Optical Communication

Notes Part D : OPTICAL COMMUNICATION SYSTEMS

10. Single Links

We have seen in Part C how to determine the SNR for an optical link. We can now use this calculation to determine the limits on the main operating parameters of the link, i.e. the bandwidth and length, and the relation between them (Fig. 10.1). Dispersion can, of course, also limit these quantities, but for single mode fibre in particular, SNR will generally be the main limitation up to very high bit rates, and we can combine the effects of dispersion with SNR by including a *dispersion penalty* in dB in our *power budget* calculations.

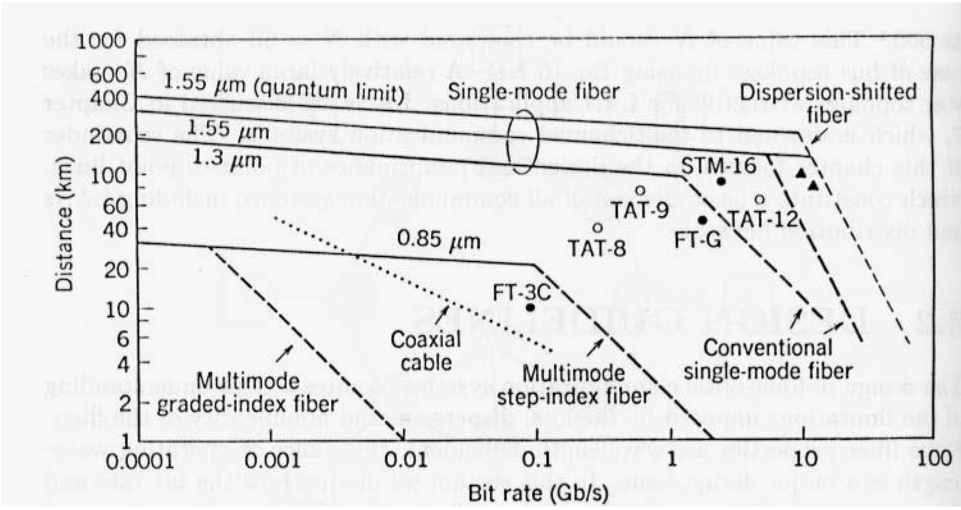


Fig. 10.1 Maximum cable length vs. bit rate for various optical links (with copper coax included for comparison). Solid lines represent power limits, dashed lines dispersion limits. From Agrawal.

In a power budget calculation, we consider the difference between transmitted power and minimum received power (this is our initial power "budget"), and from it, subtract the various losses. For example, if a laser source launches 0 dBm into a fibre, and the receiver requires –42 dBm for the required bit-rate, then we have 42 dB of loss which can be assigned to cable attenuation, connector and splice losses, system safety margin (includes allowance for ageing) and any additional factors such as dispersion penalty. If we take, in this example, 10 dB for losses except cable loss, and if the cable has 3 dB/km at the operating wavelength, then the maximum length without repeaters is just over 10 km, i.e. (42dB-10dB)/(3dB/km).

The required receiver SNR, for a given bandwidth, is determined for digital (amplitude modulated) systems by the acceptable bit error rate (BER), defined as the probability of an error in the interpretation of each bit. We define the average photocurrent for bit values 1 and 0 as I_1 and I_0 respectively, with these having variance σ_1 and σ_0 respectively. The relation between signal variance and noise level depends on the statistics of the noise sources. If we choose our decision level, i.e. the threshold for defining a received value as 1 or 0, optimally, then our BER will be as in Fig 10.2, as a function of the parameter Q:

$$Q = \frac{I_1 - I_0}{\sigma_1 + \sigma_0}$$

Note that for on-off keying ($I_0 = 0$), if $\sigma_1 \cong \sigma_0$ (reasonable for thermal noise), then $Q \cong SNR/2$. Thus for an acceptable BER of 10^{-9}, we need SNR of about 12.

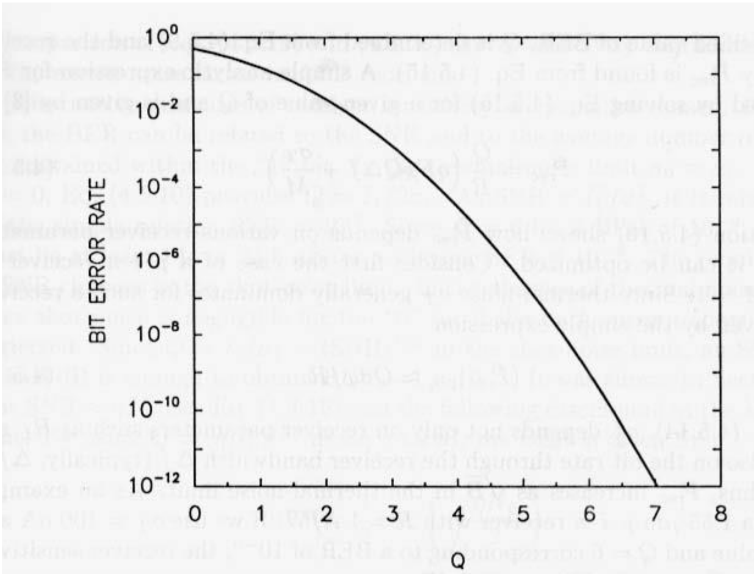


Fig. 10.2 Bit error rate vs. Q. From Agrawal.

Dispersion, and the resulting pulse broadening, has two deleterious effects. Firstly, it can cause inter-symbol interference because of the spread of signal power into adjacent symbol slots. In practice, pulse shaping is used to limit such effects. However, an unavoidable consequence of pulse spreading is the loss of signal power at the point (in time) of measurement, which we can calculate as a loss in SNR.

Consider a Gaussian pulse of transmitted rms width σ_o (in time), in the case where the rms spectral width σ_λ is dominated by the source, and the spectrum is also Gaussian. Assume that, to limit inter-symbol interference, the relation between the pulse width and slot size is $4B\sigma = 1$, with B the bit rate, i.e. the pulse fits (effectively) completely within the slot. Now we can obtain the pulse spreading for a cable of length L and dispersion parameter D from the equations given in Part 5, i.e.:

$$\sigma_g/\sigma_o = [1 + (DL\sigma_\lambda/\sigma_o)^2]^{1/2}$$

The dispersion penalty is just $10\log_{10}$ of this quantity, and is plotted in Fig. 10.3.

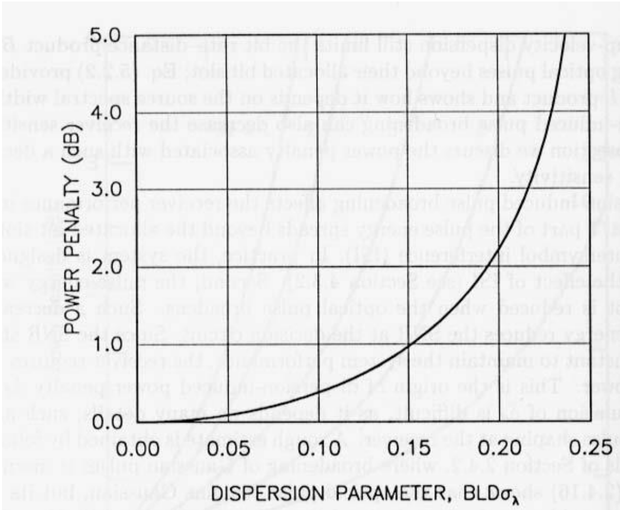


Fig. 10.3 Dispersion penalty vs. bit rate-length product. From Agrawal.

It is not always reasonable to assume a Gaussian spectrum. For example, Fabry-Perot lasers typically have a spectrum consisting of a number of distinct lines corresponding to longitudinal modes, and this causes what is termed mode partition noise. In DFB lasers, in which there is only one spectral line, there remains another factor, which is the dependence of output wavelength on drive current, which is termed chirp. This is complex to analyse, but a reasonable model can be derived assuming rectangular pulses, with the chirp confined to the leading and trailing edges of the pulse. In this model the penalty in dB, due to the chirped part of the spectrum moving out of the symbol slot, is given by (in dB):

$$\delta_c = -10\log_{10}(1 - 4BLD\Delta\lambda_c)$$

with $\Delta\lambda_c$ the spectral shift resulting from chirp. A more realistic model includes the effect of the chirp duration t_c, and is plotted in Fig. 10.4.

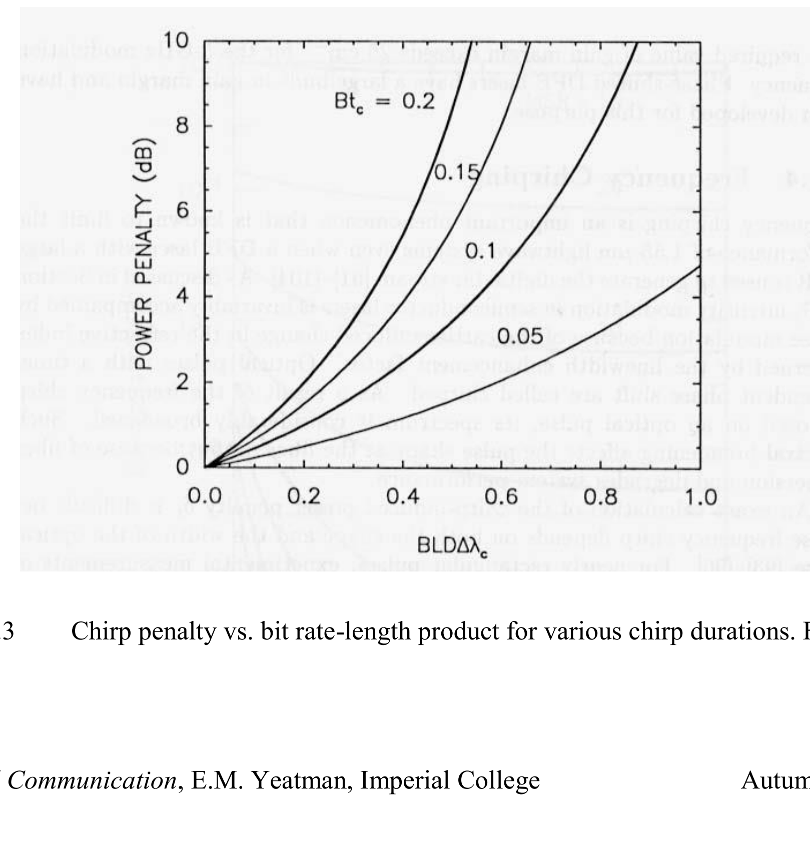


Fig. 10.3 Chirp penalty vs. bit rate-length product for various chirp durations. From Agrawal.

11. Coherent Communications

Coherent receivers are ones in which the phase of the carrier is used in the detection process (even if the demodulated signal is not effectively a measure of the phase). Such systems, besides allowing the use of phase modulation, also have considerable advantages, even with amplitude modulation, in terms of noise rejection, and most radio frequency receivers are in effect coherent. The main difficulty in optics is that most optical transmitters tend to have very poor coherence, i.e.. a wide intrinsic spectrum and a lot of phase noise. However, improvements in laser technology are increasingly making coherent techniques practical.

The coherent transmitter will need to have a very narrow line-width laser which is externally controlled to minimise frequency drift. Typically its output will be modulated in an external modulator rather than directly through the laser drive current, as drive current usually effects both amplitude and optical frequency. The coherent receiver, as in RF, always functions by mixing the incoming signal with a local oscillator - in the optical system, the local oscillator must be a laser. An optical coupler works as a coherent adder, since its output signal has an electric field amplitude given by the sum of the two input fields. If the phase of one of the signals varies, the two signals will interfere constructively or destructively. The mixer must be nonlinear in order to produce the sum and difference frequencies, and luckily the photodiode is an ideal square-law device, giving an output current proportional to the optical field squared (i.e. the intensity).

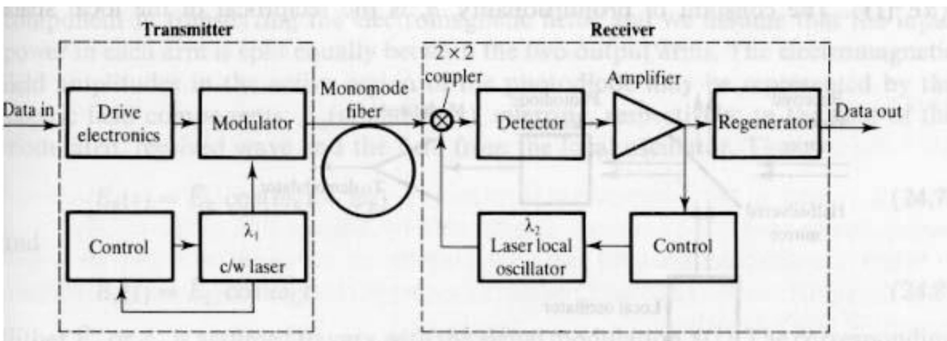


Figure 11.1 A schematic block diagram of a coherent system, (from Gowar, p. 547).

The figure above shows the generic arrangement of transmitter and receiver. Let us now consider what happens at the output of the coupler. We know that the photodiode gives an output current which is the responsivity times the received optical flux (or power):

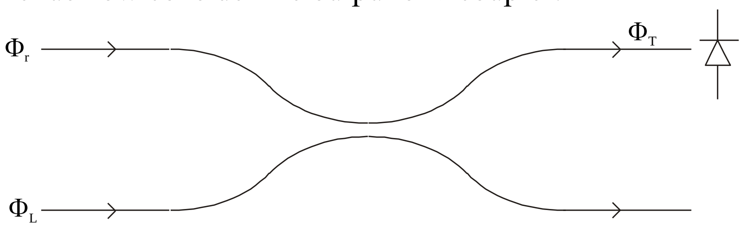
$$I = R\Phi$$

The optical flux is proportional to the square of the optical electric field, the receiver area A and the reciprocal of the impedance Z:

$$\Phi = \frac{\hat{E}^2}{2Z} \quad \text{where} \quad Z = \left(\frac{\mu_r\mu_o}{\varepsilon_r\varepsilon_o}\right)^{1/2}$$

and $\hat{E}$ is the optical field amplitude.

Let us now consider the output of a coupler:



where Φ_r and Φ_L are the received and local signals respectively. If the received signal is phase modulated, then:

$$E_r(t) = \hat{E}_r \cos(\omega_r t + \phi_r)$$
$$E_L(t) = \hat{E}_L \cos(\omega_L t)$$

The flux at the detector is now:

$$\Phi_T(t) = \frac{1}{2}\ \frac{A}{Z}\ \langle [E_r(t) + E_L(t)]^2 \rangle$$

where < > means a time average, and the extra ½ factor is due to the splitting of power between the two output ports.

$$\begin{aligned}[E_r(t) + E_L(t)]^2 &= \tfrac{1}{2}\hat{E}_r^2\ (1 + \cos 2(\omega_r t + \phi_r)) \\ &+ \tfrac{1}{2}\hat{E}_L^2\ (1 + \cos 2\omega_L t) \\ &+ \tfrac{1}{2}\hat{E}_r\hat{E}_L\ (\cos(\omega_r t + \omega_L t + \phi_r) + \cos(\omega_r t - \omega_L t + \phi_r))\end{aligned}$$

Since the photodiode detects the time average and cannot respond to optical frequencies, the $2\omega_r$, $2\omega_L$ and $\omega_r + \omega_L$ components vanish and we get:

$$\Phi_T(t) = \frac{1}{4}\ \frac{A}{Z}\ (\hat{E}_r^2 + \hat{E}_L^2 + \hat{E}_r\hat{E}_L \cos(\omega_r t - \omega_L t + \phi_r))$$

Of course the source and local oscillator wavelengths (and thus frequencies) must be close enough so that $(\omega_r - \omega_L)$ is within the detector frequency response. The first two terms are simply a constant DC level and the third contains the signal. Since the local intensity will be much higher than the received intensity, the signal has been boosted from $\hat{E}_r^2$ to $\hat{E}_r\hat{E}_L$, which can easily be enough to overcome amplifier and thermal noise.

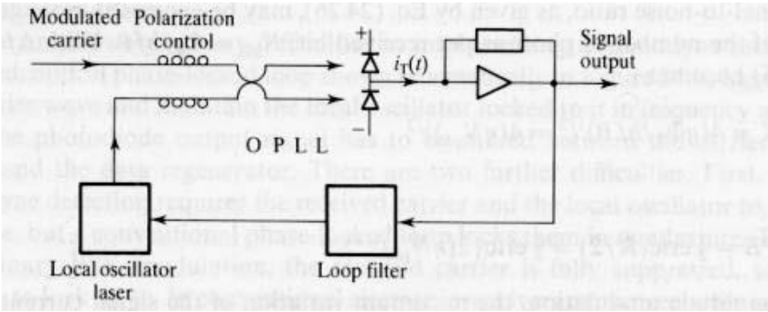


Figure 11.2 A balanced coherent receiver (from Gowar, p. 549).

In a balanced receiver as above, the DC contributions cancel and the signal current doubles. However, the shot noise will be mainly due to Φ_L, and it will be different between the two detectors and won't cancel. The noise signal will effectively be determined by:

$$\left(\Phi_{sh}^*\right)^2 = 2 \cdot 2eI = 2eR(\Phi_r + \Phi_L) \cong 2eR\Phi_L$$

This will dominate if it is greater than the combined amplifier/thermal noise. The signal current will be:

$$i(t) = 2R(\Phi_r \Phi_L)^{1/2} \cos \phi_r$$

Here we have assumed homodyne detection, i.e. $\omega_r = \omega_L$. This gives the best noise performance but requires locking the local oscillator to the signal with an optical phase-locked loop. The signal to noise ratio is then:

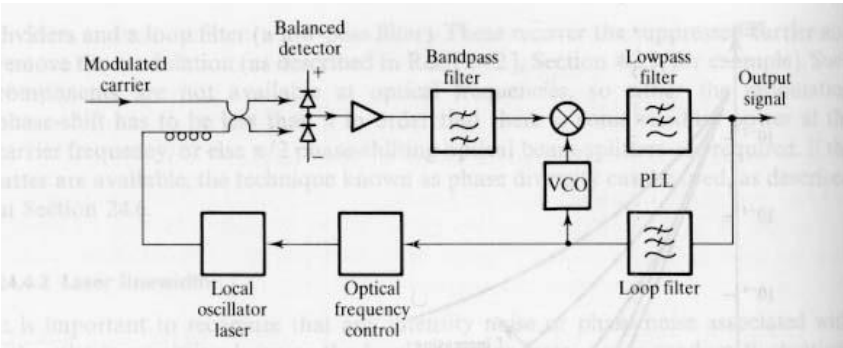
$$\text{SNR} \cong \frac{4R(\Phi_r \Phi_L)^{1/2}}{(2eR\Phi_L)^{1/2}} B_A^{1/2} = 4\left(\frac{R\Phi_r}{2eB_A}\right)^{1/2}$$

where B_A is the amplifier bandwidth. Here we have assumed a PSK signal with $\cos \phi_r = \pm 1$, so that the signal variation is 2i(t) as given above. Now, in the Nyquist limit the amplifier bandwidth is half the bitrate $B_A = B/2$. If we take the detector quantum efficiency to be perfect, $\eta = 1$, then $R = e/(h\upsilon)$ where $h\upsilon$ is the energy per photon. If $\dot{N}$ is the number of photons per second, then $\Phi = \dot{N}(h\upsilon)$. Combining all this, we obtain:

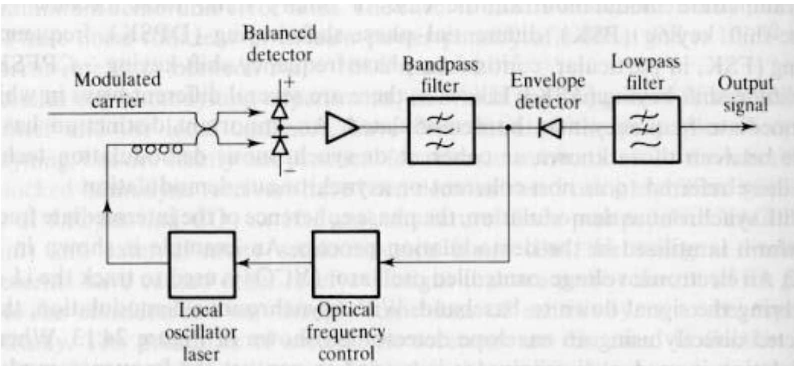
$$\text{SNR} = 4\left(\frac{\dot{N}}{B}\right)^{1/2}$$

$\dot{N}/B$ is the average number of photons per bit. A typical optical SNR requirement is 12, which gives a BER of 10^{-9} as we have seen. Using this, we obtain: $\dot{N}/B = (12/4)^2 = 9$ photons per bit. This represents an extraordinary sensitivity.

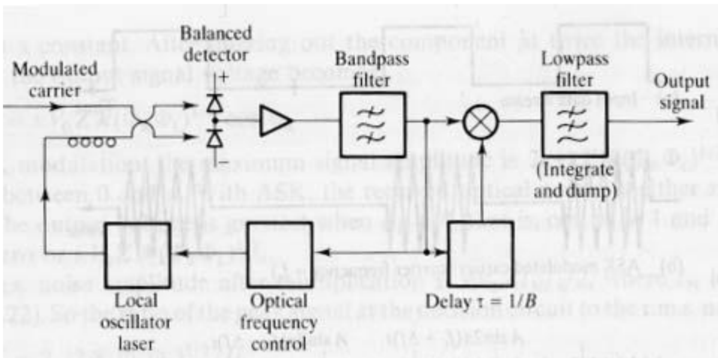
The heterodyne receiver mixes the incoming signal with a local oscillator of different frequency, so that the detector output has the signal information at an intermediate frequency $\omega_{IF} = \omega_L - \omega_r$. As in a normal RF receiver, a second detection stage is necessary to shift down to the baseband. This can be done asynchronously, i.e. with no second stage mixer, simply by envelope detection (for amplitude modulation). A synchronous detector has a local oscillator at the intermediate frequency. Self-synchronous receivers mix the IF signal with a copy of itself delayed by one bit. These are illustrated below:



a)



b)



c)

Fig. 11.3 Heterodyne receivers: a) with synchronous demodulation; b) with asynchronous (envelope) demodulation; c) DPSK with delay-loop demodulation; (from Gowar, pp. 558-559).

More recently, phase modulated (PSK) systems are being investigated which do not require local oscillators in the receiver; these operate by delaying part of the incoming optical signal and then mixing it with itself. This makes the receiver cheaper and less complex, but loses the SNR advantage gained from the local oscillator (although other advantages of phase modulation, such as constant transmitted power, are maintained).

Ref: Gower (2nd ed.) chapter 24; also Miller and Kaminow chapter 21 (see syllabus list)

12. Optical Amplifiers

Coherent detection is a powerful method for improving signal-to-noise ratio, and thus increasing the length-bandwidth product possible without electronic repeaters. However, lack of suitable sources and difficulty of sufficiently reducing dispersion mean that this has still not come into significant commercial use. An alternative method to improve SNR has come into wide adoption, however, and this is the fibre amplifier.

It has been known for several decades that an optical signal could be amplified by stimulated emission occurring as the signal passed through a medium with a suitable excited state. The breakthrough, at Southampton University in 1987, was the demonstration of high gain amplification in an erbium doped single mode silica fibre, for a 1.55 μm signal. This showed that such amplifiers were practical, and commercialisation of EDFAs (erbium doped fibre amplifiers) followed very quickly, with tremendous impact on the capacity of fibre systems.

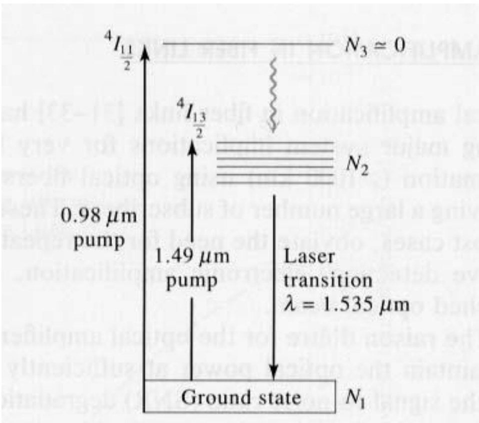


Fig. 12.1 Energy levels and relevant transitions for the erbium ion Er^{3+} in silica fibre amplifiers. From A. Yariv, Optical Electronics, 4th Ed., Fort Worth: Saunders College Pub., 1991, p. 450.

Essentially an optical amplifier works like a laser but without the resonant cavity. Electrons in the amplifying medium, or in a dopant, are energised to an excited state, and signal photons passing through the medium cause stimulated emissions, each of which produces a photon which is an exact copy in phase, wavelength, direction and polarisation of the original. The transition must have the same energy as the signal photons, and in practice there will be a band of allowable transition energies which will correspond to some equivalent band of signal wavelengths. This gives the highly desired property that any optical signal within the wavelength band is amplified, regardless of its modulation properties (particularly bit-rate), or its precise wavelength. Even a wavelength multiplexed signal will be amplified without needing separation. Practical implementation of such amplifiers has given an enormous boost to optical networks, allowing link designs to be made almost without concern about signal-to-noise ratios! One serious limitation is that much of the currently installed fibre network is intended for operation at 1.3 μm. A practical amplifier for this wavelength window has not been developed, and use of 1.55μm signals in these fibres entails large values of dispersion, which then becomes the limiting factor in the length-bandwidth performance. Optical amplifiers increase signal amplitude, but do not correct dispersion, although they can be used in conjunction with dispersion compensating fibre lengths.

Most lasers for optical communications are semiconductor diode lasers, as we have discussed previously. In an amplifier, the signal passes along the gain path only once, and so it is difficult to get high gain in a short distance, which is all that is practical in semiconductor structures. A doped fibre can be many meters long, and this is a major advantage. However, the generation of the excited state can not be carried out electrically in a fibre, which is an insulator. Instead, we supply the energy optically, as we do in any glass laser, using a so-called pump source. This source can be an incoherent one, such as a flash lamp, but in the fibre amplifier the pump is usually a semiconductor diode laser. Also, the excited state in the fibre is within individual dopant ions, so the energy levels are discrete, rather than in bands as in the semiconductor. Absorption of a pump photon produces the excited state, from which stimulated emission can occur.

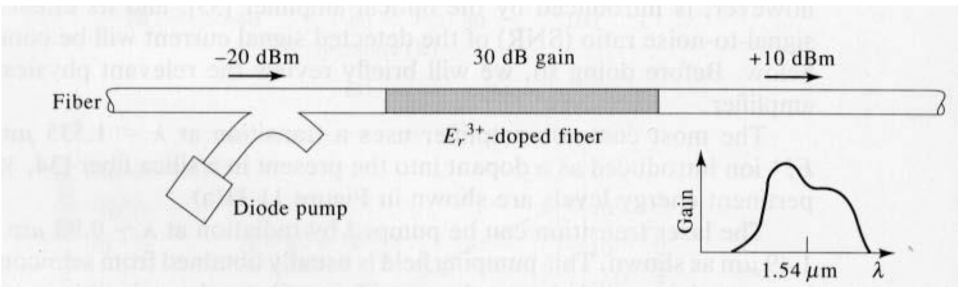


Fig. 12.2 Structure of an EDFA. From Yariv.

A single transition, however, is not satisfactory. This is because the transition rates for absorption and spontaneous emission are the same, so if this rate is high, the excited state would decay spontaneously too quickly, while if the rate was low the pump energy would not be sufficiently absorbed. In addition, we do not want the pump photons to stimulate emissions, only the signal photons. These problems are solved using a 3-level system, as illustrated in Fig. 12.1, in which the initial excited state decays very quickly to an intermediate state, whose spontaneous decay lifetime is long. The Er^{3+} ion in silica, as illustrated, provides the right combination of energy levels and transition rates. In practice, other dopant oxides are added to the glass to improve the solubility of the Er and to increase the spontaneous decay lifetime, which is then typically about 10 ms. Even better performance is possible with a 4-level system, in which the stimulated transition is down to a lower intermediate state, from which the electron rapidly decays to the ground state. The advantage is that in the 3-level system, signal photons can be absorbed by exciting a transition from the ground state directly to the intermediate state, whereas the 4-level system does not have this problem.

The amplifier, as indicated in Fig 12.2, consists of the pump laser, a coupler for combining the pump and signal, and a length of doped fibre which forms the gain path. The EDFA (erbium doped fibre amplifier) can give 30 dB gain over a spectral range of more than 40 nm as illustrated in Fig. 12.3; this corresponds to a signal bandwidth of about 5000 GHz!

While the EDFA can easily boost the optical power to levels sufficient to overcome receiver noise, and to reduce the effects of shot noise, there is a price to be paid in the form of ASE : Amplified Spontaneous Emission. Spontaneous emission from the intermediate state cannot be eliminated altogether, and these spontaneous photons will then be amplified just as the signal photons are. This produces a noise power which can be shown to be uniformly distributed within the optical frequency range of interest, such that within a band of optical frequency $\Delta\nu$ the noise power at the output of the amplifier is given by:

$$\Phi_A = \mu\, h\nu\, \Delta\nu\, (G-1)$$

where G is the optical power gain of the amplifier, and μ is a dimensionless atomic inversion factor of the transition. This factor depends on the populations of the ground and excited states, N_1 and N_2 respectively, according to:

$$\mu = N_2/(N_2 - N_1)$$

which takes the minimum value 1 for full inversion (note that amplification is not possible unless $N_2 > N_1$).

The resultant noise in the detector is created not by the ASE itself but by the mixing (or beating) of the ASE with the actual signal. This mixing can be analysed in the same way as the mixing of the signal and local oscillator in the coherent receiver. The ASE field at frequency ν is :

$$E_A(t) = \hat{E}_A \cos(\nu t)$$

and since $I_{ph} = e\Phi/h\nu$, neglecting detector efficiency, the detector current will be

$$I_{ph}(t) = \frac{e}{h\nu} \frac{A}{Z} \left\langle \left[E_r(t) + E_A(t)\right]^2 \right\rangle$$

$$\left[E_r(t) + E_A(t)\right]^2 = \hat{E}_r^{\,2}(1 + \cos 2\omega_r t)/2 + \hat{E}_A^{\,2}(1 + \cos 2\nu t)/2$$
$$+ \hat{E}_r \hat{E}_A \left(\cos(\omega_r t + \nu t) + \cos(\omega_r t - \nu t)\right)$$

The first term is the signal itself. The second term will be negligible, since $\hat{E}_A^2 << \hat{E}_r^2$, and the third is averaged to zero by the detector. The fourth term contributes noise within the signal bandwidth B if $\omega_r - B < \nu < \omega_r + B$, so that the relevant bandwidth $\Delta\nu = 2B$. Then we have:

$$I_A(t) = \frac{e}{h\nu}\frac{A}{Z}\hat{E}_r\hat{E}_A\,(\cos(\omega_r t - \nu t))$$

Note that $E_r(t)$ is the field of the amplified signal, so that if Φ_o is the signal flux at the amplifier input, we have :

$$G\Phi_o = \frac{A}{2Z}\hat{E}_r^2$$

and integrating the ASE power gives

$$2\mu\, h\nu\, B\,(G-1) = \frac{A}{2Z}\hat{E}_A^2$$

Combining these equations, and taking the mean square of the beat current we arrive at :

$$(I_A^*)^2 = 4\mu\frac{e^2}{h\nu}G(G-1)\Phi_o \qquad = 4\mu e(G-1)I_{ph}$$

(remembering that $(I^*)^2 = I^2/B$, and using in the second formulation $I_{ph} \cong eG\Phi_o/h\nu$) .

Comparing this to the equivalent shot noise term, and remembering that the signal current has been increased by G, we find that (taking $G-1 \cong G$) the SNR has been decreased by a factor $\sqrt{2\mu}$. Since $\mu >= 1$, such that the optical SNR is worse than in the shot noise limited case by at least $\sqrt{2}$, or equivalently, the electrical noise is worse by at least 2, or 3 dB. The noise figure (NF) of an optical amplifier is usually given in dB and refers to electrical SNR, and so the minimum NF (obtained for full inversion, $\mu = 1$) is 3dB.

Typical values for gain and noise figure for an EDFAS are shown in Fig 12.3.

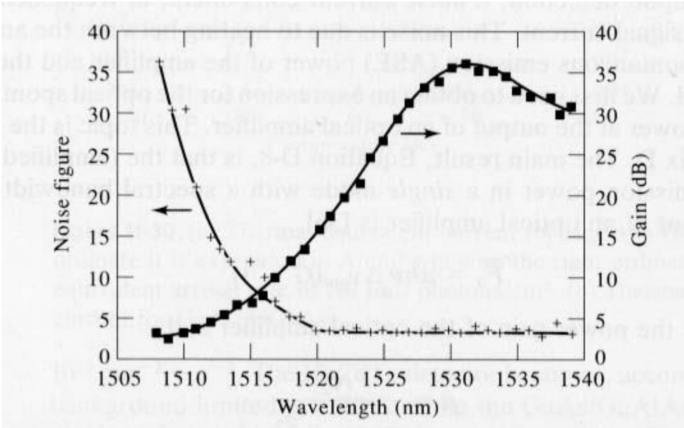


Fig. 12.3 Gain and noise figure vs. wavelength for a typical EDFA. From Yariv.

Part D: Summary

Single Links

power budget calculations; length × bit-rate limits
bit-error rate, relation to signal variance
pulse spreading for Gaussian spectrum; resulting power penalty
chirp penalty

Coherent Communications

relation between optical power and electric field amplitude
combining received signal with local oscillator in 4-port coupler
mixing in photodetector, resulting difference frequency term
shot noise in coherent receiver; SNR
minimum # photons/bit
configurations for heterodyne receivers

Optical Amplifiers

use of stimulated emission for gain
energy levels in Er^{3+} ion; optically pumped amplification at 1.5 μm
structure of the EDFA
3- and 4-level laser systems
amplified spontaneous emission noise; amplifier noise figure

Part D: Problems

1. a) A fibre optic link has two connectors both of which have power reflection coefficients R, where R << 1. The secondary reflection results in an additional signal arriving at the photodetector with a delay substantially greater than the bit period. Mixing of this reflection and the main signal in the detector causes beat noise. Find an expression for the mean square spectral density $(I_R^*)^2$ of this noise current, and the resulting signal to noise ratio, neglecting other noise sources. The signal optical power is Φ_o, and propagation loss can be neglected.

   b) For R = 0.01, and signal wavelength and bandwidth 1.55 µm and 100 Mhz respectively, find the range of Φ_o for which the reflection noise is more significant than shot noise.

NOTE: Some previous years' notes have an extra section in Part D on networks, and associated problems, but this material is not covered this year. Also, this year the detailed content on Coherent Comms (Part 11) will not be examined but is included for your interest (however, you should know the basics, i.e. what coherent comms is and what limits its application in optical systems).