are concerned only with one channel, we could align the center wavelengths exactly by temperature-tuning the individual mux/demuxes. However, other channels could become even more misaligned in the process (tuning one channel tunes the others as well). In addition, the lasers themselves will have a tolerance regarding their center wavelength. In a cascaded system, wavelength inaccuracies cause additional power penalties due to added signal loss and crosstalk (see Problems 5.19 and 5.20).

5.7 Dispersion

Dispersion is the name given to any effect wherein different components of the transmitted signal travel at different velocities in the fiber, arriving at different times at the receiver. A signal pulse launched into a fiber arrives smeared at the other end as a consequence of this effect. This smearing causes intersymbol interference, which in turn leads to power penalties. Dispersion is a cumulative effect: the longer the link, the greater the amount of dispersion.

Several forms of dispersion arise in optical communication systems. The important ones are *intermodal dispersion*, *polarization-mode dispersion*, and *chromatic dispersion*. Of these, we have already studied intermodal dispersion and chromatic dispersion in Chapter 2 and quantified the limitations that they impose on the link length and/or bit rate.

Intermodal dispersion arises only in multimode fiber, where the different modes travel with different velocities. Intermodal dispersion was discussed in Section 2.2. The link length in a multimode system is usually limited by intermodal dispersion and not by the loss. Clearly, intermodal dispersion is not a problem with single-mode fiber.

Polarization-mode dispersion (PMD) arises because the fiber core is not perfectly circular, particularly in older installations. Thus different polarizations of the signal travel with different group velocities. PMD is proving to be a serious impediment in very high-speed systems operating at 10 Gb/s bit rates and beyond. We discuss PMD in Section 5.7.4.

The main form of dispersion that we are concerned with is *chromatic dispersion*, which has a profound impact on the design of single-mode transmission systems (so much so that we often use the term *dispersion* to mean "chromatic dispersion"). Chromatic dispersion arises because different frequency components of a pulse (and also signals at different wavelengths) travel with different group velocities in the fiber and thus arrive at different times at the other end. We discussed the origin of chromatic dispersion in Section 2.4. Chromatic dispersion is a characteristic of the fiber, and different fibers have different chromatic dispersion profiles. We discussed the chromatic dispersion profiles of many different fibers in Section 2.5.9. As with other kinds of dispersion, the accumulated chromatic dispersion increases with the link length. Chromatic dispersion and the system limitations imposed by it are discussed in detail in the next two sections.

5.7.1 Chromatic Dispersion Limits: NRZ Modulation

In this section, we discuss the chromatic dispersion penalty for NRZ modulated signals. We will consider RZ modulated signals in Section 5.7.2.

The transmission limitations imposed by chromatic dispersion can be modeled by assuming that the pulse spreading due to chromatic dispersion should be less than a fraction ϵ of the bit period, for a given chromatic dispersion penalty. This fraction has been specified by both ITU (G.957) and Telcordia (GR-253). For a penalty of 1 dB, $\epsilon = 0.306$, and for a penalty of 2 dB, $\epsilon = 0.491$. If D is the fiber chromatic dispersion at the operating wavelength, B the bit rate, $\Delta\lambda$ the spectral width of the transmitted signal, and L the length of the link, this limitation can be expressed as

$$|D|LB(\Delta\lambda) < \epsilon. \tag{5.15}$$

D is usually specified in units of ps/nm-km. Here, the ps refers to the time spread of the pulse, the nm is the spectral width of the pulse, and km corresponds to the link length. For standard single-mode fiber, the typical value of D in the C-band is 17 ps/nm-km. For this value of D, $\lambda = 1.55$ μm, and $\epsilon = 0.491$ (2 dB penalty), (5.15) yields the condition $BL < 29$ (Gb/s)-km, assuming $\Delta\lambda = 1$ nm. This limit is plotted in Figure 5.19. Thus even at a bit rate of 1 Gb/s, the link length is limited to < 29 km, which is a severe limitation. This illustrates the importance of (1) using nearly monochromatic sources, for example, DFB lasers, for high-speed optical communication systems, and (2) devising methods of overcoming chromatic dispersion.

Narrow Source Spectral Width

We now consider the case of using sources with narrow spectral widths. Even for such a source, the spectral width of the transmitted signal depends on whether it is directly modulated or whether an external modulator is used. SLM DFB lasers have unmodulated spectral widths of typically less than 50 MHz. Directly modulating a DFB laser would ideally cause its spectral width to correspond to the modulation bandwidth (for example, about 2.5 GHz for a 2.5 Gb/s on-off modulated signal). In practice, however, the spectral width can increase owing to chirp. As the modulation current (and thus optical power) varies, it is accompanied by changes in carrier density within the laser cavity, which, in turn, changes the refractive index of the

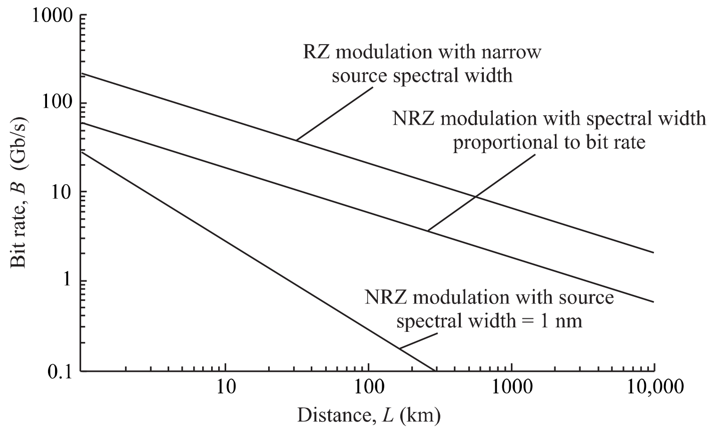

Figure 5.19 Chromatic dispersion limits on the distance and bit rate for transmission over standard single-mode fiber with a chromatic dispersion value of $D = 17$ ps/nm-km. A chromatic dispersion penalty of 2 dB has been assumed in the NRZ case; this implies that the rms width of the dispersion-broadened pulse must lie within a fraction 0.491 of the bit period. For sources with narrow spectral width, the spectral width of the modulated signal in GHz is assumed to be equal to the bit rate in Gb/s. For RZ transmission, the rms output pulse width is assumed to be less than the bit interval.

cavity, causing frequency variations in its output. The magnitude of the effect depends on the variation in current (or power), but it is not uncommon to observe spectral widths over 10 GHz as a consequence of chirp. Chirp can be reduced by decreasing the extinction ratio. The spectral width can also be increased because of back-reflections from connectors, splices, and other elements in the optical path. To prevent this effect, high-speed lasers are typically packaged with built-in isolators.

For externally modulated sources, the spectral width is proportional to the bit rate. Assuming the spectral width is approximately equal to the bit rate, a 10 Gb/s externally modulated signal has a spectral width of 10 GHz, which is a practical number today. At 1.55 μm, this corresponds to a spectral width of 0.083 nm, using the relation $\Delta\lambda = (c/f^2)|\Delta f| = (\lambda^2/c)|\Delta f|$. Substituting $\Delta\lambda = (\lambda^2/c)B$ in (5.15), we get

$$|D|LB^2\lambda^2/c < \epsilon,$$

or

$$B\lambda\sqrt{|D|L/c} < \sqrt{\epsilon}. \tag{5.16}$$

For $D = 17$ ps/nm-km, $\lambda = 1.55\ \mu$m, and $\epsilon = 0.491$ (2 dB penalty), (5.16) yields the condition $B^2L < 3607$ (Gb/s)^2-km. This limit is also plotted in Figure 5.19.

Note that the chromatic dispersion limitations are much more relaxed for narrow spectral width sources. This explains the widepsread use of narrow spectral width SLM lasers for high-bit-rate communication. In addition, external modulators are used for long-distance transmission (more than a few hundred kilometers) at 2.5 Gb/s and in most 10 Gb/s systems.

5.7.2 Chromatic Dispersion Limits: RZ Modulation

In this section, we derive the system limitations imposed by chromatic dispersion for unchirped Gaussian pulses, which are used in RZ modulated systems. The results can be extended in a straightforward manner to chirped Gaussian pulses.

Consider a fiber of length L. From (2.13), the width of the output pulse is given by

$$T_L = \sqrt{T_0^2 + \left(\frac{\beta_2 L}{T_0}\right)^2}.$$

This is the half-width of the pulse at the $1/e$-intensity point. A different, and more commonly used, measure of the width of a pulse is its *root-mean square* (rms) *width* T^{rms}. For a pulse, $A(t)$, this is defined as

$$T^{\text{rms}} = \sqrt{\frac{\int_{-\infty}^{\infty} t^2 |A(t)|^2\, dt}{\int_{-\infty}^{\infty} |A(t)|^2\, dt}}. \tag{5.17}$$

We leave it as an exercise (Problem 2.10) to show that for Gaussian pulses whose half-width at the $1/e$-intensity point is T_0,

$$T^{\text{rms}} = T_0/\sqrt{2}.$$

If we are communicating at a bit rate of B bits/s, the bit period is $1/B$ s. We will assume that satisfactory communication is possible only if the width of the pulse as measured by its rms width T^{rms} is less than the bit period. (Satisfactory communication may be possible even if the output pulse width is larger than the bit period, with an associated power penalty, as in the case of NRZ systems.) Therefore, $T_L^{\text{rms}} = T_L/\sqrt{2} < 1/B$ or

$$BT_L < \sqrt{2}.$$

Through this condition, chromatic dispersion sets a limit on the length of the communication link we can use at bit rate B without dispersion compensation. T_L is a function of T_0 and can be minimized by choosing T_0 suitably. We leave it as an exercise (Problem 5.22) to show that the optimum choice of T_0 is

$$T_0^{\text{opt}} = \sqrt{\beta_2 L},$$

and for this choice of T_0, the optimum value of T_L is

$$T_L^{\text{opt}} = \sqrt{2|\beta_2|L}.$$

The physical reason there is an optimum pulse width is as follows. If the pulse is made too narrow in time, it will have a wide spectral width and hence greater dispersion and more spreading. However, if the pulse occupies a large fraction of the bit interval, it has less room to spread. The optimum pulse width arises from a trade-off between these two factors. For this optimum choice of T_0, the condition $BT_L < \sqrt{2}$ becomes

$$B\sqrt{2|\beta_2|L} < \sqrt{2}. \tag{5.18}$$

Usually, the value of β_2 is specified indirectly through the *dispersion parameter D*, which is related to β_2 by the equation

$$D = -\frac{2\pi c}{\lambda^2}\beta_2. \tag{5.19}$$

Thus (5.18) can be written as

$$B\lambda\sqrt{\frac{|D|L}{2\pi c}} < 1. \tag{5.20}$$

For $D = 17$ ps/nm-km, (5.20) yields the condition $B^2L < 46152$ (Gb/s)^2-km. This limit is plotted in Figure 5.19. Note that this limit is higher than the limit for NRZ modulation when the spectral width is determined by the modulation bandwidth (for example, for external modulation of an SLM laser). However, for both RZ and NRZ transmission, the bit rate B scales as $1/\sqrt{L}$.

Note that we derived the dispersion limits for unchirped pulses. The situation is much less favorable in the presence of frequency chirp. A typical value of the chirp parameter κ of a directly modulated semiconductor laser at 1.55 μm is -6, and β_2 is also negative so that monotone pulse broadening occurs. We leave it as an exercise to the reader (Problem 5.31) to calculate the chromatic dispersion limit with this value of κ and compare it to the dispersion limit for an unchirped pulse at a bit rate of 2.5 Gb/s.

If the chirp has the right sign however, it can interact with dispersion to cause pulse compression, as we saw in Section 2.4. Chirped RZ pulses can be used to take advantage of this effect.

Large Source Spectral Width

We derived (2.13) for the width of the output pulse by assuming a nearly monochromatic source, such as a DFB laser. In practice, this assumption is not satisfied for many sources such as MLM Fabry-Perot lasers. This formula must be modified to account for the finite spectral width of the optical source. Assume that the frequency spectrum of the source is given by

$$F(\omega) = B_0 W_0 e^{-(\omega-\omega_0)^2/2W_0^2}.$$

Thus the spectrum of the source has a Gaussian profile around the center frequency ω_0, and W_0 is a measure of the frequency spread or bandwidth of the pulse. The *rms spectral width* W^{rms}, which is defined in a fashion similar to that of the rms temporal width in (5.17), is given by $W^{\text{rms}} = W_0/\sqrt{2}$. As in the case of Gaussian pulses, the assumption of a Gaussian profile is chiefly for mathematical convenience; however, the results derived hold qualitatively for other source spectral profiles. From this spectrum, in the limit as $W_0 \to 0$, we obtain a monochromatic source at frequency ω_0. Equation (2.13) for the width of the output pulse is obtained under the assumption $W_0 << 1/T_0$. If this assumption does not hold, it must be modified to read

$$\frac{T_z}{T_0} = \sqrt{\left(1 + \frac{\kappa\beta_2 z}{T_0^2}\right)^2 + (1 + W_0^2 T_0^2)\left(\frac{\beta_2 z}{T_0^2}\right)^2}. \tag{5.21}$$

From this formula, we can derive the limitation imposed by chromatic dispersion on the bit rate B and the link length L. We have already examined this limitation for the case $W_0 \ll 1/T_0$. We now consider the case $W_0 \gg 1/T_0$ and again neglect chirp.

Consider a fiber of length L. With these assumptions, from (5.21), the width of the output pulse is given by

$$T_L = \sqrt{T_0^2 + (W_0\beta_2 L)^2}.$$

In this case, since the spectral width of the pulse is dominated by the spectral width of the source and not by the temporal width of the pulse ($W_0 \gg 1/T_0$), we can make T_0 much smaller than the bit period $1/B$ provided the condition $W_0 \gg 1/T_0$ is still satisfied. For such short input pulses, we can approximate T_L by

$$T_L = W_0|\beta_2|L.$$

Therefore, the condition $BT_L < \sqrt{2}$ translates to

$$BL\beta_2 W^{\text{rms}} < 1.$$

The key difference from the case of small source spectral width is that the bit rate B scales *linearly* with L. This is similar to the case of NRZ modulation using a source with a large spectral width, independent of the bit rate. As in the case of NRZ modulation, chromatic dispersion is much more of a problem when using sources with nonnegligible spectral widths.

In fact, the two conditions (for NRZ and RZ) are nearly the same. To see this, express the spectral width of the source in wavelength units rather than in angular frequency units. A spectral width of W in radial frequency units corresponds to a spectral width in wavelength units of $(\Delta\lambda) = -2\pi c W/\lambda^2$. Using this and the relation $D = -2\pi c \beta_2/\lambda^2$, the chromatic dispersion limit $BL\beta_2 W^{\text{rms}} < 1$ becomes

$$BL|D|(\Delta\lambda) < 1 \tag{5.22}$$

which is the same as (5.15) with $\epsilon = 1$.

As we have seen, the parameter β_2 is the key to group velocity or chromatic dispersion. For a given pulse, the magnitude of β_2 governs the extent of pulse broadening due to chromatic dispersion and determines the system limitations. β_2 can be minimized by appropriate design of the fiber as discussed in Section 2.4.2.

5.7.3 Dispersion Compensation

Dispersion management is a very important part of designing WDM transmission systems, since dispersion affects the penalties due to various types of fiber nonlinearities, as we will see in Section 5.8. We can use several techniques to reduce the impact of chromatic dispersion: (1) external modulation in conjunction with DFB lasers, (2) fiber with small chromatic dispersion, and (3) chromatic dispersion compensation. The first alternative is commonly used today in high-speed systems. New builds over the past few years have used nonzero-dispersion-shifted fibers (NZ-DSF) that have a small chromatic dispersion value in the C-band. Dispersion compensation can be employed when external modulation alone is not sufficient to reduce the chromatic dispersion penalty on the installed fiber type. We now discuss this option.

Along with the development of different fiber types, researchers have also developed various methods of compensating for chromatic dispersion. The two most popular methods use dispersion compensating fibers and chirped fiber Bragg gratings.

Dispersion Compensating Fibers

Special chromatic dispersion compensating fibers (DCFs) have been developed that provide negative chromatic dispersion in the 1550 nm wavelength range. For example, DCFs that can provide total chromatic dispersion of between −340 and

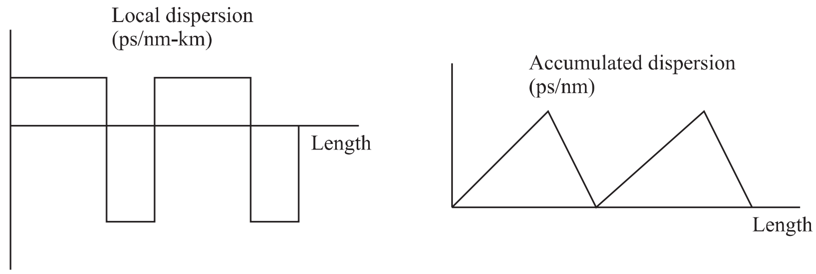

Figure 5.20 The chromatic dispersion map in a WDM link employing chromatic dispersion compensating fiber. (a) The (local) chromatic dispersion at each point along the fiber. (b) The accumulated chromatic dispersion from the beginning of the link up to each point along the fiber.

−1360 ps/nm are commercially available. An 80 km length of standard single-mode fiber has an accumulated or total chromatic dispersion, at 17 ps/nm-km, of $17 \times 80 = 1360$ ps/nm. Thus a DCF with −1360 ps/nm can compensate for this accumulated chromatic dispersion, to yield a net zero chromatic dispersion. Between amplifier spans is standard single-mode fiber, but at each amplifier location, dispersion compensating fiber having a negative chromatic dispersion is introduced. The *chromatic dispersion map*—the variation of accumulated chromatic dispersion with distance—of such a system is shown in Figure 5.20. Even though the chromatic dispersion of the fibers used is high, because of the alternating signs of the chromatic dispersion, this approach leads to a small value of the accumulated chromatic dispersion so that we need not worry about penalties induced by chromatic dispersion.

One disadvantage of this approach is the added loss introduced in the system by the DCF. For instance the −1360 ps/nm DCF has a loss of 9 dB. Thus a commonly used measure for evaluating a DCF is the *figure of merit* (FOM), which is defined as the ratio of the absolute amount of chromatic dispersion per unit wavelength to the loss introduced by the DCF. The FOM is measured in ps/nm-dB, and the higher the FOM, the more efficient the fiber is at compensating for chromatic dispersion. The FOM for the DCF in the preceding example is thus 150 ps/nm-dB. DCF with a chromatic dispersion of −100 ps/nm-km and a loss of 0.5 dB/km is now available. The FOM of this fiber is 200. There is intensive research under way to develop DCFs with higher FOMs.

The FOM as defined here does not fully characterize the efficiency of the DCF because it does not take into account the added nonlinearities introduced by the DCF due to its smaller effective area. A modified FOM that does take this into account has been proposed in [FTCV96].

The preceding discussion has focused on standard single-mode fiber that has a large chromatic dispersion in the C-band, about 17 ps/nm-km. In systems that use NZ-DSF, the chromatic dispersion accumulates much more slowly, since this fiber has a chromatic dispersion in the C-band of only 2–4 ps/nm-km. Thus these systems need a much smaller amount of chromatic dispersion compensating fiber. In many newly designed submarine systems, NZ-DSF with a small but negative chromatic dispersion is used. The use of negative chromatic dispersion fibers permits higher transmit powers to be used since modulation instability is not an issue (see Section 2.5.9). In this case, the accumulated chromatic dispersion is negative and can be compensated with standard single-mode fiber. This avoids the use of special chromatic dispersion compensating fibers, with their higher losses and susceptibility to nonlinear effects. The use of standard single-mode fiber for chromatic dispersion compensation also reduces the cabling loss due to bending. Terrestrial systems do not adopt this approach since the use of negative chromatic dispersion fiber precludes the system from being upgraded to use the L-band since the chromatic dispersion zero for these fibers lies in the L-band. This is not an issue for submarine systems since these systems are not upgradable once they have been deployed.

Chirped Fiber Bragg Gratings

The fiber Bragg grating that we studied in Section 3.3.4 is a versatile device that can be used to compensate for chromatic dispersion. Such a device is shown in Figure 5.21. The grating itself is linearly chirped in that the period of the grating varies linearly with position, as shown in Figure 5.21. This makes the grating reflect different wavelengths (or frequencies) at different points along its length. Effectively, a chirped Bragg grating introduces different delays at different frequencies.

In a regular fiber, chromatic dispersion introduces larger delays for the lower-frequency components in a pulse. To compensate for this effect, we can design chirped gratings that do exactly the opposite—namely, introduce larger delays for the higher-frequency components, in other words, compress the pulses. The delay as a function of frequency is plotted in Figure 5.21 for a sample grating.

Ideally, we want a grating that introduces a large amount of chromatic dispersion over a wide bandwidth so that it can compensate for the fiber chromatic dispersion over a large length as well as a wide range of wavelengths. In practice, the total length of the grating is limited by the size of the phase masks available. Until recently, this length used to be a few tens of centimeters. With a 10-cm-long grating, the maximum delay that can be introduced is 1 ns. This delay corresponds to the product of the chromatic dispersion introduced by the grating and the bandwidth over which it

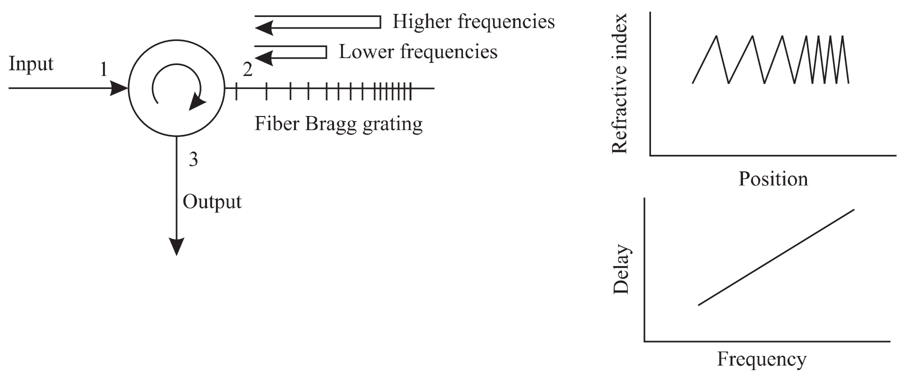

Figure 5.21 Chirped fiber Bragg grating for chromatic dispersion compensation.

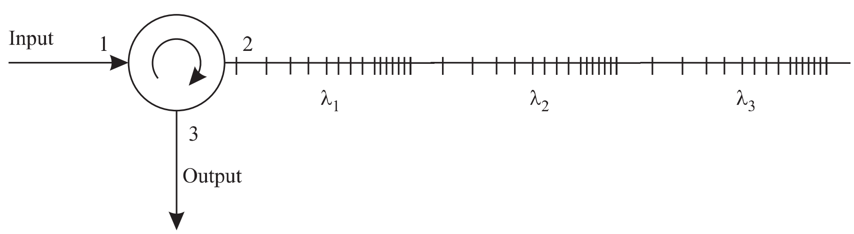

Figure 5.22 Chirped fiber Bragg gratings for compensating three wavelengths in a WDM system.

is introduced. With such a grating, we introduce large chromatic dispersion over a small bandwidth, for example, 1000 ps/nm over a 1 nm bandwidth, or small chromatic dispersion over a wide bandwidth, for example, 100 ps/nm over a 10 nm bandwidth. Note that 100 km of standard single-mode fiber causes a total chromatic dispersion of 1700 ps/nm. When such chirped gratings are used to compensate for a few hundred kilometers of fiber chromatic dispersion, they must be very narrow band; in other words, we would need to use a different grating for each wavelength, as shown in Figure 5.22.

Chirped gratings are therefore ideally suited to compensate for individual wavelengths rather than multiple wavelengths. In contrast, DCF is better suited to compensate over a wide range of wavelengths. However, compared to chirped gratings, DCF introduces higher loss and additional penalties because of increased nonlinearities.

Recently, very long gratings, about 2 m in length, have been demonstrated [Bre01]. These gratings have been shown to compensate for the accumulated chromatic dispersion, over the entire C-band, after transmission over 40 km of standard single-mode fiber. Such a grating may prove to be a strong competitor to DCF.

Dispersion Slope Compensation

One problem with WDM systems is that since the chromatic dispersion varies for each channel (due to the nonzero slope of the chromatic dispersion profile), it may not be possible to compensate for the entire system using a common chromatic dispersion compensating fiber. A typical spread of the total chromatic dispersion, before and after compensation with DCF, across several WDM channels, is shown in Figure 5.23. This spread can be compensated by another stage of *chromatic dispersion slope compensation* where an appropriate length of fiber whose chromatic dispersion slope is opposite to that of the residual chromatic dispersion is used.

As we remarked in Section 2.5.9, it is difficult to fabricate positive chromatic dispersion fiber with negative slope (today), so that this technique can only be used for systems employing positive dispersion, positive slope fiber for transmission (and negative dispersion, negative slope fiber for dispersion, and dispersion slope, compensation). Thus, in submarine systems that use negative dispersion, positive slope fiber, dispersion slope compensation using dispersion compensating fiber is not possible. Moreover, if such systems employ large effective area fiber to mitigate nonlinear effects, the spread in chromatic dispersion slopes is enhanced, since large effective area fibers have larger dispersion slopes. One way to minimize the chromatic dispersion slope spread is to use a *hybrid fiber design*. In such a design, each span of, say, 50 km uses two kinds of fiber: large effective area fiber (with a consequent large dispersion slope) in the first half of the span and a reduced slope fiber in the second half. Since nonlinear effects are significant only at the high power levels that occur in the first half of the span, the use of large effective area fiber in this half mitigates these effects, as effectively as using large effective area fiber for the whole span. The use of reduced slope fiber in the second half reduces (but does not eliminate) the overall spread in dispersion slope across channels (compared to using large effective area fiber in the whole span).

A second method of dispersion slope compensation is to provide the appropriate chromatic dispersion compensation for each channel separately at the receiver after the channels are demultiplexed. Although individual channels can be compensated using appropriately different lengths of DCF, chirped fiber gratings (see Section 5.7.3) are commonly used to compensate individual channels since they are much more compact.

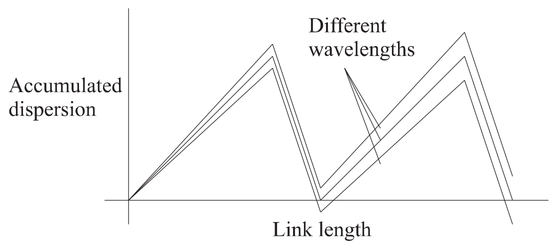

Figure 5.23 Variation of total chromatic dispersion in a WDM system across different channels, after chromatic dispersion compensation with a DCF.

A third method of overcoming the dispersion slope problem is termed *mid-span spectral inversion* (MSSI). Roughly speaking, in this method, the spectrum of the pulse is inverted in the middle of the span; that is, the shorter and longer wavelengths of the pulse are interchanged. Recall that a pulse that is nominally at some frequency has a finite (nonzero) spectral width. Here we are referring to the different spectral components, or wavelengths, of a single pulse, and not the different wavelength channels in the system. This process is called *phase conjugation,* and it reverses the sign of the chromatic dispersion in the two halves of the span. Even if the chromatic dispersion values of different channels are equal, the chromatic dispersion in the two halves of the span cancels for each channel. Currently, the two other techniques, namely, chromatic dispersion compensating fiber and chirped fiber gratings, appear to be more suitable for commercial deployment.

5.7.4 Polarization-Mode Dispersion (PMD)

The origin of PMD lies in the fact that different polarizations travel with different group velocities because of the ellipticity of the fiber core; we discussed this in Section 2.3.3. Moreover, the distribution of signal energy over the different state of polarizations (SOPs) changes slowly with time, for example, because of changes in the ambient temperature. This causes the PMD penalty to vary with time as well. In addition to the fiber itself, PMD can arise from individual components used in the network.

The time-averaged differential time delay between the two orthogonal SOPs on a link is known to obey the relation [KK97a, Chapter 6]

$$\langle \Delta\tau \rangle = D_{\text{PMD}}\sqrt{L},$$

where $\langle \Delta\tau \rangle$ is called the differential group delay (DGD), L is the link length, and D_{PMD} is the fiber PMD parameter, measured in ps/$\sqrt{\text{km}}$. The PMD for typical fiber lies between 0.5 and 2 ps/$\sqrt{\text{km}}$. However, carefully constructed new links can have PMD as low as 0.1 ps/$\sqrt{\text{km}}$.

In reality, the SOPs vary slowly with time, and the actual DGD $\Delta\tau$ is a random variable. It is commonly assumed to have a Maxwellian probability density function (see Appendix H). This means that the square of the DGD is modeled by a more familiar distribution—the exponential distribution. The larger the DGD, the larger is the power penalty due to PMD. Thus, the power penalty due to PMD is also time varying, and it turns out that it is proportional to $\Delta\tau^2$ and thus obeys an exponential distribution (see Problem 5.23). If the power penalty due to PMD is large, it is termed a *PMD outage* and the link has effectively failed. For a DGD of $0.3T$, where T is the bit duration, the power penalty is approximately 0.5 dB for a receiver limited by thermal noise and 1 dB for a receiver with signal-dependent noise (ITU standard G.691).

Using the Maxwellian distribution, the probability that the actual delay will be greater than three times the average delay is about 4×10^{-5} (see Appendix H). Given our earlier reasoning, this means that in order to restrict the PMD outage probability (PMD $\geq$ 1 dB) to 4×10^{-5}, we must have the average DGD to be less than $0.1T$; that is,

$$\langle \Delta\tau \rangle = D_{\text{PMD}}\sqrt{L} < 0.1T. \tag{5.23}$$

This limit is plotted in Figure 5.24. Observe that for a bad fiber with PMD of 2 ps/$\sqrt{\text{km}}$, the limit is only 25 km. This is an extreme case, but it points out that PMD can impose a significant limitation.

Note that we have not said anything about the distribution of the length of time for which there is a PMD outage. In the above example, the DGD may exceed three times the average delay, and we may have one PMD outage with an average duration of one day once every 70 years, or one with an average duration of one minute every 17 days. This depends on the fiber cable in question, and typical outages last for a few minutes. Thus an outage probability of 4×10^{-5} can also be interpreted as a cumulative outage of about 20 minutes per year.

The limitations due to intermodal dispersion, chromatic dispersion, and PMD are compared in Figure 5.25.

PMD gives rise to intersymbol interference (ISI) due to pulse spreading, just as all other forms of dispersion. The traditional (electronic) technique for overcoming ISI in digital systems is *equalization,* discussed in Section 4.4.9. Equalization to compensate for PMD can be carried out in the electronic domain and is discussed in [WK92, YS06]. However, electronic equalization becomes more difficult at very high bit rates of 40 Gb/s and beyond. At such high bit rates, optical PMD compensation must be used.

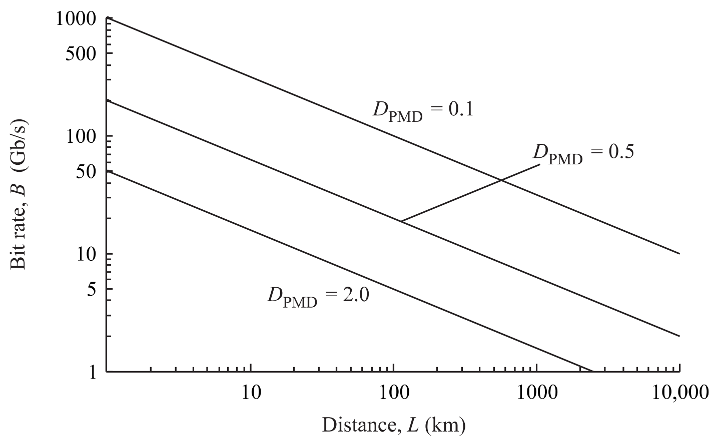

Figure 5.24 Limitations on the simultaneously achievable bit rates and distances imposed by PMD.

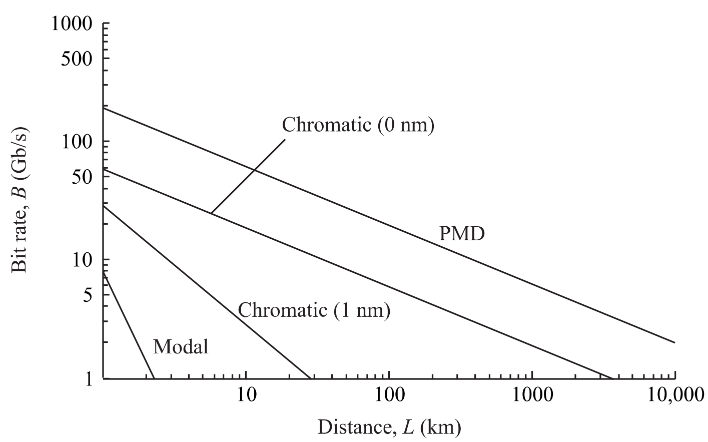

Figure 5.25 Limitations on the simultaneously achievable bit rates and distances imposed by intermodal dispersion, chromatic dispersion with a source spectral width of 1 nm, chromatic dispersion with spectral width proportional to the modulation bandwidth, and PMD with $D_{PMD} = 0.5$. NRZ modulation transmission over standard single-mode fiber with a chromatic dispersion value of 17 ps/nm-km is assumed.

To understand how PMD can be compensated optically, recall that PMD arises due to the fiber birefringence and is illustrated in Figure 2.7. The transmitted pulse consists of a “fast” and a “slow” polarization component. The principle of PMD compensation is to split the received signal into its fast and slow polarization components and to delay the fast component so that the DGD between the two components is compensated. Since the DGD varies in time, the delay that must be introduced in the fast component to compensate for PMD must be estimated in real time from the properties of the link.

The PMD effect we have discussed so far must strictly be called first-order polarization-mode dispersion. First-order PMD is a consequence of the fact that the two orthogonal polarization modes in optical fiber travel at slightly different speeds, which leads to a differential time delay between these two modes. However, this differential time delay itself is frequency dependent and varies over the bandwidth of the transmitted pulse. This effect is called second-order PMD. Second-order PMD is an effect that is similar to chromatic dispersion and thus can lead to pulse spreading.

PMD also depends on whether RZ or NRZ modulation is used; the discussion so far pertains to NRZ modulation. For RZ modulation, the use of short pulses enables more PMD to be tolerated since the output pulse has more room to spread—similar to the case of chromatic dispersion. However, second-order PMD depends on the spectral width of the pulse; narrower pulses have larger spectral widths. This is similar to the case of chromatic dispersion (Section 5.7.2). Again, as in the case of chromatic dispersion, there is an optimum input pulse width for RZ modulation that minimizes the output pulse width [SKA00, SKA01].

In addition to PMD, some other polarization-dependent effects influence system performance. One of these effects arises from the fact that many components have a polarization-dependent loss (PDL); that is, the loss through the component depends on the state of polarization. These losses accumulate in a system with many components in the transmission path. Again, since the state of polarization fluctuates with time, the signal-to-noise ratio at the end of the path will also fluctuate with time, and careful attention needs to be paid to maintain the total PDL on the path to within acceptable limits. An example is a simple angled-facet connector used in some systems to reduce reflections. This connector can have a PDL of about 0.1 dB, but hundreds of such connectors can be present in the transmission path.

5.8 Fiber Nonlinearities

As long as the optical power within an optical fiber is small, the fiber can be treated as a linear medium; that is, the loss and refractive index of the fiber are independent of the signal power. However, when power levels get fairly high in the system, we have to worry about the impact of nonlinear effects, which arise because, in reality, both the loss (gain) and refractive index depend on the optical power in the fiber. Nonlinearities can place significant limitations on high-speed systems as well as WDM systems.

As discussed in Chapter 2, nonlinearities can be classified into two categories. The first occurs because of scattering effects in the fiber medium due to the interaction of light waves with phonons (molecular vibrations) in the silica medium. The two main effects in this category are stimulated Brillouin scattering (SBS) and stimulated Raman scattering (SRS). The second set of effects occurs because of the dependence of refractive index on the optical power. This category includes four-wave mixing (FWM), self-phase modulation (SPM), and cross-phase modulation (CPM). In Chapter 2, we looked at the origins of all these effects. Here we will examine the limitations that all these nonlinearities place on system designers.

Except for SPM and CPM, all these effects provide gains to some channels at the expense of depleting power from other channels. SPM and CPM, on the other hand, affect only the phase of signals and can cause spectral broadening, which in turn, leads to increased chromatic dispersion penalties.

5.8.1 Effective Length in Amplified Systems

We discussed the notion of the effective length of a fiber span in Section 2.5.1. In systems with optical amplifiers, the signal gets amplified at each amplifier stage without resetting the effects due to nonlinearities from the previous span. Thus the effective length in such a system is the sum of the effective lengths of each span. In a link of length L with amplifiers spaced l km apart, the effective length is approximately given by

$$L_e = \frac{1 - e^{-\alpha l}}{\alpha} \frac{L}{l}. \tag{5.24}$$

Figure 5.26 shows the effective length plotted against the actual length of the transmission link for unamplified and amplified systems. The figure indicates that, in order to reduce the effective length, it is better to have fewer amplifiers spaced further apart. However, what matters in terms of the system effects of nonlinearities is not just the effective length; it is the product of the launched power P and the effective length L_e. Figure 5.6 showed how P varies with the amplifier spacing l. Now we are interested in finding out how PL_e grows with the amplifier spacing l. This is shown in Figure 5.27. The figure shows that the effect of nonlinearities can be reduced by reducing the amplifier spacing. Although this may make it easier

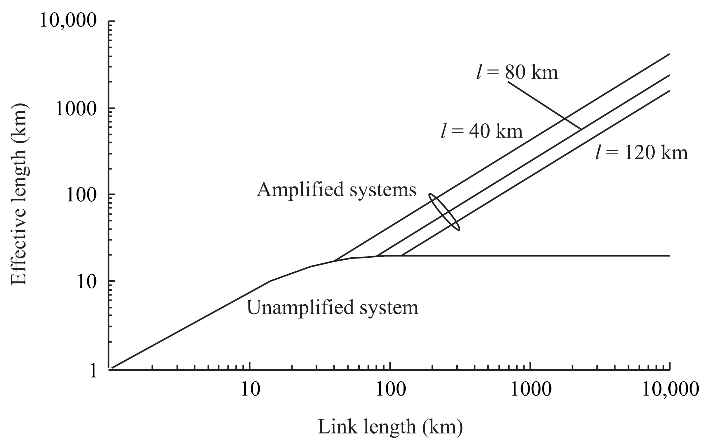

Figure 5.26 Effective transmission length as a function of link length, l.

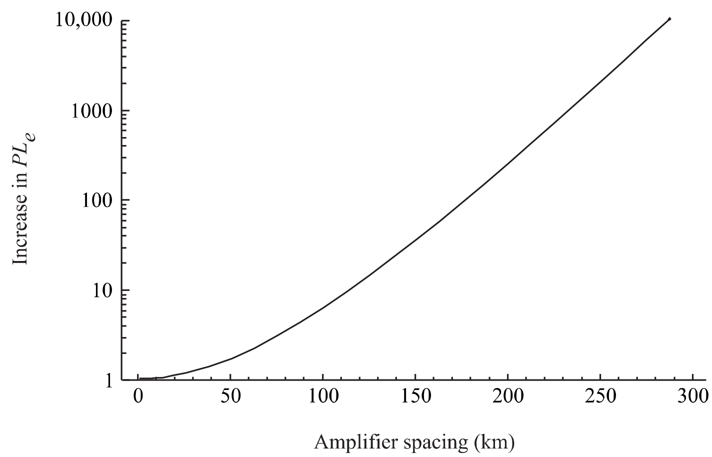

Figure 5.27 Relative value of PL_e versus amplifier spacing. The ordinate is the value relative to an amplifier spacing of 1 km. $\alpha = 0.22$ dB/km.

to design the amplifiers (they need lower gain), we will also need more amplifiers, resulting in an increase in system cost.

The effect of a scattering nonlinearity depends on PL_e and thus increases with an increase in the input power and the link length. The longer the link, the greater is the amount of power that is coupled out from the signal (pump) into the Stokes wave. For a given link length, an approximate measure of the power level at which the effect of a nonlinearity starts becoming significant is the *threshold power*. For a given fiber length, the threshold power of a scattering nonlinearity is defined as the incident optical power per channel into the fiber at which the pump and Stokes powers at the fiber output are equal. In amplified systems, the threshold power is reduced because of the increase in the effective length. This makes amplified systems more susceptible to impairments due to nonlinearities.

5.8.2 Stimulated Brillouin Scattering

The calculation of the threshold power for SBS P_{th} is quite involved, and we simply state the following approximation for it from [Smi72]:

$$P_{\text{th}} \approx \frac{21 b A_e}{g_B L_e}.$$

Here, A_e and L_e are the effective area and length of the fiber, respectively (see Section 2.5.1), $g_B \approx 4 \times 10^{-11}$ m/W is called the Brillouin gain coefficient, and the value of b lies between 1 and 2 depending on the relative polarizations of the pump and Stokes waves. Assuming the worst-case value of $b = 1$, $A_e = 50\ \mu\text{m}^2$, and $L_e = 20$ km, we get $P_{\text{th}} = 1.3$ mW. Since this is a low value, some care must be taken in the design of optical communication systems to reduce the SBS penalty.

The preceding expression assumes that the pump signal has a very narrow spectral width and lies within the narrow 20 MHz gain bandwidth of SBS. The threshold power is considerably increased if the signal has a broad spectral width, and thus much of the pump power lies outside the 20 MHz gain bandwidth of SBS. An approximate expression that incorporates this effect is given by

$$P_{\text{th}} \approx \frac{21 b A_e}{g_B L_e} \left(1 + \frac{\Delta f_{\text{source}}}{\Delta f_B}\right),$$

where Δf_{source} is the spectral width of the source. With $\Delta f_{\text{source}} = 200$ MHz, and still assuming $b = 1$, the SBS threshold increases to $P_{\text{th}} = 14.4$ mW.

The SBS penalty can be reduced in several ways:

1. Keep the power per channel to much below the SBS threshold. The trade-off is that in a long-haul system, we may have to reduce the amplifier spacing.
2. Since the gain bandwidth of SBS is very small, its effect can be decreased by increasing the spectral width of the source. This can be done by directly modulating the laser, which causes the spectral width to increase because of chirp. This may cause a significant chromatic dispersion penalty. The chromatic dispersion

penalty can, however, be reduced by suitable chromatic dispersion management, as we will see later. Another approach is to dither the laser slightly in frequency, say, at 200 MHz, which does not cause as high a penalty because of chromatic dispersion but increases the SBS threshold power by an order of magnitude, as we saw earlier. This approach is commonly employed in high-bit-rate systems transmitting at high powers. Regardless of the bit rate, the use of an external modulator along with a narrow spectral width source increases the SBS threshold by only a small factor (between 2 and 4) for amplitude-modulated systems. This is because a good fraction of the power is still contained in the optical carrier for such systems.

3. Use phase modulation schemes rather than amplitude modulation schemes. This reduces the power present in the optical carrier, thus reducing the SBS penalty. In this case, the spectral width of the source can be taken to be proportional to the bit rate. However, this may not be a practical option in most systems.

5.8.3 *Stimulated Raman Scattering*

We saw in Section 2.5 that if two or more signals at different wavelengths are injected into a fiber, SRS causes power to be transferred from the shorter-wavelength channels to the longer-wavelength channels (see Figure 2.16). Channels up to 150 THz (125 nm) apart are coupled due to SRS, with the peak coupling occurring at a separation of about 13 THz. Coupling occurs for both copropagating and counter-propagating waves.

Coupling occurs between two channels only if both channels are sending 1 bits (that is, power is present in both channels). Thus the SRS penalty is reduced when chromatic dispersion is present because the signals in the different channels travel at different velocities, reducing the probability of overlap between pulses at different wavelengths at any point in the fiber. This is the same *pulse walk-off* phenomenon that we discussed in the case of CPM in Section 2.5.7. Typically, chromatic dispersion reduces the SRS effect by a factor of 2.

To calculate the effect of SRS in a multichannel system, following [Chr84], we approximate the Raman gain shape as a triangle, where the Raman gain coefficient as a function of wavelength spacing $\Delta\lambda$ is given by

$$g(\Delta\lambda) = \begin{cases} g_R \frac{\Delta\lambda}{\Delta\lambda_c}, & \text{if } 0 \leq \Delta\lambda \leq \Delta\lambda_c, \\ 0 & \text{otherwise.} \end{cases}$$

Here $\Delta\lambda_c = 125$ nm, and $g_R \approx 6 \times 10^{-14}$ m/W (at 1.55 μm) is the peak Raman gain coefficient.

Consider a system with W equally spaced channels $0, 1, \ldots, W-1$, with $\Delta\lambda_s$ denoting the channel spacing. Assume that all the channels fall within the Raman gain bandwidth; that is, the system bandwidth $\Lambda = (W-1)\Delta\lambda_s \leq \Delta\lambda_c$. This is the case of practical interest given that the Raman gain bandwidth is 125 nm and the channels within a WDM system must usually be spaced within a 30 nm band dictated by the bandwidth of optical amplifiers. The worst affected channel is the channel corresponding to the lowest wavelength, channel 0, when there is a 1 bit in all the channels. Assume that the transmitted power is the same on all channels. Assume further that there is no interaction between the other channels, and the powers of the other channels remain the same (this approximation yields very small estimation errors). Assume also that the polarizations are scrambled. This is the case in practical systems. In systems that use polarization-maintaining fiber, the Raman interaction is enhanced, and the equation that follows does not have the factor of 2 in the denominator. The fraction of the power coupled from the worst affected channel, channel 0, to channel i is given approximately by [Buc95]

$$P_o(i) = g_R \frac{i\Delta\lambda_s}{\Delta\lambda_c}\frac{PL_e}{2A_e}.$$

This expression can be derived starting from the coupled wave equations for SRS that are similar in form to (2.14) and (2.15); see [Buc95] for details and [Zir98] for an alternative derivation with fewer assumptions. So the fraction of the power coupled out of channel 0 to all the other channels is

$$P_o = \sum_{i=1}^{W-1} P_o(i) = \frac{g_R \Delta\lambda_s P L_e}{2\Delta\lambda_c A_e}\frac{W(W-1)}{2}. \tag{5.25}$$

The power penalty for this channel is then

$$-10\log(1-P_o).$$

In order to keep the penalty below 0.5 dB, we must have $P_o < 0.1$, or, from (5.25),

$$WP(W-1)\Delta\lambda_s L_e < 40{,}000 \text{ mW-nm-km}.$$

Observe that the total system bandwidth is $\Lambda = (W-1)\Delta\lambda_s$ and the total transmitted power is $P_{\text{tot}} = WP$. Thus the result can be restated as

$$P_{\text{tot}}\Lambda L_e < 40{,}000 \text{ mW-nm-km}.$$

The preceding formula was derived assuming that no chromatic dispersion is present in the system. With chromatic dispersion present, the right-hand side can be relaxed to approximately 80,000 mW-nm-km.

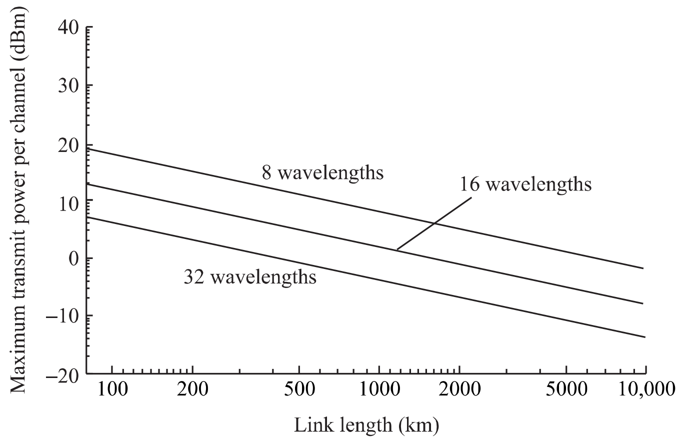

Figure 5.28 Limitation on the maximum transmit power per channel imposed by stimulated Raman scattering. The channel spacing is assumed to be 0.8 nm, and amplifiers are assumed to be spaced 80 km apart.

If the channel spacing is fixed, the power that can be launched decreases with W as $1/W^2$. For example, in a 32-wavelength system with channels spaced 0.8 nm (100 GHz) apart, and $L_e = 20$ km, $P \leq 2.5$ mW. Figure 5.28 plots the maximum allowed transmit power per channel as a function of the link length.The limit plotted here corresponds to $P_{\text{tot}}\Lambda L_e < 80{,}000$ mW-nm-km.

Although SRS is not a significant problem in systems with a small number of channels due to the relatively high threshold power, it can pose a serious problem in systems with a large number of wavelengths. To alleviate the effects of SRS, we can (1) keep the channels spaced as closely together as possible and/or (2) keep the power levels below the threshold, which will require us to reduce the distance between amplifiers.

5.8.4 Four-Wave Mixing

We saw in Section 2.5 that the nonlinear polarization causes three signals at frequencies ω_i, ω_j, and ω_k to interact to produce signals at frequencies $\omega_i \pm \omega_j \pm \omega_k$. Among these signals, the most troublesome one is the signal corresponding to

$$\omega_{ijk} = \omega_i + \omega_j - \omega_k, \; i \neq k, \; j \neq k. \tag{5.26}$$

Depending on the individual frequencies, this beat signal may lie on or very close to one of the individual channels in frequency, resulting in significant crosstalk to