# Lecture 7 - Digital Communication and Storage II

In this lecture I shall be looking at how we can use the transmitters, receivers and fibres that have been introduced in earlier lectures to make a practical fibre-optic communication system. I shall also look at CD players as an important example of digital storage.

## 1 Optical Fibre Communications in Practice

We saw in lecture 3 how analogue audio (or video) signals can be digitised so that any information may be encoded in binary format. To transmit large amounts of binary information efficiently a large bandwidth ($> 100\,\mathrm{MHz}$) is required. The advent of optical fibres with losses of less than $1\,\mathrm{dB\,km^{-1}}$ and with high information carrying capacity has meant that they have become very attractive alternatives to twisted pair or coaxial cables in many types of communication links. Because of the skin effect, the attenuation of wire based systems increases as the square root of the frequency of the carrier. At a frequency of $100\,\mathrm{MHz}$, for example, $9.5\,\mathrm{mm}$ diameter coaxial cable has an attenuation of $\sim 20\,\mathrm{dB\,km^{-1}}$, significantly worse than the optical losses in a silica fibre even at a wavelength of $850\,\mathrm{nm}$ ($\sim 2\,\mathrm{dB\,km^{-1}}$).

A schematic diagram of an optical transmission system is shown in figure 7.1. The emitter is usually an LED or semiconductor laser, whilst the detector may be a p-i-n or avalanche photodiode. Also included in this diagram are what are known as *repeater* units, which are present to counter the effects of fibre transmission losses and dispersion. In a repeater the weak, and possibly broadened, optical pulses are detected and converted to (weak) electrical pulses. These pulses are then reshaped, amplified and re-timed, to reproduce, as far as is possible, the original noise-free pulse train. Finally, a new set of optical pulses is generated for transmission down the next section of fibre. A separate power supply line must be be provided for the repeater units and, as might be imagined, their presence adds greatly to the cost of a fibre-optic link. Modern systems, as described later, replace these electrical repeaters by optical fibre amplifiers, and soliton based schemes.

The main advantages of optical fibre links include relatively low signal attenuations, high bandwidths (up to several gigahertz or more), small physical size and weight, the elimination of ground loop problems an immunity from electrical interference. Fibre cables may also survive better than copper cables in certain corrosive or hazardous

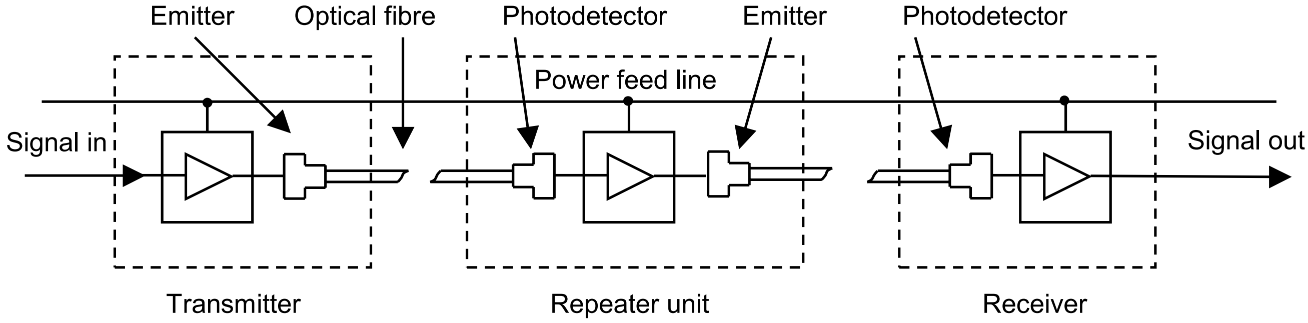


Figure 7.1 Schematic diagram of the main components of a fibre optical communication system involving a repeater.

| Digital rates (Mbps) | | Approximate number of telephone channels (1 channel = 64 kbps) | |
|---|---|---|---|
| *Europe*: | | | |
| 2.048 | | 32 | |
| 8.448 | | 120 | (4 × 30) |
| 34.368 | | 480 | (4 × 120) |
| 39.364 | | 1920 | (4 × 480) |
| 565.992 | | 7680 | (2 × 7680) |
| 1120 | | 15 530 | (2 × 7680) |
| 2400 | | 30 720 | (2 × 15 530) |
| *United States*: | | | |
| 1.544 | T1 | 24 | |
| 6.312 | T2 | 96 | (4 × T1) |
| 44.736 | T3 | 672 | (7 × T2) |
| 274.176 | T4 | 4032 | (6 × T3) |

Table 7.1 Digital rates used in telecommunications in Europe and the United States.

environments such as in sea water or inside nuclear reactors. The clear superiority of fibre-based communication has meant that all modern trunk telecommunication lines are fibre based. A single telephone call requires a bandwidth of 4 kHz ($64\,\mathrm{kb\,s^{-1}}$). By using multiplexing techniques, many channels are carried simultaneously. Standard rates for Europe and America are listed in table 7.1. Another obvious area where fibres are dominant is in computer communications. Fast Ethernet communication over large distances ($> 100\,\mathrm{m}$) use fibres. Multimode fibres can transmit at $100\,\mathrm{Mb\,s^{-1}}$ over a distance of 2 km. Faster gigabit Ethernet requires single mode fibres.

## 1.1 Wavelength Considerations

The two crucial characteristics of optical fibres that depend on wavelength are attenuation and dispersion. These were discussed at length in lecture 4. Here I shall give a brief summary of the results and their relevance to practical fibre communication systems. The attenuation versus wavelength for a silica-based bibre is shown in figure 7.2. Material dispersion implies that the pulse spread $\Delta\tau$ over a fibre of length $L$ can be written as

$$\Delta\tau \approx -\frac{L}{c}\,\lambda_0^2\,\frac{\mathrm{d}^2 n}{\mathrm{d}\lambda_0^2}\,\frac{\Delta\lambda_0}{\lambda_0}.$$

The pulse spread per unit length per unit wavelength interval, $\Delta\tau/L\Delta\lambda_0$, for silica is shown in figure 7.3. The first generation of optical fibre transmission systems were based GaAs or AlGaAs, and hence operated around 850 nm. This was by no means ideal, at 850 nm silica fibres have attenuations of $\sim 2\,\mathrm{dB\,km^{-1}}$ whilst the dispersion is $\sim 80\,\mathrm{ps\,nm^{-1}}$. Thus light from a typical LED with a linewidth of 50 nm would suffer a dispersion of $4\,\mathrm{ns\,km^{-1}}$. Thus the bandwidth would be reduced to 100 MHz over a distance of 2.5 km, in agreement with current FDDI Ethernet fibre links. Since lasers have much narrower linewidths, they can achieve higher bandwidths over the same length of fibre than LEDs.

The minimum attenuation occurs at $\sim 1.55\,\mu\mathrm{m}$ ($\sim 0.15\,\mathrm{dB\,km^{-1}}$). The minimum for the dispersion occurs at a slightly shorter wavelength of $1.3\,\mu\mathrm{m}$ (see figure 7.3). Thus all long distance, high capacity telecommunication links now operate at wavelengths of $1.3\,\mu\mathrm{m}$ or $1.55\,\mu\mathrm{m}$, and suitable p-i-n and APD detectors for these wavelengths based on InGaAs are now readily available.

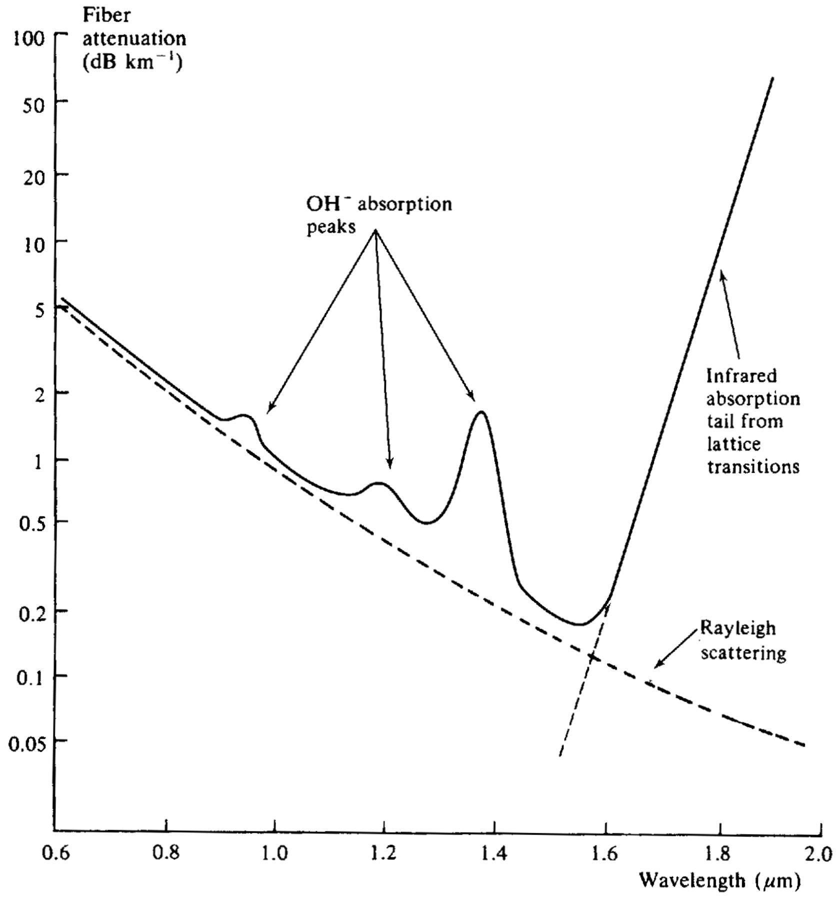


Figure 7.2 Typical attenuation versus wavelength plot for a silica based optical fibre. The contribution from Rayleigh scattering is shown, as are the other two main loss mechanisms, the IR absorption tail and the $OH^-$ absorption peaks.

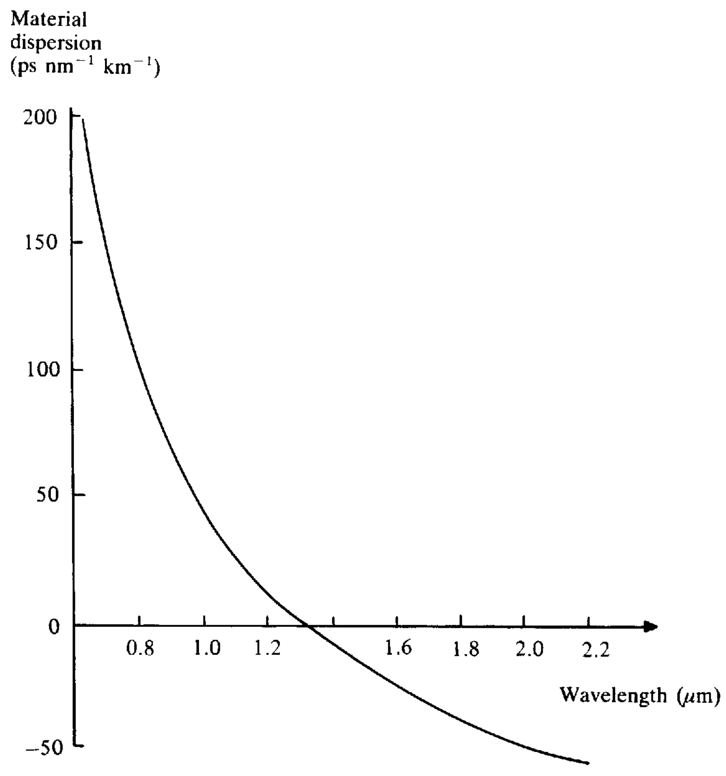


Figure 7.3 Material dispersion for pure silica.

## 1.2 Fibre Coupling Considerations

It is important for an emitter to be able to couple a useful amount of power into a fibre, and also be able to achieve the desired bandwidth. Compared with semiconductor lasers, LED sources are easy to drive, have long lifetimes and are cheap. They have greater linewidths, however, and are much less efficient at launching power into fibres. This is mainly because of their larger emitting areas relative to fibre cores, greater beam divergence and incoherent output. For a source of area $A_s$ and a fibre of core area $A_c$ and numerical aperture NA the coupling efficiency $\eta_c$is

$$\eta_c = \frac{A_c(\mathrm{NA})^2}{A_s n_0^2}$$

where $n_0$ is the refractive index of the medium outside the fibre. A typical multimode fibre has NA = 0.3, so in air we have $\eta_c = 9\%$ — only a small fraction of the light emitted from an LED in contact with a fibre is transmitted.

The two main types of LED most often used in fibre optical systems are the surface-etched well emitter (also known as the 'Burrus' type) and the edge emitter, both of which are illustrated in figure 7.4. In order to couple the strongly asymmetric emission pattern from an edge emitting LED or laser into a fibre an anamorphic lens is often used, as illustrated in figure 7.5. The coupling criterion for lasers into single mode fibres is more complex since it relies on the use of Gaussian beam optics, however, for a laser with a beam waist $\omega_B$ and a single mode fibre with mode size $\omega_0$ the coupling efficiency can be shown to be

$$\eta_c = \frac{4\omega_0^2\omega_B^2}{(\omega_0^2 + \omega_B^2)^2}$$

Thus if the focussed spot has a diameter of twice the fibre mode size, the coupling efficiency will be $(4/5)^2 = 0.64$. This is obviously much better than can be achieved using an LED.

Lasers also tend to have narrower linewidths than LEDs, enabling high transmission rates over long distances. A semiconductor laser with a Fabry–Perot-type cavity usually has a (multimode) linewidth of $\sim 3\,\mathrm{nm}$. By using a distributed feedback system the laser can be made to operate in a single longitudinal mode giving linewidths $\sim 10\,\mathrm{MHz}$. With external cavities this can be reduced to $\sim 10\,\mathrm{kHz}$. Thus lasers are better sources for high-bandwidth optical fibre communications than LEDs, although they are more expensive and require good temperature control and output stabilization.

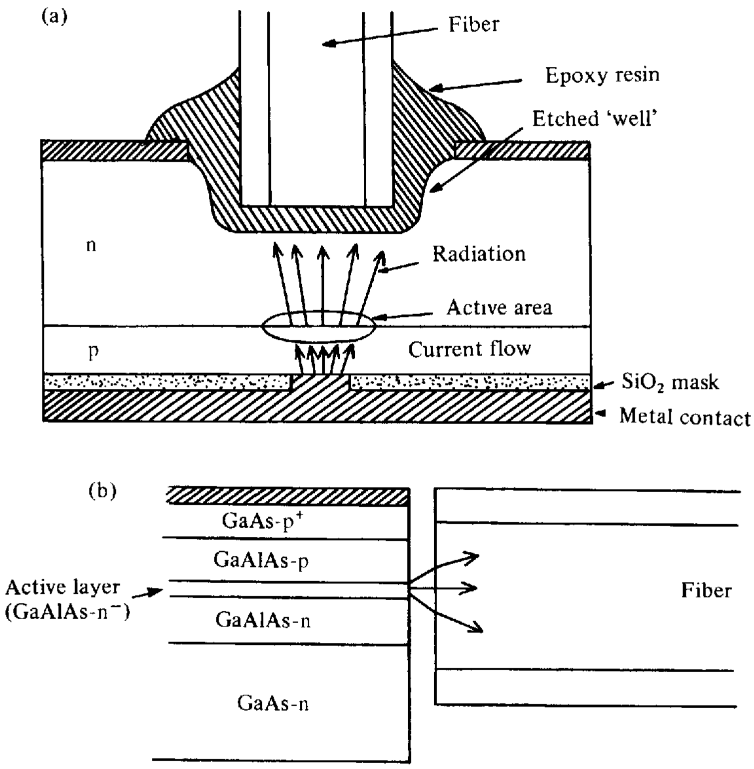


Figure 7.4 Two types of emitter designed for more efficient coupling into optical fibres; (a) the etched well or "Burrus" type and (b) an edge emitter. In (a) the active light-emitting area is restricted to a small region just below the end of the fibre by the use of an $SiO_2$ mask. The fibre is held in position by the use of a transparent epoxy resin, which also helps to reduce Fresnel losses. In the edge emitter, the radiation is confined to a narrow light-guiding layer with a structure very similar to that of a double heterostructure laser.

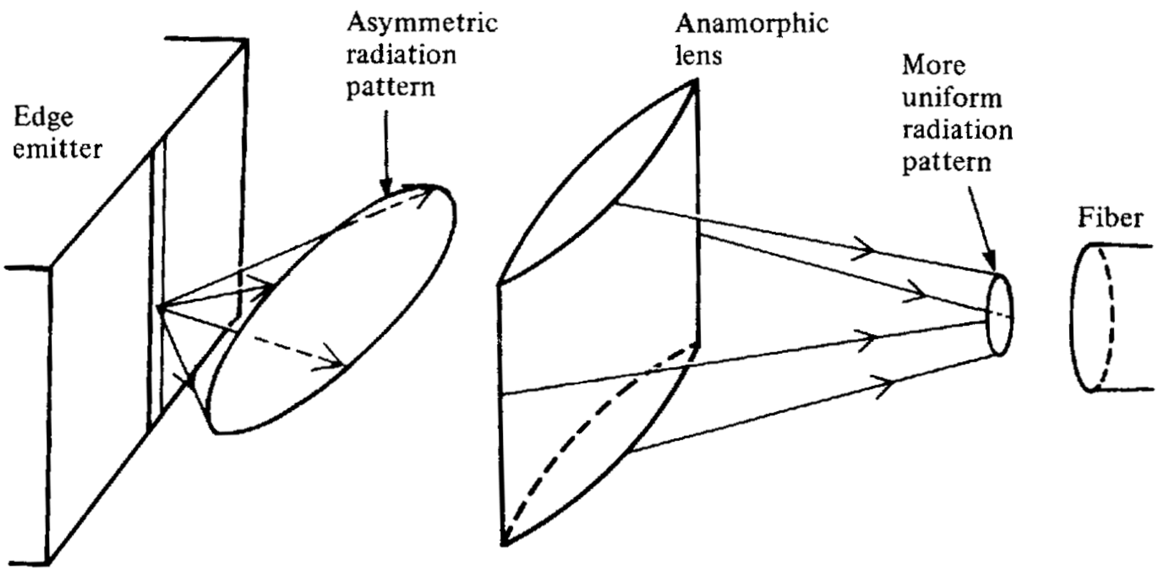


Figure 7.5 Strongly asymmetrical radiation patterns from edge emitters can be coupled more efficiently into a fibre using an anamorphic lens (I.e. one with differing curvatures in two orthogonal directions.

## 1.3 Erbium Doped Fibre Amplifiers

Electronic repeaters are both expensive and complicated. They can therefore form a considerable part of the cost of a long fibre trunk route. Complex systems are also less reliable as there are more components to go wrong! A much more elegant solution would be to develop an 'all optical' amplifier that did not require the optical signal to be turned into an electrical signal and then back again into an optical one. In lecture 4 the erbium doped fibre laser was introduced. Without it's cavity this becomes a perfect optical amplifier. These only work at 1.55 $\mu$m, amplifiers at 1.3 $\mu$m are still in development. For completeness I have included the energy diagram of the $Er^{3+}$ system in figure 7.6. The first excited state ($^4I_{13/2}$) has a relatively long lifetime and gain is easily obtained once sufficient population inversion has been achieved between it and the ground state ($^4I_{15/2}$). Because the ions are incorporated into an amorphous solid matrix the energy levels are considerably broadened and gain may be obtained over a range of some 30–40 nm. One big advantage is that, unlike the repeater unit, the operation of the device is independent of both the modulation coding and of the bit rate (up to $\sim$ 100 GHz).

The most effective pump wavelength is 980 nm. A semiconductor laser can be used as a pump source. Small-signal gains as high as 45 dB and incremental gains of 10 dB $mW^{-1}$ of pump power having been achieved. Figure 7.7 shows a typical gain curve of a length of $Er^{3+}$-doped silica fibre. Ideally the gain curve should be flat so that all signals within the gain profile will be amplified by the same amount; in this case this is clearly not so. This can be corrected by using a suitably designed optical filter. The physical layout of an erbium fibre amplifier is shown in figure 7.8. A means has to be found to inject both the signal and the pumping radiation into the core of the doped fibre section. This is normally achieved by means of a biconical fused fibre coupler (see lecture 8). In this device the fractional amount of radiation coupled from one input fibre to the other output fibre varies with the coupling length $z$ as $\sin^2(Cz)$. The value of $C$ depends on wavelength and since pump and signal are at different wavelengths, it is possible to have a situation where the two different input fibres both couple with nearly 100% efficiency into the same output fibre as illustrated in figure 7.8. Polarization insensitive optical isolators are used to prevent unwanted feedback into the gain region from reflections outside, since this might give rise to unwanted laser action. In the absence of optical pumping the amplifier will become highly absorbing, therefore backup pump lasers are essential in case of pump laser failure.

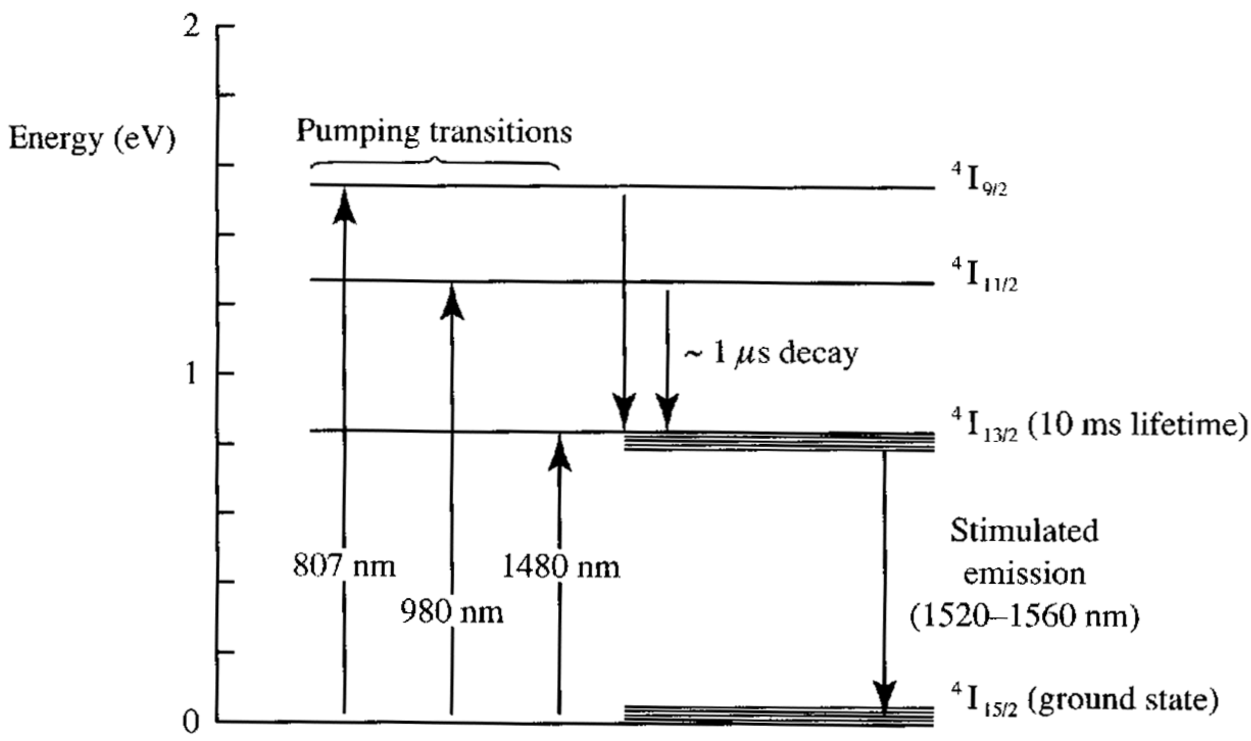


Figure 7.6 The energy levels for $Er^{3+}$-doped silica fibre that are involved in the amplification of light over the wavelength range 1520 nm to 1560 nm using stimulated emission.

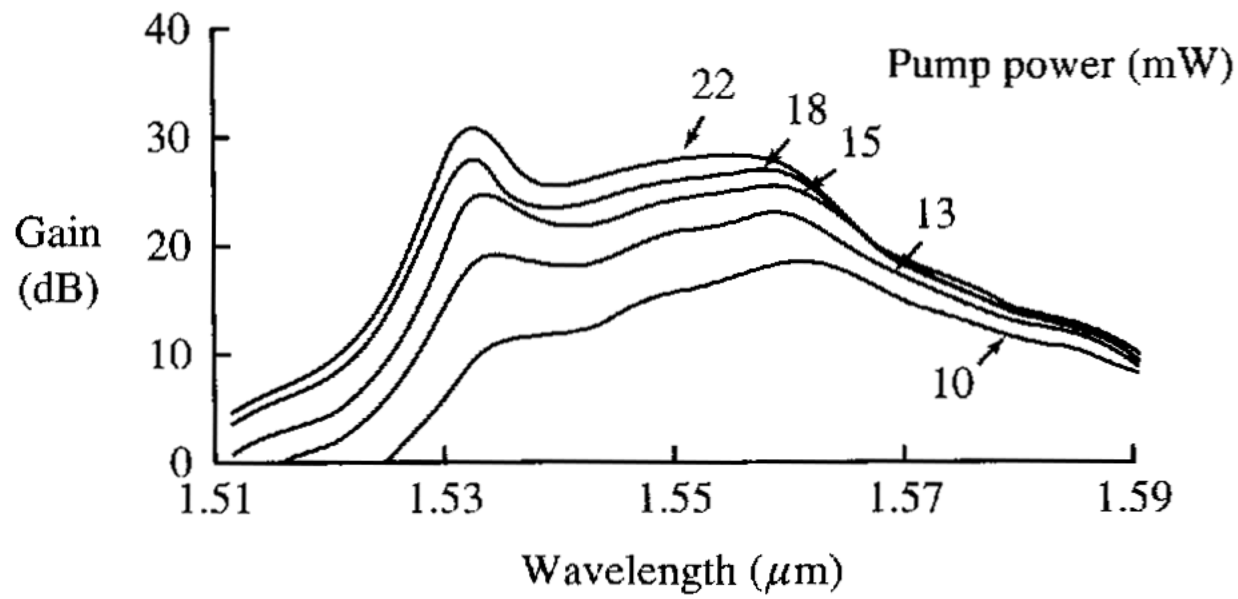


Figure 7.7 Variation of gain with wavelength for a typical erbium-doped fibre amplifier. The results at several different pump powers with a wavelength of 1485 nm are illustrated.

Although the $Er^{3+}$-doped fibre amplifier has proved very successful there are other systems that have been proposed for use as optical amplifiers; one of the most important of these is the semiconductor laser amplifier. If the end facets of a semiconductor laser are anti-reflection coated then the device will function as an amplifier. The advantages of the semiconductor laser amplifier are a much greater freedom in the choice of operating wavelengths (via band structure engineering), and a broader and smoother gain curve. The main disadvantages are that the gain is dependent on the direction of polarization and that coupling losses between fibres and the amplifier can be significant unless great care is taken. It is possible that semiconductor laser amplifiers will find greater use in integrated optical circuits as discussed in the next lecture.

## 1.4 System Design

Three main factors are paramount in influencing system choice, namely bandwidth, maximum transmission distance and, of course, total system cost. For short distances ($\sim$ 500 m) and for bandwidths up to $\sim$ 10 MHz a cost-effective combination is a red-emitting LED, a high NA all-plastic fibre and a Si photodiode detector. More exacting distance and bandwidth requirements may be met with a move to silica-based fibres with an LED or laser source operating at $\sim$ 850 nm. The detector may be either a Si photodiode or a Si APD.

When the transmission rate exceeds 400 Mbs or so, the usable fibre length is determined primarily by fibre dispersion rather than by fibre attenuation. It then becomes necessary to move to single mode fibres and operating wavelengths close to the material dispersion minimum (1.3 $\mu$m in silica fibres). Transmission rates of several gigabits per second over fibres of several tens if not hundreds of kilometres in length are already in use.

The development of the erbium doped optical fibre amplifiers has made 1.55 $\mu$m an attractive wavelength to work at, with the added advantage that fibre losses are about a factor of two smaller than at a wavelength of 1.3 $\mu$m. The downside of this is, of course, that material dispersion will be larger. There are two ways around this: first, if the installation is a new one then dispersion-shifted (or dispersion-flattened) fibre can be used. Secondly, if standard fibre is used then a dispersion compensation scheme may be employed. Here a length of fibre which exhibits a high dispersion in the *opposite sense* to that of the main fibre is inserted just before the receiver. Given a correct length, this

| | |
|---|---|
| Transmitter output | 0 dBm |
| Receiver sensitivity | −50 dBm |
| Required margin | 50 dBm |
| | |
| System loss: | |
| Fibre loss 2 dB km$^{-1}$, 15 km | 30 dB |
| Total splicing loss (0.5 dB ×10) | 5 dB |
| Detector coupling loss | 1 dB |
| Headroom | 5 dB |
| | |
| Total attenuation | 41 dB |

link can cancel out the dispersion in the main fibre, producing a dispersionless link. The increased loss due to the link will be small at 1.55 $\mu$m and can be ignored.

To determine whether or not a chosen system will perform satisfactorily, number of checks must be made. Both detector and emitter must be capable of handling the required signal bandwidth. The fibre dispersion over the length required must not degrade the signal excessively, while for a given bit error rate (BER) there will be a minimum average signal power that must reach the detector. If the power launched into the fibre is known, together with the fibre attenuation, the maximum length of fibre that can be used may be calculated. Allowance must be made for any splices and joins, and a safety margin of $\sim$ 5 dB is usually included. Such a calculation is called a *flux budget*.

As an example, suppose that we have to construct a link of length 15 km and bandwidth 100 Mbps. Components are chosen with the following properties: receiver sensitivity −50 dBm (at 100 Mbps), fibre loss 2 dB km$^{-1}$ and a launch power of 0 dBm (1 mW). Assume 10 splices are needed, each with a loss of 0.5 dB. So from the table we have and excess power margin of $50 - 41 = 9$ dB, which should be sufficient for the operation of the link.

Figure 7.9 shows typical launch powers for both LEd and laser sources and also the received powers required by state-of-the-art receivers to achieve a $10^{-9}$ BER as a function of bit rate. It is obvious from the figure that as the required bit rate increases the relative advantage of the laser/APD combination also increases. Recent transatlantic communication systems such as TAT12 installed in 1995 use a wavelength of 1.55 $\mu$m and achieve a bit rate of 5 Gbps. The link uses erbium doped fibre amplifiers spaced $\sim$ 45 km

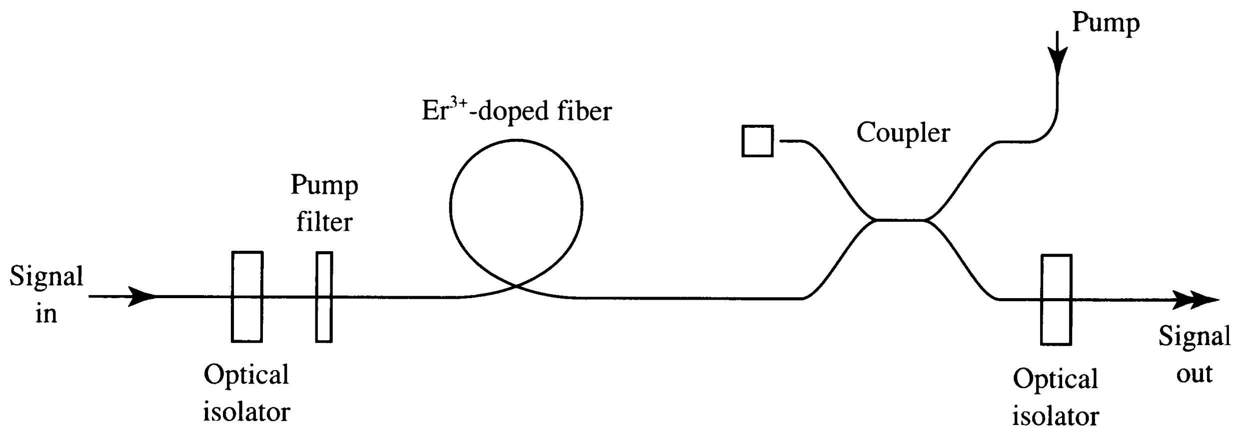


Figure 7.8 The physical layout of an erbium-doped fibre amplifier. In this case the pump and signal beams are moving in opposite directions. A filter is inserted to prevent the small amount of pump radiation remaining after the gain section from propagating further along the fibre. In addition, there are optical isolators at either end of the amplifier; these are to prevent any power at the signal wavelength from being reflected back into the gain section, which could result in the device acting as a laser. More powerful pumping can be achieved with two pump beams moving in opposite directions.

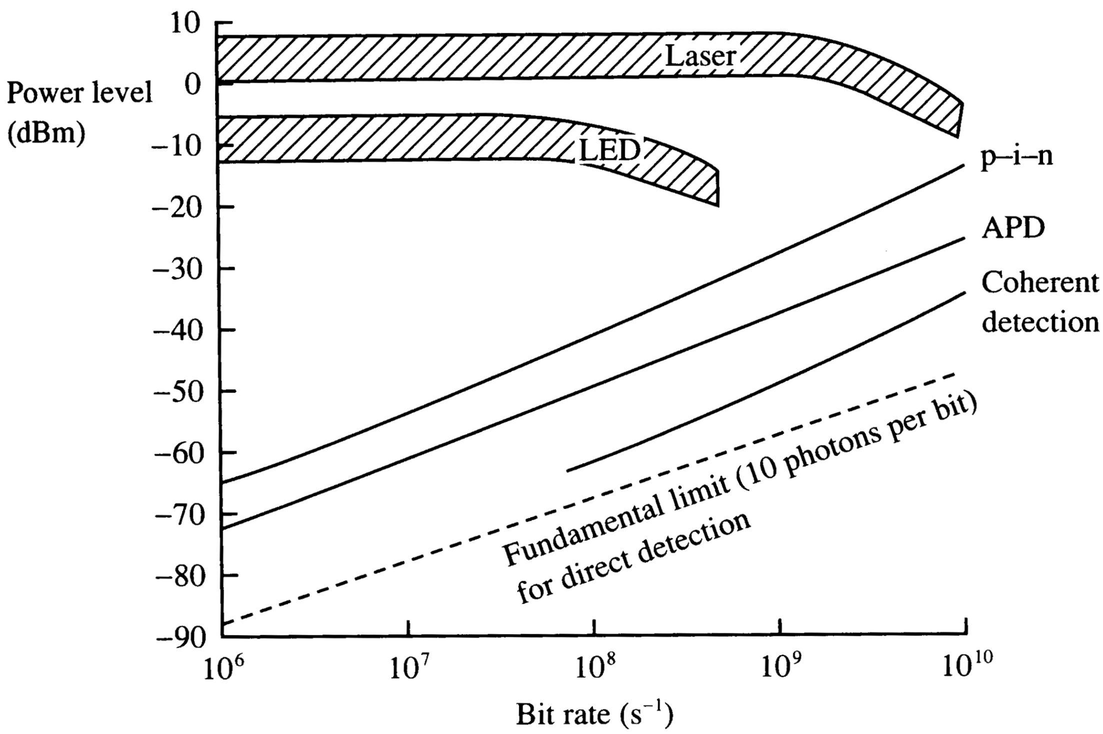


Figure 7.9 Simple illustration of the system flux budget for representative systems. The upper shaded areas show typical launch powers from LED and laser sources, whilst the lower curves show typical receiver sensitivities for a $10^{-9}$ BER at 1.3 μm wavelength. The curves for the p-I-n and APD detectors are for direct detection; a further curve chows results for coherent detection. The fundamental detection limit corresponding to 10 photons per bit is also shown,

apart.

# 2 Optical Disc Storage Systems

*Optical storage* and *optical memory* are terms commonly associated with various existing or proposed mass data storage systems based on the concept of using optical radiation to record and retrieve information. The recording process, which may be either reversible or irreversible, uses the energy of the radiation to produce local modifications in the physical, chemical or structural properties of an optically sensitive medium. In the retrieval or readout mode, a beam of radiation interrogates the medium and the features written on its surface modulate the phase, polarization or intensity of the reflected or transmitted beams, thereby reconstructing the pattern of stored information.

In this part of the lecture I shall be looking at the ubiquitous CD player in some detail, and will then discuss other methods of digital disc storage in a general way.

## 2.1 Replicable storage - CDs and CD-ROMs

The development of replicable optical storage discs was based on the pioneering work of Philips with the *video long playing* (VLP) system, often referred to as "Laser Vision". In 1973, the first 13-bit digital recordings were released on long-playing records, and in 1978 some of the major recording companies began to produce 16-bit recordings. In 1982, the CD was conceived according to a standard proposed jointly by Philips and Sony, which went on to become enormously successful.

The CD is a single-sided plastic disc, 120 mm in diameter, that has a normal playing time of about 75 minutes. In its player system, the playing surface is face down, which allows the surface to be accessed by a laser spot that is made to scan the disc, moving from the outer to the inner radius. Servo mechanisms control both the laser head motion and the disc rpm, so that the playing surface passes the focused laser beam at a constant linear velocity of $\sim 1.3\,\mathrm{ms}^{-1}$. Viewed from above, the disc rotates clockwise at speeds between 200 and 500 rpm. The digital information is stored on the surface of the CD in the form of a continuous spiral of very small pits ($\approx 0.8-3.0\,\mu$m long by $\approx 0.5\,\mu$m wide and $0.12\,\mu$m deep) which are coated by a layer of Al and a transparent plastic overcoat for protection (see figure 7.10). The adjacent turns of the spiral are separated by a track pitch of $1.5-1.7\,\mu$m. The readout signal is produced by reflecting an AlGaAs

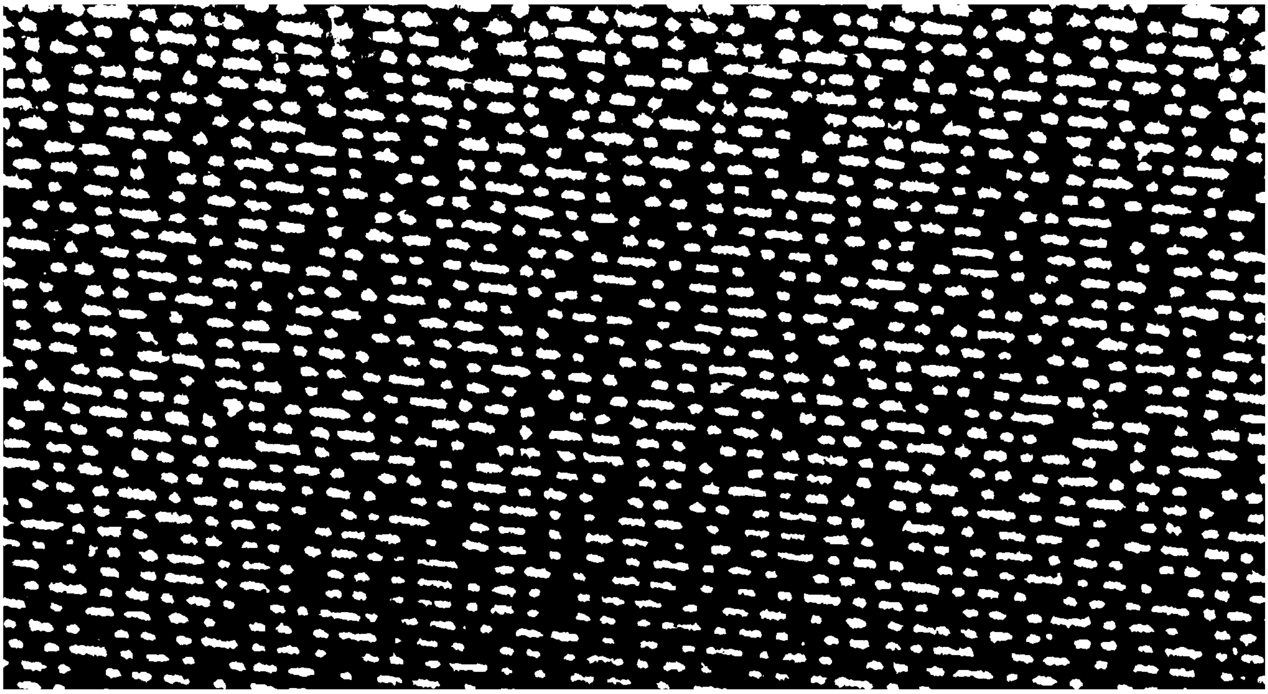

Figure 7.10 Scanning electron micrograph of a compact disc surface showing the recorded data.

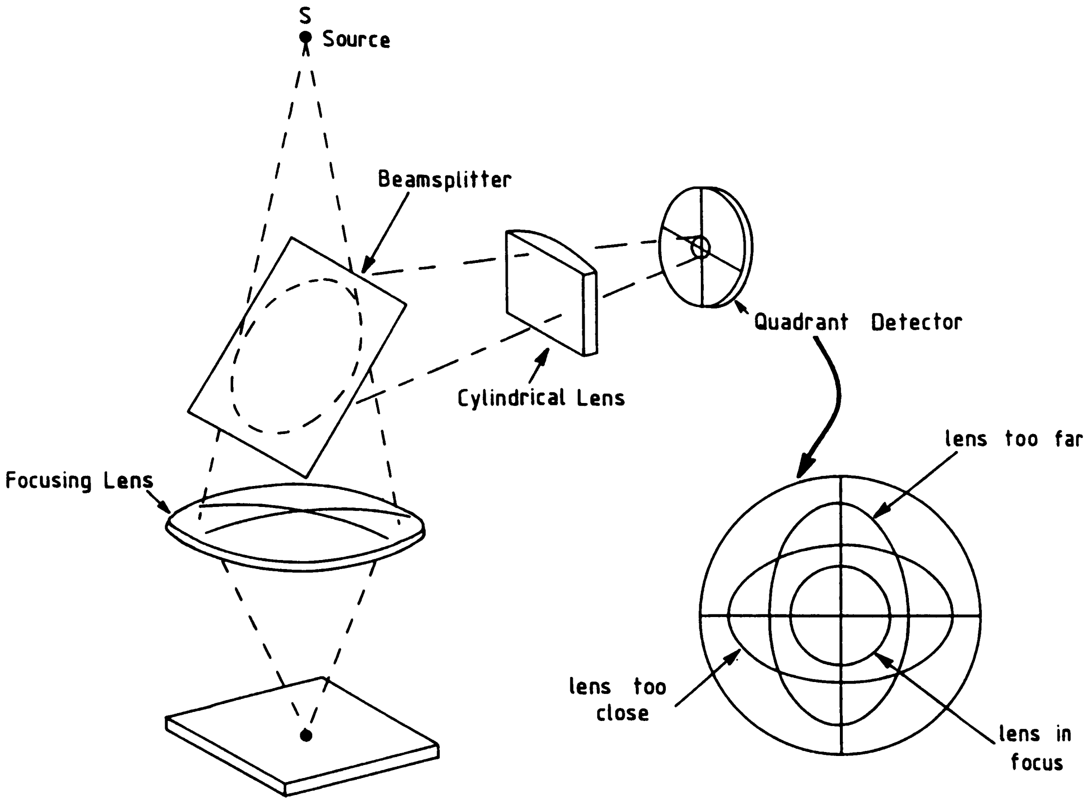


Figure 7.11 Astigmatic focus technique.

diode laser beam ($770 - 790\,$nm wavelength with a spot diameter of $\approx 0.7\,\mu$m produced by focussing through a lens of numerical aperture 0.45) from the surface containing the pits. As this is done from below, the pits appear as bumps. Light is reflected back to a semiconductor photodetector, so that changes in signal strength are produced by the (pit) bit pattern embossed on the surface of the disc. Control and timing information is included in the bit stream.

The optical pickup works by monitoring the light reflected back into the lens. Light is received both from the bottoms of the pits and from the surrounding "land". Because of the fuzziness of the light spot the detector in the pickup cannot "see" the pits and the land separately; it receives only a single aggregate signal. Since the pit bottoms lie closer to the pickup than the land does, there is a phase difference between the contributions made by each to the aggregate signal, and so there is destructive interference. The size and shape of the pits is chosen so that this interference is nearly total. Therefore, much less light is received when the centre of the light spot falls on a pit than when it falls on one of the gaps between the pits along the track. As the disc rotates, the pickup follows the track, so the reflected signal corresponds to the pattern of pits and gaps.

In order to keep the pickup on the track as the disc rotates an electrical signal is obtained optically which indicates whether the pickup is deviating to one side or the other, and is used to drive a servo control.

Two quite different ways are used in commercial players to derive this tracking error signal. In Philips player mechanisms the aperture of the lens is split in half by a prism so that the light returning from each half focuses in a different place. A focus error causes each (de)focussed spot to move laterally, in opposite directions. There is a split photodiode at each focus point (4 diodes in all), the signals within each split pair are subtracted and the two differences are combined to give a focus error signal.

In Japanese players there is a weak cylindrical lens in the returning light beam, giving an astigmatic image. Focus errors cause the image to become elliptical. This image (circular at the ideal focus point) is cast on a 4-quadrant photodiode. Defocussed (elliptical) images fall more on one opposite pair of quadrants than on the other pair; appropriate addition and subtraction of outputs gives a focus error signal. This method is illustrated in figure 7.11.

Focus correction is then done by a servo loop which moves the main lens axially by a moving-coil or moving-magnet mechanism. The tolerance is very small, since the depth of focus is $\sim 2\,\mu$m, whereas the disc may move up and down by as much as

200 $\mu$m. For tracking error sensing, Philips players use the single-spot method (using the same 4 photodiodes already mentioned to subtract total powers in the two halves of the aperture). Japanese players use the three-spot system as a rule, where the detector looks to see whether a greater cancellation of reflected light is observed from one or other of two subsidiary light spots placed to either side of centre, which shows how far off-centre the pits are in relation to the pickup as they pass by. Philips players move the lens on a radial arm controlled by two springs. This is accurate enough for fine and course adjustment. Japanese players move the lens on a slide for course movements and use a two-spring suspension for fine movements.

CDs are mastered as follows. Blue laser light from an argon laser is focussed onto a glass substrate coated with a thin layer of photoresist. The thickness of this layer directly determines the pit depth. Whenever the light is turned on, the photoresist is exposed. In subsequent developing, the exposed areas are dissolved to leave pits, since it is a positive photoresist which is used.

The developed master is then metallised and a layer of nickel is electroplated onto its surface. This is then stripped off to form a negative replica. This can then be used as a stamper in a mould. Hot molten plastic is injected into the mould under pressure. The plastic solidifies in the mould to form the disc, which is subsequently metallised, lacquered (to protect the metal) and labelled.

When the data is written to or read from the CD, the transition from a pit to a land or from a land to a pit, and the length of the land or pit represents the number of zeros (see figure 7.12). Now let us consider the digital data to be put on the disc, and the means by which it is encoded. Audible sounds contain frequencies (if you're young) up to $\sim 20\,\mathrm{kHz}$. Nyquist tells us that we need to sample the audio waveform at 40 kHz or more to record frequencies up to 20 kHz. Moreover, each sample must measure the voltage to enough accuracy to cover the required range of loudness ($\sim 80\,\mathrm{dB}$), and this fixes the number of digits that must constitute each sample (16 bits is more than adequate). In practice, samples of each stereo channel are taken at 44 100 samples per second. Each sample is encoded on 16 bits (2 bytes). This means that samples are generated at a rate of $(44.1 \times 10^3$ samples per second$) \times$ (16 bits per sample) $\times$ (2 channels) $= 1.4112$ $\mathrm{Mbit\,s^{-1}}$. The information is next heavily re-encoded to make possible the detection and correction of errors by the player. The technical name for the encoding scheme is a "8-14 cross-interleaved Reed-Solomon code". It has the effect of multiplying the bit rate by a factor 588/192, bringing the final bit rate up to 4.3218 $\mathrm{Mbit\,s^{-1}}$.

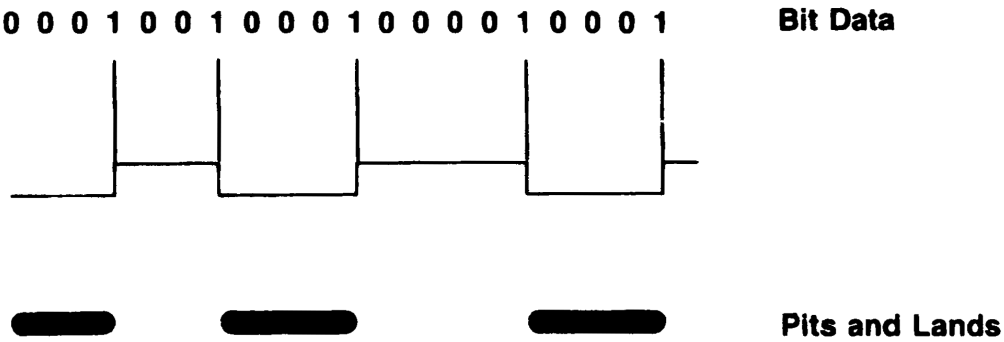


Figure 7.12 Transition between one and zero and zero and one on a CD-ROM. Ones are represented by transitions from pits to lands or from lands to pits.

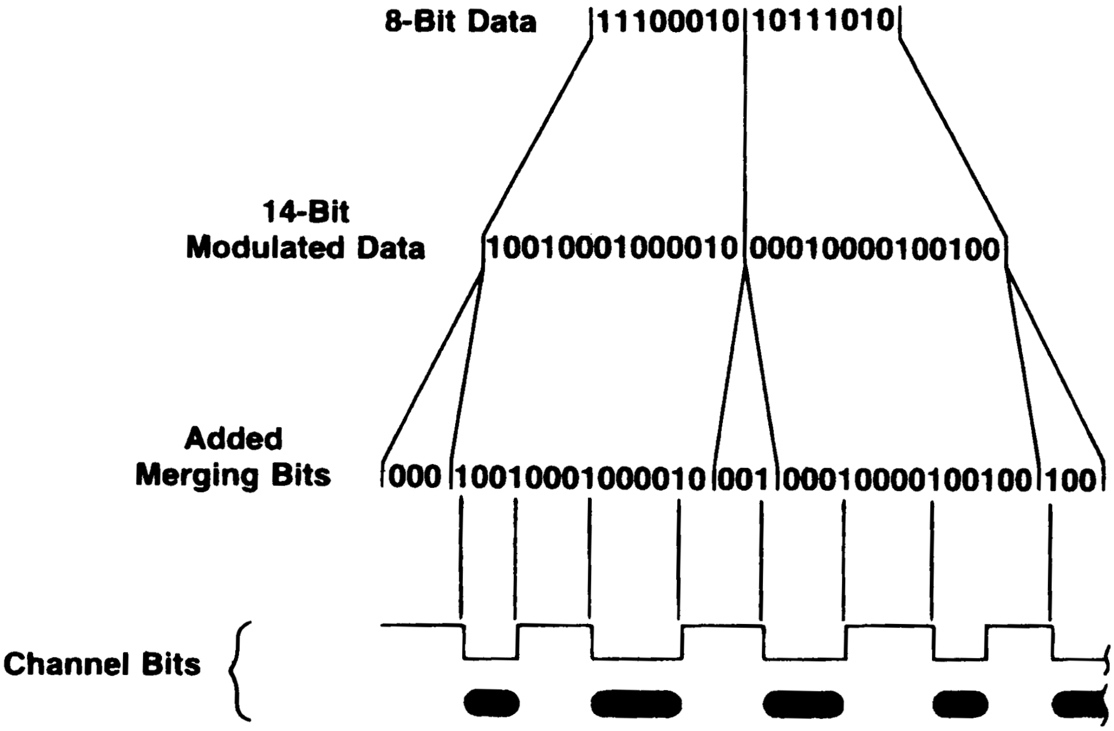


Figure 7.13 Eight-to-fourteen bit encoding example.

| | | |
|---|---|---|
| synchronization pattern | 27 bits | |
| control code | 17 bits | |
| data re-encoded as $24 \times (14+3) =$ | 408 bits | originally 24 bytes (192 bits) |
| error correction | 136 bits | |
| total bits in frame | 588 bits | hence the factor 588/192 |

The ratio 588/192 arises as follows. Each byte (8 bits) in the original signal is re-encoded in a new code that assigns it 14 bits. The 14-bit words are designed so that if an odd bit is misread in the player the error can be detected and put right, the process is known as *eight-to-fourteen modulation* or EFM. Next 3 "merge" bits are added on the end of each 14, which make the join conform to the rules, and make the total number of 0s and 1s equal when encoding is complete. This process is inflicted on each byte in a group of 24; such a group of 24 bytes is the information in a "frame", the basic unit in which information is put onto the disc. Each frame starts with a synchronization pattern of 27 bits, which makes sure that the reading electronics knows where the frame starts. After the synchronization pattern comes a 17-bit control code which tells the reader whether the frame holds audio data, or something else like video or computer data. Finally come 136 bits which are used for further error detection and correction. In summary, each final frame is made up as shown in the table, and is illustrated in figure 7.13. The whole scheme is known as *run-length-limited* (RLL) coding. The effect of this is to reduce the BER from its raw value for a CD system of $\sim 10^{-5} - 10^{-7}$ to a deliverable BER of $10^{-12}$. The 17-byte components belonging to a frame are shuffled round the disc as well, so that any scratch or piece of dirt corrupts data in several small isolated places whose errors can be corrected, rather than destroying a large contiguous sequence of bits.

The really clever thing about all this manipulation is the following. The original signal at $1.4112\,\mathrm{Mbit\,s^{-1}}$ had transitions from 0 to 1 and back again in adjacent bits, at least sometimes, so it needed to be represented by Fourier components up to 706 kHz (plus harmonics). The final code constrains the number of 0s between consecutive 1s to be never less than a lower bound $d$ or more than an upper bound $k$ (the 3 "merge" bits make sure that this is true at the joins of the 14-bit words too). The result is that the frequency requirement is reduced, typically to $4.3218\,\mathrm{Mhz} \div 6 = 720\,\mathrm{kHz}$, barely higher than before. This illustrates a beautiful and rather surprising theorem due to Shannon, that you can build as much error correction as you like into a long encoded signal, with only a negligible increase in the frequency range occupied. The writing laser actually uses pulses, so that a sequence of minimum diameter features is overlapped to achieve the desired mark length in the medium. This is illustrated in figure 7.14.

## 2.2 Other optical disc types

A new standard has emerged recently, DVD (digital video disc or digital versatile disc), which puts more information on a disc. This is done in several ways. First, a shorter wavelength laser ($\sim 640\,\mathrm{nm}$) is used, so the area devoted to a bit can be reduced laterally and longitudinally, the new specification being a pit length of $0.4\,\mu\mathrm{m}$ and a track pitch of $0.74\,\mu\mathrm{m}$. This means that the data is packed approximately 4.6 times more densely, and this alone will raise the disc capacity to $\sim 3\,\mathrm{GB}$. Second, data is stored at more than one depth in the disc using layers (an idea that isn't new and appears in some of the early patents around 1973). There are other minor improvements as well. The number of merge bits is reduced from 3 to 2. The space devoted to error correction is reduced substantially by using a more compact coding that is just as good in principle as the EFM method discussed above, but puts a higher demand on the "reading" computer's ability to calculate—a demand that can realistically be met now. One side of a DVD can actually hold 4.7 GB on a single layer, 8.5 GB on 2 layers and up to 17 GB when used double sided.

Another class of optical discs, used mainly for data archiving, is known as WORM (write once read many) or CDR. Here a powerful laser in the writer burns a pit in a layer of Te/Se alloy. One concern is the lifetime of the discs, as the active layer can become unstable. The projected lifetime of these discs is only about 10 years. Other active layers involving polymers are also in use, together with organic or cyanine dye layers in CDRs from Philips and other manufacturers. These more recent technologies have pushed the life expectancy of CDR discs up to beyond 50 years. A popular new product is the rewritable or erasable CD. Some rely on a reversible phase change from crystalline to amorphous to crystalline again. These systems also have long term stability problems and are probably not suitable for long term archiving.

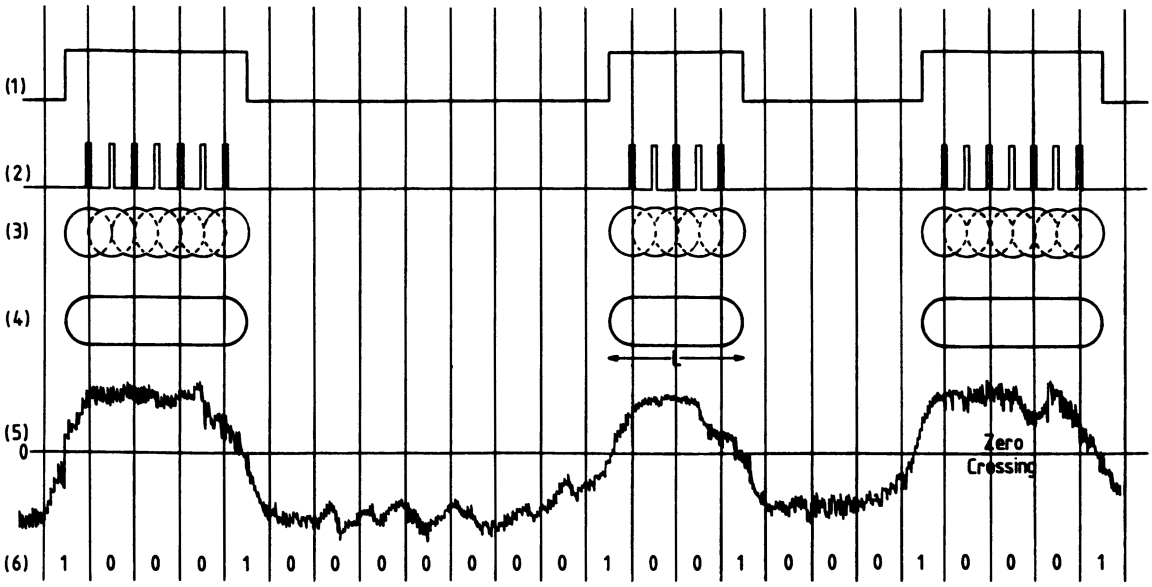


Figure 7.14 Writing and reading process: (1) user data and error correction bits encoded into a RLL pulselength-modulated channel signal; (2) the electrical drive to the laser - note the use of five pulses to write the minimum bit length L; (3) The overlapping pits or marks from continuous features (4) - note the minimum feature length is 3; (5) readout signal - a noisy copy of the write signal; (6) the recovered bit stream.