW/√Hz

Receiver Noise

- Signal to Noise Ratios pin Receivers
 - It is possible to approach the lower shot noise limit e.g. by cooling the detector

Shot noise limited SNR=
$$\frac{(RP_{in})^2}{2q(I_P + I_{dark})\Delta f} \approx \frac{RP_{in}}{2q\Delta f}$$

Ideally, the number of photons for a '0' bit would be zero and the number for a '1' bit would be equivalent to the minimum detectable optical power, $P_{in}(min)$. If N_p = minimum number of photons received during one bit slot to indicate a '1' bit then

$$P_{in}(\min) = N_P h f B$$

For λ =1.55µm, *B*=1Gb/s and *SNR*=20dB: $N_p \approx 100$ photons, $P_{in}(min) \approx 13$ nW.

Receiver Noise

APD Receiver Noise

When compared with pin photodiodes it would appear that a larger SNR can be achieved for the same P_{in} .

$$I_P = MRP_{in} = R_{APD}P_{in}$$

 $SNR \propto P_{in}^2$ for thermal noise, so it seems that SNR is improved by M^2 . Actually, it is not improved because the shot noise power is also increased to

where $\sigma_s^2 = 2qM^2(RP_{in} + I_{dark})\Delta fF_E$

 F_E is the excess noise factor due to random generation of secondary carriers. If one carrier dominates the $F_E(max)=2$.

Receiver Noise

• APD Receiver Noise

APD thermal SNR is improved by a factor of M^2

$$SNR = \frac{R_L R^2 M^2}{4k_B TF} \frac{P_{in}^2}{\Delta f}$$

However, shot noise limited SNR is reduced by a factor of F_E

$$SNR = \frac{RP_{in}}{2qF_E\Delta f}$$

Receiver Noise

Receiver Performance

One receiver can be said to be more sensitive than another if it achieves the same performance for a lower incident optical power. Digital techniques are mostly used within optical communications so that absolute optical intensities which may vary with channel loss are unimportant. When dealing with digital receivers we define the sensitivity in terms of the minimum average received optical power (P_{rec}) to achieve a given bit error rate (*BER*) performance, usually *BER* \approx 10⁻⁹.

$$P_{rec} = \overline{N}_P h f B$$

 $\underline{N_P}$ is the average number of photons per bit (includes different numbers of photons per bit and the 0:1 bit ratio) and *B* is the bit rate.

Receiver Noise

Receiver Performance

The sensitivity will be degraded by :

 Reducing the bit extinction ratio i.e. having a non-zero number of photons for a '0' bit.

- Intensity noise such as the transmitter RIN.

- Timing jitter (>10% of bit period) due to noise in the clock recovery gives fluctuations in the sampling time and thus additional noise.

- Thermal and shot noise.

Real receiver sensitivities are \approx 25dB above theoretical limits for p-i-n diodes and \approx 20dB above for APDs.

Modulation and Demodulation

• If we consider the electric field of a general sinusoidal optical signal we can write:

$$E_{c}(t) = A_{c}(t)\cos(\omega_{c}t + \phi_{c})$$

Hence, we may modulate $E_c(t)$ by varying either its instantaneous amplitude, $A_c(t)$, frequency, $\omega_c(t)$ or phase, $f_c(t)$. However, as we have seen, direct detection optical detectors respond only to the optical power. Hence, the photocurrent related to the above signal is given by:

$$I_{c}(t) = RP_{c}(t) = R\left\{A_{c}(t)^{2}\cos^{2}(\omega_{c}t + \phi_{c})\right\} = R\frac{A_{c}(t)^{2}}{2}\left\{1 + \cos(2\omega_{c}t)\right\} = R\frac{A_{c}(t)^{2}}{2}$$

The cos term is removed by filtering or time averaging [2()> = 1/2]

Modulation and Demodulation

 The detector will only respond to amplitude (intensity) modulation. Any frequency or phase information is lost. This leaves the choice between analogue (i.e. multilevel) amplitude modulation and digital modulation. Analogue modulation would necessitate the measurement of the absolute power level to recover the modulation amplitude information. This would require very stable and linear transmitters and fibre channels. Non-linearity would distort the modulation envelope and lead to errors. Likewise, any variation in the loss at any point within the system would corrupt the data being transmitted.

Modulation and Demodulation

• Consequently, digital techniques are most commonly used. The direct detection constraints are significantly relaxed as now we need only differentiate between two optical power levels, that for a logic '0' bit and that for a logic '1' bit. The most commonly used digital techniques are pulse width modulation (*PWM*) and straight binary or binary coded decimal (*BCD*) for both return to zero and non-return to zero formats.

In *PWM* the length of each pulse encodes the signal amplitude hence, this techniques is sensitive to distortion by dispersion induced pulse broadening.



Modulation and Demodulation

BCD encodes the information in the bit sequence and is thus the simplest to implement reliably. Thus the *BCD* scheme is the most frequently used.



Coherent Detection

 The received optical signal light is coherently mixed with light from a second, narrow linewidth laser (local oscillator). If the input signal has an amplitude given by

$$E_c(t) = A_c e^{-j(\omega_c t + \phi_c)}$$

and the local oscillator amplitude is given by

$$E_{lo}(t) = A_{lo}e^{-j(\omega_{lo}t + \phi_{lo})}$$

then the optical power incident on the detector from the resultant mixed beam is easily shown to be,

$$P(t) \propto \left[E_c(t) + E_{lo}(t)\right]^2 = \left\{P_c + P_{lo} + 2\sqrt{P_c P_{lo}}\cos(\omega_{IF} + \phi_c - \phi_{lo})\right\}$$

where

 $-P_{c} = K A_{c}^{2}, \quad P_{lo} = K A_{lo}^{2}$ -K is a coefficient of proportionality $-\omega_{lF} = (\omega_{c} - \omega_{lo}) \text{ is the intermediate frequency}$

Coherent Detection

The photocurrent is thus,

$$I(t) = RP(t) = \left\{ R[P_c + P_{lo}] + 2R\sqrt{P_c P_{lo}} \cos(\omega_{IF} + \phi_c - \phi_{lo}) \right\}$$

and contains terms involving the signal amplitude, frequency and phase hence, unlike with direct detection, for coherent detection any of these quantities may be modulated to transfer information.

Homodyne Coherent Detection

• If we make the intermediate frequency, $\omega_{IF} = 0$ i.e. ($\omega_c = \omega_{Io}$) then,

$$I(t) = RP(t) = \left\{ R[P_c + P_{lo}] + 2R\sqrt{P_c P_{lo}} \cos(\phi_c - \phi_{lo}) \right\}$$

Typically, $P_{lo} >> P_c$ and we can arrange that the phase of the local oscillator is locked to that of the signal using an optical phase locked loop hence, $f_c - f_{lo} = constant$. The information carrying part of the homodyne signal is thus,

$$I(t) = 2R\sqrt{P_c(t)P_{lo}}$$

Therefore, intensity modulated signals only can be recovered and relative to *IM/DD*, the signal power has increased by a factor of

$$4P_{lo} / P_c >> 1$$

However, the requirements that $(\omega_c = \omega_{lo})$ and $(\phi_c - \phi_{lo}) = constant$ are difficult to maintain and necessitate very expensive laser sources.

Heterodyne Coherent Detection

• If we allow the intermediate frequency, $\omega_{IF} \neq 0$ but have $(f_c - f_{lo}) \approx 0.1 - 5$ GHz, then we have an intermediate frequency in the microwave range and standard microwave components can be used in the detector. Assuming, $P_{lo} >> P_c$ and bandpass filtering to remove DC terms leaves

$$I(t) = 2R\sqrt{P_c(t)P_{lo}}\cos(\omega_{IF}(t)t + \phi_c(t) - \phi_{lo})$$

Hence, intensity, frequency or phase modulation can be used.

Therefore, relative to *IM/DD*, the average signal power has increased by a factor of

$$2P_{lo} / P_c >> 1$$
 Remember, $<\cos^2()> = 1/2$

This is only half the 'gain' of the homodyne scheme (this is known as the 3dB heterodyne penalty) but benefits from simpler, less expensive components.

SNR for Coherent Detection

• Increasing the local oscillator power such that

$$P_{lo} >> \frac{\sigma_{T^2}}{2qR\Delta f} >> \frac{I_{dark}}{R}$$

ensures that shot limited detection is achievable and hence,

Heterodyne Coherent Detection $SNR = (R P_c)/(q \Delta f) = 2 \times IM/DD SNR$

Homodyne Coherent Detection $SNR = (2 R P_c)/(q \Delta f) = 4 \times IM/DD SNR$

Coherent detection therefore offers enhanced sensitivity and a greater variety of modulation schemes.

SNR for Coherent Detection

• Figure shows how amplitude, phase and frequency modulation can be implemented digitally.



- Amplitude shift keying (ASK): Modulate A_c ; ω_c and f_c constant
- Phase shift keying (PSK): Modulate f_c ; A_c and ω_c constant
- Frequency shift keying (FSK): Modulate ω_c ; A_c and f_c constant

SNR for Coherent Detection

 As we discussed earlier, direct modulation of the laser injection current leads to amplitude-frequency/phase cross modulation. Hence, to implement a coherent detection scheme we require an external modulator to implement the ASK, PSK or FSK. Insertion of the external modulator between the laser and the fibre, or spliced in-line with the fibre, introduces an additional loss component (≈1dB) but this is more than compensated by the coherent detections >3dB gain.

External Modulators

 The most common optoelectronic modulators are produced using integrated optic technology. Figure shows such a device configured as an amplitude modulator.



The light from the input fibre enters a channel waveguide which splits and then recombines. The difference in the two optical paths is small and so complete interference occurs at the output. The optical phase difference between the two paths can be controlled electro-optically and thus the output amplitude is directly related to the voltage applied to the modulation electrodes. Modulation up to 20GHz can be achieved this way.

External Modulators

Other *AM* devices are based on electro-optic Pockel's cells, acousto-optic Bragg cells or multiple quantum well electro-absorption.

PM and *FM* integrated optic devices can also be produced using the principle of electro-refraction or using acousto-optic Bragg shifting. Direct *FM* modulation of the laser is used but requires the correction of the cross modulated, non-uniform intensity during demodulation.